



BIG DATA And Advanced Analytics 2023

1

IX

международная
научно-практическая
конференция

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

BIG DATA И АНАЛИЗ ВЫСОКОГО УРОВНЯ

BIG DATA AND ADVANCED ANALYTICS

СБОРНИК НАУЧНЫХ СТАТЕЙ
IX МЕЖДУНАРОДНОЙ НАУЧНО-ПРАКТИЧЕСКОЙ КОНФЕРЕНЦИИ

В двух частях

Часть 1

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

Минск БГУИР 2023

УДК 004.6(082)

ББК 32.973.3

Б59

Редакционная коллегия:

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

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

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

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

Рецензенты:

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

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

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

BIG DATA и анализ высокого уровня = **BIG DATA and Advanced Analytics** : сб. науч. Б59 ст. IX Междунар. науч.-практ. конф. В 2 ч. Ч. 1 (Республика Беларусь, Минск, 17–18 мая 2023 года) / редкол. : В. А. Богуш [и др.]. – Минск : БГУИР, 2023. – 484 с.

ISBN 978-985-543-712-4 (ч. 1).

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

УДК 004.6(082)

ББК 32.973.3

ISBN 978-985-543-712-4 (ч. 1)

ISBN 978-985-543-711-7

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

ОРГАНИЗАЦИОННЫЙ КОМИТЕТ КОНФЕРЕНЦИИ



Председатель, Богуш В.А.

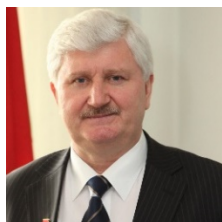
Ректор Белорусского государственного университета информатики и радиоэлектроники, доктор физико-математических наук, профессор



Заместитель председателя, Дик С.К.

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

ЧЛЕНЫ ОРГАНИЗАЦИОННОГО КОМИТЕТА



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



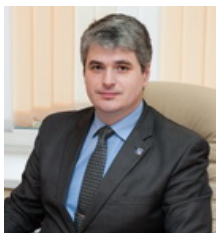
Болдырев А.С., директор института радиотехнических систем и управления Южного федерального университета (г. Таганрог), кандидат физико-математических наук, доцент, Российская Федерация



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



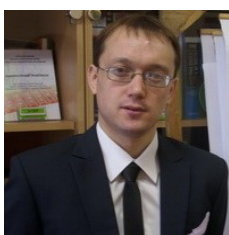
Казак Т.В., заведующий кафедрой инженерной психологии и эргономики БГУИР, доктор психологических наук, профессор, Республика Беларусь



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



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



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



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



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



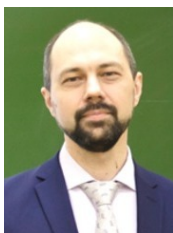
Кругликов С. В., генеральный директор Объединенного института проблем информатики Национальной академии наук Беларуси



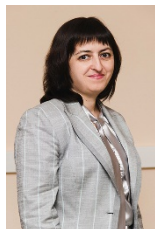
Шнейдеров Е.Н., проректор по учебной работе Белорусского государственного университета информатики и радиоэлектроники, кандидат технических наук, доцент, Республика Беларусь



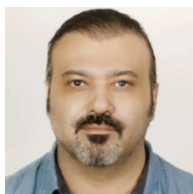
Стемплицкий В.Р., проректор по научной работе БГУИР, начальник научно-исследовательской части, научный руководитель НИИ "Компьютерное проектирование микро- и нанoeлектронных систем" кандидат технических наук, доцент, Республика Беларусь



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



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



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

ОРГАНИЗАТОРЫ КОНФЕРЕНЦИИ



Министерство образования Республики
Беларусь



Учреждение образования
«Белорусский государственный университет
информатики и радиоэлектроники»



Национальная академия наук Беларуси



Объединенный институт проблем
информатики



Ташкентский университет информационных
технологий, Узбекистан



Южный федеральный университет, Российская
Федерация



Северо-Кавказский федеральный университет,
Российская Федерация



Евразийский национальный университет им. Л.Н.
Гумилева



Технологический университет Шарифа, Иран

**СПОНСОР
КОНФЕРЕНЦИИ**



Облачный хостинг hoster.by

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

ОГЛАВЛЕНИЕ

S.S. Beknazarova Basics of digital processing of images.....	12
D.R. Rahel Features of searching for associations in large volumes of marketing data.....	18
Z.Y.Chen, Z.X. Yang, H.Li Human physical activity recognition algorithm based on smartphone data and long short time memory neural network.....	21
Gaowei Qie Heart rate estimation from photoplethysmogram and acceleration smartphone data based on convolutional neuralnetwork and long short time memor.....	29
S.K. Kurbanov Creating of integrated spatial model for 3D GIS	35
O.N. Karasik, A.A. Prihozhy Profiling of energy consumption by algorithms of shortest paths search in large dense graphs.....	44
B.A.Al-Nami Converting images from one extension to another without losing its quality, problems and solutions	51
N.A.Naim Modern trends in bigdata and datascience.....	62
A.I. Tukhtasinov, V.S.Sodikov Fractal modeling of big data.....	65
M.Ya Mansurova Some issues of Big Data application in modeling business processes of e-business systems	68
N.A.Naim Application of big data technology	79
V.S.Sodikov, A.I. Tukhtasinov Analyzing the large amount of data.....	82
И.А.Оганезов, Н.В. Щербина, А.В. Буга Оценка эффективности автоматизации установки для охлаждения молока с использованием частотного управления	86
М.В. Тумилович, Л.П. Пилинович, А.Г. Кравцов Моделирование влияния характеристик фильтруемой среды на закономерности тангенциальной фильтрации в пористых градиентных материалах	93
И.В. Марахина Направления трансформации web-копирайтинга в условиях развития искусственного интеллекта	100
Н.В. Щербина Математическая модель прогнозирования показателей релаксации работников локомотивных бригад.....	103
И.Ф. Киринович, К. Дадаев Использование базовых принципов эргономики в web-разработке	114
В.В. Верняховская, О.М. Раптунович Построение индивидуального пути обучения IT-специалиста и сетевое взаимодействие вузов в процессе обучения.....	117
И.Ф. Киринович, К. Дадаев Особенности улучшения пользовательского интерфейса мобильного приложения.....	121
Н.В. Лещёв, О.Н. Шкор Будущее маркетинга: автоматизация взаимодействия с клиентами с помощью интернета.....	124
В.В. Шаталова, Т.В. Казак Новые подходы в подготовке HR-специалистов в условиях цифровой трансформации	128
П.Ю. Бранцевич Сравнительный анализ электроэнцефалограмм.....	132
О.Н. Шкор Инновационные тренды в сфере образования и роль преподавателя	145

А.В. Гордиевич, Л.Ф. Васковская, П.В. Камлач Исследование влияния применения фильтров на сигнал электрокардиограммы	149
Т.В.Казак, А.В. Свороб, А.Н. Василькова Как большие данные и искусственный интеллект могут принести результаты.....	156
Д.Ю. Маркевич, О.Н. Шкор Потенциальное влияние WEB3 на B2B рынок.....	161
Р.В.Козарь, А.А. Навроцкий, Н.С. Конойко Альтернативные методы классификации при решении задач распознавания эндоскопических снимков	164
И.П. Сидорчук, А.А. Охрименко, Е.Г. Крысь Государственно-правовое регулирование процессов цифровизации в Беларуси.....	170
Д.И Черемисинов О методике преподавания программирования на старших курсах вузов.....	180
И.Д. Марковская, О.Н. Шкор Использование cgm систем как инструмента работы с Big Data в маркетинге	188
В.П. Корячко, В.И. Орешков Снижение влияния утечки данных на точность моделей в машинном обучении.....	192
Т.В. Казак, А.А. Войтович, А.Н. Василькова Применение блокчейна и больших данных в сфере образования	199
В.В. Савченко, В.А. Дубовский Разработка подхода и обобщенного алгоритма адаптивного взаимодействия систем верхнего уровня высокоавтоматизированных транспортных средств.....	204
В.А. Стельмах, О.Н. Шкор Интерактивное web 3.0 приложение как инструмент конкурентоспособности	212
А.В.Гордиевич, А.Д. Скрипка, П.В. Камлач Автоматизированная система контроля «диспетчер здоровья»	215
Н.А.Кириенко Полиномиальная реализация систем булевых функций в задачах логического проектирования.....	227
М.В.Романюк, А.Н.Марков Организация и управление зоной Wi-Fi в учреждении образования для работы с BIGDATA	235
Н.И. Потапенко, В.А. Буд-Гусаим, А.Н. Василькова Big Data и машинное обучение	239
В.Н.Шведко, Т.В.Казак Организация психопрофилактической работы в войсках	244
Л.Р.Коркин, Саевич К.Ф. Программа для подсчета микроорганизмов	247
Д.А. Мирзаев, И.Ж. Бозорова, Ф.Х. Мамасидикова Перспективы развития IT-аудита системы бухгалтерского учёта за счёт внедрения ИКТ.....	252
М.Ю.Чиркова, О.Н.Шкор Big data и advanced analytics в телемедицине	258
Ю.Е. Седельников, Д.А. Веденькин, А.Ф. Гильфанова Применение метода Ильина-Морозова и в задачах формирования мультисфокусированного электромагнитного излучения	261
Н.И. Потапенко, К.Ю. Назарук, А.Н. Василькова Использование искусственного интеллекта для генерации музыки	267
Ю.И. Голубович, С.Н. Нестеренков, Д.В. Басак Применение больших данных в работе морских портов и терминалов.....	272

А.И. Калько Система идентификации рукописных текстов как объект защиты и обработки массового объема документов	276
В.Л.Ланин, И.Б. Петухов Моделирование резонансных явлений при ультразвуковой микросварке проволочных выводов в корпусах типа ТО-5	285
А.А. Логинова, М.Д. Попов, А.Р. Денисов Особенности применения алгоритмов процессной аналитики (process mining) для анализа поведения студентов	290
Н.А.Шиманский, А.В.Баглов, Л.С.Хорошко Автоматизация обработки результатов исследования структуры и свойств наноматериалов.....	296
В.В. Синицына Технологии больших данных в алгоритмах рекolorизации изображений для людей с аномальной трихромазией	301
Л.И. Архипова, Л.Ф. Медведева Martech как тренд цифровой трансформации бизнеса	312
С.А. Затонский Инструментарий алгоритмизации открытых и Big Data – данных в целях развития международного бизнеса компаний.....	320
А.А.Бакурина, О.Н. Шкор Big Data и advanced analytics в медицине	329
В.П. Корячко, В.И. Орешков Выбор числа кластеров в задачах кластеризации с использованием метода силуэтов.....	333
Л.Д.Черемисинова Минимизация графа достижимых состояний параллельного автомата ...	341
Е.И. Баяк, С.Н. Нестеренков, Д.А. Жалейко Применение Big Data для анализа социальных сетей.....	348
А.И.Калько, О.Д.Хадарович Приложение для демонстрации данных в виде виртуальных 3D-Объектов с дополненной реальностью.....	351
Ж.Б. Балтаев, С.Д. Арзикулов, М.А. Миржамолова Методы оценки риска нарушений информационной безопасности при облачных вычислениях в финансовых и банковских системах	359
И.П. Логинова Анализ быстродействия при реализации параллельных вычислений для решения задач оптимизации в системе логического проектирования	365
С.Н.Нестеренков, Д.В. Басак, В.В.Куц Применение Apache Spark для обработки больших данных в мобильных системах	379
Е.А.Сальникова Облачные технологии и технология больших данных в образовании	382
Л.И. Архипова, С.В. Наркевич Business development в реализации проектов цифрового маркетинга.....	386
В.Л.Ланин, А.Д. Хацкевич моделирование индукционного нагрева шариков припоя для Flip-Chip монтажа в Comsol Multiphysics.....	391
Е.С. Боженко, Е.Ю. Полякова Теория и практика маркетинга больших данных в организациях высшего образования	396
М.В. Спринджук Анализ биоинформации для мониторинга опасной коронавирусной инфекции	407
Н.И. Куприянов, П.В.Камлач, А.В. Слижёва Моделирование имитационной модели аппарата искусственной вентиляции легких.....	415
Н.В. Новицкая, С.В.Матийко Управление планированием и прогнозированием финансов в автоматизированной программе 1С ERP с использованием Big Data.....	424

А.И.Лаппо, В.Л.Ланин Моделирование лазерного нагрева при движении лазера по заданной траектории	428
Г.А. Пискун, В.Ф.Алексеев, Т.М. Воронко Актуальность использования языка javascript для разработки моделей машинного обучения	432
А.В. Санец, О.С. Медведев, Т.Ю. Шлыкова Нейросеть как фактор оптимизации тайм-менеджмента специалиста при решении текущих производственных задач.....	439
Н.В. Щербина Методика выработки навыка релаксации для работников локомотивных бригад	443
Д.А.Фролова, П.Д. Булахова Внедрение CRM И оптимизация бизнес-процессов в сфере туризма	450
О.Н. Шкор, Э.В. Котович Использование Big Data and advanced analytics, а также Data science для оптимизации производственных и бизнес решений в сфере розничной торговли	454
К.И. Котельников, В.Е. Буюков, А.Н. Марков Применение Big Data в сфере образования	459
Г.А.Пискун, В.Ф.Алексеев, Т.М. Воронко Обзор программной платформы APACHE HADOOP для обработки и хранения больших данных.....	465
В.Ф.Алексеев, Г.А.Пискун, Д.В. Лихачевский, И.Н. Тонкович, Э.В. Асадчая, А.Д. Ларькин Организация разделов электронной формы регистрации участников научной конференции с формированием онлайн базы данных.....	472

UDC 00.013.19

BASICS OF DIGITAL PROCESSING OF IMAGES



S.S. Beknazarova

*Head of department of Audiovisual technologies of Tashkent university of information technologies named after Muhammad al-Khwarizmi, Doctor of technical science, professor
saida.beknazarova@gmail.com*

S.S. Beknazarova

Doctor of technical Sciences, Professor, head of department of Audiovisual technologies, faculty of Television technologies, Tashkent University of information technologies named after Muhammad al-Khwarizmi. Author of 189 research papers on audio, video, multimedia resources, and media education technologies.

Abstract. The article describes replacing the continuous image with a discrete different ways. For example, an image can be replaced with them by choosing a system of any orthogonal functions and calculating the image display coefficients according to this system (based on this). The variety of bases allows you to create different discrete images of a permanent image. However, the most common is periodic sampling, in particular sampling with a rectangular raster, as mentioned above. Such a sampling method can be considered as one of the options for using an orthogonal basis that has been modified as its elements-using functions.

Keywords: distributed system, lossless data processing, ensuring continuity of data flow, various types of information, image analysis, quality image.

Introduction.

Images obtained in information systems are digital. Therefore, transferring them to this form is a mandatory operation if it is necessary to use digital processing, transmission, storage. As with one-dimensional signals, this transformation involves two procedures. The first consists of replacing a continuous frame with a discrete one and is usually called sampling, while the second replaces a set of constant light values with a set of quantized values and is called quantization. In digital form, a binary number is assigned to each of the quantized values of brightness, which allows the image to be inserted into the computer [1].

The two-dimensional nature of the image compared to traditional signals includes additional possibilities for optimizing the digital image to reduce the amount of digital data received. In this regard, the question of the best placement of quantization levels, as well as the use of various rasts [1.1 ... 1.3] and other aspects of this problem were studied. However, it must be said that in most cases, sampling and uniform quantization of brightness are practiced, based on the use of a rectangular raster. This is due to the simplicity of performing the corresponding operations and the relatively small benefit of using optimal transformations. When a rectangular raster is used in the final form, the digital image is usually a matrix, whose rows and columns correspond to the rows and columns of the image [2-4].

Replacing a continuous image with a discrete one can be done in different ways. For example, an image can be replaced with them by choosing a system of any orthogonal functions and calculating the image display coefficients according to this system (based on this). The variety of bases allows you to create different discrete images of a permanent image. However, the most common is periodic sampling, in particular sampling with a rectangular raster, as mentioned above. Such a sampling method can be

considered as one of the options for using an orthogonal basis that has been modified as its elements-using functions. Next, mainly after [1.1], we will take a closer look at the main features of rectangular discretization [5].

let $x(t_1, t_2)$ be a continuous image, the corresponding discrete $x(i_1, i_2)$, obtained continuously by rectangular sampling. This means that the relationship between them is determined by the following expression:

$$x(i_1, i_2) = x(i_1 \Delta t_1, i_2 \Delta t_2) \quad (1)$$

Where $\Delta t_1, \Delta t_2$ - vertical and horizontal steps or sampling intervals, respectively.

Picture.1 the rectangle shows the location of the samples in the plane with the samples (t_1, t_2) .

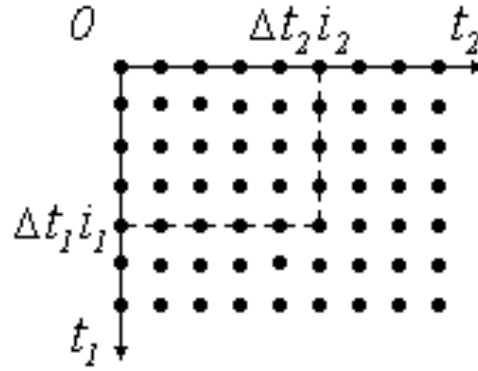


Figure 1. Sample location for rectangular sampling

The main question that arises when replacing a continuous image with a discrete one is to determine the conditions under which such a replacement is complete, i.e. not accompanied by a loss of information in a continuous signal. If it is possible to restore a continuous signal with a discrete signal, there will be no losses. From a mathematical point of view, the question is to restore a continuous signal in two-dimensional intervals between nodes whose values are known, or, in other words, to carry out two-dimensional interpolation. This question can be answered by analyzing the spectral characteristics of continuous and discrete images [6].

2D continuous frequency spectrum $X_H(\Omega_1, \Omega_2)$ continuous signal $x_H(t_1, t_2)$ determined by two-dimensional direct Fourier transform:

$$X_H(\Omega_1, \Omega_2) = \iint_{-\infty-\infty}^{\infty \infty} x_H(t_1, t_2) \exp(-j\Omega_1 t_1 - j\Omega_2 t_2) dt_1 dt_2 \quad (2)$$

This corresponds to a two-dimensional inverse continuous Fourier transform:

$$x_H(t_1, t_2) = \frac{1}{4\pi^2} \iint_{-\infty-\infty}^{\infty \infty} X_H \cdot (\Omega_1, \Omega_2) \exp(j\Omega_1 t_1 + j\Omega_2 t_2) d\Omega_1 d\Omega_2 \quad (3)$$

The last relation is true for any values of t_1, t_2 , including $t_1 = i_1 \Delta t_1$ in rectangular grid nodes $t_1 = i_1 \Delta t_1, t_2 = i_2 \Delta t_2$ Therefore, taking into account (1), for signal values in nodes (3), the relationship can be written in the following form:

$$x(i_1, i_2) = \frac{1}{4\pi^2} \iint_{-\infty-\infty}^{\infty \infty} X_H \cdot (\Omega_1, \Omega_2) \exp(j\Omega_1 i_1 \Delta t_1 + j\Omega_2 i_2 \Delta t_2) d\Omega_1 d\Omega_2 \quad (4)$$

To shorten

$$\frac{-\pi + 2\pi k_1}{\Delta t_1} \leq \Omega_1 \leq \frac{\pi + 2\pi k_1}{\Delta t_1}, \quad \frac{-\pi + 2\pi k_2}{\Delta t_2} \leq \Omega_2 \leq \frac{\pi + 2\pi k_2}{\Delta t_2},$$

the area of a rectangle in a two-dimensional frequency field we mark with $S(k_1, k_2)$. Separate the calculation of the integral in (4) by the entire frequency field

$$x(i_1, i_2) = \frac{1}{4\pi^2} \sum_{k_1=-\infty}^{\infty} \sum_{k_2=-\infty}^{\infty} \iint_{S(k_1, k_2)} X_H(\Omega_1, \Omega_2) \times \\ \times \exp(j\Omega_1 i_1 t_1 + j\Omega_2 i_2 t_2) d\Omega_1 d\Omega_2.$$

$\Omega_1 \rightarrow \Omega_1 - 2\pi k_1 / \Delta t_1$, $\Omega_2 \rightarrow \Omega_2 - 2\pi k_2 / \Delta t_2$ by changing the variables according to the rule of the We integration space of k_1 and k_2

we will achieve independence from the numbers:

$$x(i_1, i_2) = \frac{1}{4\pi^2} \int_{-\frac{\pi}{\Delta t_1}}^{\frac{\pi}{\Delta t_1}} \int_{-\frac{\pi}{\Delta t_2}}^{\frac{\pi}{\Delta t_2}} \sum_{k_1} \sum_{k_2} X_H\left(\Omega_1 + \frac{2\pi k_1}{\Delta t_1}, \Omega_2 + \frac{2\pi k_2}{\Delta t_2}\right) \times \\ \times \exp(j\Omega_1 i_1 \Delta t_1 + j\Omega_2 i_2 \Delta t_2) d\Omega_1 d\Omega_2.$$

Here for any integer values of k and i

$$\exp(-j2\pi ki) = 1$$

taken into account. This expression is very close in form to the inverse Fourier transform. The only difference is the wrong form of the exponential factor. To give it the desired shape, we use normalized frequencies $\omega_1 = \Omega_1 \Delta t_1$, $\omega_2 = \Omega_2 \Delta t_2$ we enter and change the variables accordingly. As a result, we get:

$$x(i_1, i_2) = \frac{1}{4\pi^2} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \frac{1}{\Delta t_1 \Delta t_2} \sum_{k_1} \sum_{k_2} X_H\left(\frac{\omega_1 + 2\pi k_1}{\Delta t_1}, \frac{\omega_2 + 2\pi k_2}{\Delta t_2}\right) \times \\ \times \exp(j\omega_1 i_1 + j\omega_2 i_2) d\omega_1 d\omega_2 \quad (5)$$

Now the expression (5) has the form of reverse Fourier Transform, which means that the function under the integral sign

$$X(\omega_1, \omega_2) = \frac{1}{\Delta t_1 \Delta t_2} \sum_{k_1} \sum_{k_2} X_H\left(\frac{\omega_1 + 2\pi k_1}{\Delta t_1}, \frac{\omega_2 + 2\pi k_2}{\Delta t_2}\right) \quad (6)$$

is a two-dimensional spectrum of a discrete image. In the plane of non-normalized frequencies (6), the expression has the following appearance:

$$X(\Omega_1 \Delta t_1, \Omega_2 \Delta t_2) = \frac{1}{\Delta t_1 \Delta t_2} \sum_{k_1} \sum_{k_2} X_H\left(\Omega_1 + \frac{2\pi k_1}{\Delta t_1}, \Omega_2 + \frac{2\pi k_2}{\Delta t_2}\right) \quad (7)$$

(7) from the two-dimensional spectrum of a discrete image Ω_1 and Ω_2 along the frequency axes, respectively $\frac{2\pi k_1}{\Delta t_1}$ and $\frac{2\pi k_2}{\Delta t_2}$ it follows that a rectangle with periods is periodic. Discrete representation

$X(\Omega_1, \Omega_2)$ spectrum apart $\frac{2\pi k_1}{\Delta t_1}$ and $\frac{2\pi k_2}{\Delta t_2}$ infinite number of continuous images that differ in frequency shifts $X_H(\Omega_1, \Omega_2)$ it is formed as a result of the sum of its Spectra. Continuous in Figure 2 (2.a-picture) and discrete (2.B-image) the connection between two-dimensional spectra of images is shown qualitatively [7].

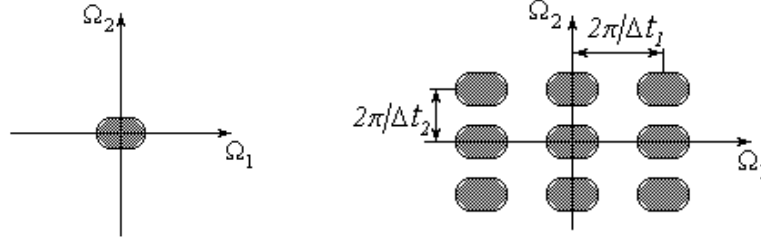


Figure 2. (2.a-picture) and discrete (2.b-image)

The result of the assembly depends mainly on the values of these frequency shifts or in other words, $\Delta t_1, \Delta t_2$ depends on the choice of sampling intervals. Let's assume, $X_H(\Omega_1, \Omega_2)$ the spectrum of a continuous image is not zero in some two-dimensional region close to zero frequency, that is, it is represented by a two-dimensional finite function. If in this case the sampling intervals $X_H(\Omega_1, \Omega_2) = 0$ with $|\Omega_1| \geq \pi/\Delta t_1, |\Omega_2| \geq \pi/\Delta t_2$ when choosing with, it does not happen to put separate branches in the formation of the aggregate (7). Consequently, each rectangle has a cross section $S(k_1, k_2)$ in only one term is different from zero. In particular, for $k_1 = 0, k_2 = 0$ we have: $X(\Omega_1, \Omega_2) = \frac{1}{\Delta t_1 \Delta t_2} X_H(\Omega_1, \Omega_2)$ pri $|\Omega_1| \geq \pi/\Delta t_1, |\Omega_2| \geq \pi/\Delta t_2$ (8)

Thus, the frequency field $S(0,0)$ in continuous and discrete spectra of images correspond to a constant factor. In this case, the spectrum of a discrete image in this frequency zone contains complete information about the spectrum of a continuous image. We note that this coincidence occurs only under conditions determined by a good selection of sampling intervals. Note that according to (8) the fulfillment of these conditions must meet the requirements $\Delta t_1, \Delta t_2$, achieved at sufficiently small values of sampling intervals:

$$\Delta t_1 \leq \pi/\Omega_{1gr}, \Delta t_2 \leq \pi/\Omega_{2gr} \quad (9)$$

Here $\Omega_{1gr}, \Omega_{2gr}$ -cutting frequencies of a two-dimensional spectrum.

Attitude (8) $x_H(t_1, t_2)$ discrete continuous image $x(i_1, i_2)$ determines the method of obtaining. To do this, it is enough to carry out two-dimensional filtering of a discrete image with a low-frequency filter with a frequency response.

$$K(j\Omega_1, j\Omega_2) = \begin{cases} \Delta t_1, \Delta t_2 \text{ pri } |\Omega_1| \leq \pi/\Delta t_1, |\Omega_2| \leq \pi/\Delta t_2 \\ 0 \text{ pri drugix } \Omega_1, \Omega_2 \end{cases} \quad (10)$$

The spectrum of the image at the output is only $S(0,0)$ contains non-zero components in the frequency zone and, according to (8), a constant image $X_H(\Omega_1, \Omega_2)$ is equal to the spectrum. It is an ideal low-frequency filter output image means it is the same with $x_H(t_1, t_2)$.

Thus, an ideal interpolation reconstruction of a continuous image is carried out using a two-dimensional filter with a rectangular frequency response (10). It is not difficult to accurately write a continuous image recovery algorithm. The two-dimensional pulsed response of the reconstruction filter, which can be easily obtained using the reverse Fourier Transform of (10), has the following form [8]:

$$x(t_1, t_2) = \frac{\sin(\pi t_1 / \Delta t_1) \sin(\pi t_2 / \Delta t_2)}{\pi t_1 / \Delta t_1 \pi t_2 / \Delta t_2}$$

The filter product can be determined using the 2D convolution of the input image and the given impulse response. Input image $x_{vx}(t_1, t_2)$ δ - expression as a two-dimensional sequence of functions.

$$x_{vx}(t_1, t_2) = \sum_{i_1} \sum_{i_2} x(i_1, i_2) \cdot \delta(t_1 - i_1 \Delta t_1) \delta(t_2 - i_2 \Delta t_2)$$

after convolution we find:

$$x(t_1, t_2) = \sum_{i_1} \sum_{i_2} x(i_1, i_2) \cdot \frac{\sin[\pi(t_1 - i_1 \Delta t_1) / \Delta t_1] \sin[\pi(t_2 - i_2 \Delta t_2) / \Delta t_2]}{\pi(t_1 - i_1 \Delta t_1) / \Delta t_1 \pi(t_2 - i_2 \Delta t_2) / \Delta t_2} \quad (11)$$

The resulting relationship shows a method of accurate interpolation reconstruction of a continuous image from a certain sequence of its two-dimensional samples. According to this expression, for proper reconstruction, two-dimensional functions of the \sin^x/x Type should be used as interpolating functions. Attitude (11) is a two-dimensional variant of the Kotelnikov-Nyquist theorem.

Once again, we note that if the two-dimensional spectrum of the signal is limited and the sampling intervals are small enough, these results are real. If at least one of these conditions is not met, the validity of the conclusions is violated. Real images rarely have Spectra with clear boundary frequencies. One of the reasons that leads to the infinity of the spectrum is the limited size of the image. Therefore, when assembling (7), the movement of terms from adjacent spectral ranges manifests itself in each of the $S(k_1, k_2)$ ranges. In this case, it will be completely impossible to accurately reproduce a permanent image. In particular, the use of a rectangular filter does not lead to precise reconstruction [9,10].

Conclusion.

The feature of optimal image reconstruction at intervals between samples (11) is the use of all samples of a discrete image, as shown in the procedure. This is not always convenient, it is often required to reconstruct the signal in the local area, relying on the small number of available discrete values. In such cases, it is recommended to apply a quasi-optimal recovery using various interpolation functions. Such a problem arises, for example, when solving the problem of connecting two images, when, due to the geometric incompatibility of these images, the existing readings of one of them can correspond to certain points located in the intervals between the nodes other.

References

- [1]. B. Zitova, J. Flusser Image registration methods: a survey // Image and Vision Computing. - 2003. - V.21. - P. 977-1000.
- [2]. R. Gonzalez, R. Woods Digital image processing / Transl. from English-M.: Technosphere, 2006 - 1070 p.
- [3]. I. S. Guzman et al. Digital image processing in information systems - Novosibirsk: NSTU, 2000. - 156 p.
- [4]. F. Durand and J. Dorsey Fast Bilateral Filtering for the Display of High-Dynamic- Range Images. In John Hughes, editor, SIGGRAPH 2002 Conference Graphics Proceedings, Annual Conference Series // ACM Press/ACM SIGGRAPH 2002 - P. 257-265.
- [5]. Hirschmüller //, Gehrig S. Stereo Matching in the Presence of Sub-Pixel Calibration Errors // IEEE Conf. on Computer Vision and Pattern Recognition. Miami, Florida, USA. - 2009. - P. 437-444.

- [6]. H. Hirschmüller Accurate and efficient stereo processing by semi-global matching and mutual information // IEEE Conf. on Computer Vision and Pattern Recognition. San Diego, California, USA. - 2005. - V. 2. - P. 807-814.
- [7]. G. Bastin, J. M. Coron and B. d'Andrea Novel, Networks and Heterogeneous Media 4(2), 177-187 (2009).
- [8]. M. Krstic and A. Smyshlyaev, Advances in Design and Control. Society for Industrial and Applied Mathematics (SIAM) 16, 13-22 (2008).
- [9]. Beknazarova, S., Bazhenov, R., Sabirova, V., Samorukov, A., Berseneva, S. Technologies of embedded project-based learning for undergraduate engineering students/ Journal of Physics Conference Series 2001(1):012028 DOI:10.1088/1742-6596/2001/1/012028.
- [10]. Beknazarova, S., Bazhenov, R., Zatonkiy, A Advantages of Freeware-based Simulation Tools for Technical and Technological Modeling/ Conference: 2021 International Conference on Industrial Engineering, Applications and Manufacturing (ICIEAM) DOI:10.1109/ICIEAM51226.2021.9446186.

ОСНОВЫ ЦИФРОВОЙ ОБРАБОТКИ ИЗОБРАЖЕНИЙ

С.С. Бекназарова

*Заведующий кафедрой АВТ Ташкентского
университета информационных технологий
имени Мухаммеда Аль-Хорезми, доктор
технических наук, профессор*

Кафедра аудиовизуальных технологий

*Ташкентский университет информационных технологий имени Мухаммеда Аль-Хорезми, Ташкент,
Узбекистан*

E-mail: saida.beknazarova@gmail.com

Аннотация. В статье описываются методы замены непрерывного изображения дискретным, которые могут осуществляться различными способами. Например, можно выбрать любую систему ортогональных функций и заменить ими изображение, вычислив коэффициенты отображения изображения по этой системе (по этому основанию). Разнообразие основ позволяет создавать различные дискретные изображения непрерывного изображения. Однако наиболее распространенной является периодическая выборка, в частности, как упоминалось выше, выборка с прямоугольным растром. Такой метод выборки можно рассматривать как один из вариантов использования ортогонального базиса, в качестве элементов которого используются модифицированные-функции.

Ключевые слова: распределенная система, обработка данных без потерь, обеспечение непрерывности потока данных, различные виды информации, анализ изображений, качество изображения.

UDC 631.15:33

FEATURES OF SEARCHING FOR ASSOCIATIONS IN LARGE VOLUMES OF MARKETING DATA



D. R. Rahel

*PhD in Economics, Associate Professor of the Department of Economics of the Belarusian State University of Informatics and Radioelectronics
ragel@bsuir.by*

D.R. Rahel

In 2000 he graduated from the Faculty of Economics of the Belarusian State University of Informatics and Radioelectronics with a degree in Economic Informatics. In 2016 he graduated from the postgraduate course of the Academy of Management under the President of the Republic of Belarus. In 2018 he defended his PhD thesis. Research interests: data mining in marketing, process modeling, data analysis, statistical forecasting, macroeconomics.

Annotation. The article describes an approach to searching for associations in baskets of goods and searching for patterns in transactions. In addition, approaches are given to the analysis of big data on the behavior of the customer base, which can be obtained in the course of market research. The stages of the analysis of consumer behavior and the search for patterns in it are described.

Keywords: data analysis, marketing data, big data, data array, marketing research, associations, analysis of the actions, buyers activity, classification, clustering, frequency pattern.

As part of marketing research, one of the main tasks is often to find links and dependencies between elements obtained from different data sources. In fact, we are talking about an associative analysis of disparate and disordered information.

If we consider customer transaction data, which collects data on customer purchases, that is, the goods purchased on a specific date or within a fixed period of time are specified, then the main goal of further marketing forecasting of customer activity is the association between groups of goods or individual goods in the considered datasets. In this case, by association we mean patterns that can be obtained analytically, which can later lead to the development of rules that help optimize the company's commercial activities and classify product categories and customer groups. Many goods are divided into "baskets" with patterns that are inherent in the goods that make them up. For example, if one of the baskets contains "bread, flour, butter", then it is likely to contain milk as well. And if there is a "hammer, glue" in the basket, then milk is unlikely to be present in it, but rather something from the categories close to these goods, such as, for example, "nails".

The search for association rules based on the received and systematized data is an activity to maximize income based on the formation of the company's product basket. Taking into account the processing of data and the formation of "baskets", further decisions are made on discounts for certain categories of goods, on the termination of the commercial operation of certain categories of goods, or the launch of new product categories with a higher commercial potential on the market.

Collecting data on transactions is now a normal practice in almost all companies. The formation of association rules for categories should be accompanied by probabilistic estimates, since the implementation of associations in practice is probabilistic. Thus, in the formation of associative rules, the condition and the consequence from it play a role. For example, the condition is "hammer, glue", and

“nails” is already a consequence that has some kind of probabilistic assessment depending on the practical implementation.

At the stage of practical implementation at the level of a business entity, the following rules are formed:

- at the level of the condition, that is, what was bought with the subsequent generalization of the most typical consequences; (1)

$$\{\text{hammer, glue}\} \rightarrow \{\text{nails}\} \quad (1)$$

- at the level of the investigation, with the formulation of the rules that eventually led to it. (2)

$$\{\text{nails}\} \rightarrow \{\text{hammer, glue, nails}\} \quad (2)$$

If we touch upon the problem statement and apply associations to the database with records of the company's commercial transactions, then for such a database (DB) with a set of data on transactions $Tr_1, Tr_2, Tr_3, \dots, Tr_n$, it is required to find a set of templates that can express any then the minimum of transactions for the considered period of time. The number of transactions that are described within the templates can be denoted as N . The parameter N , in turn, can be represented as the number of transactions selected from the general set that fall under this template. At the same time, each transaction in the data set is expressed as a binary vector, one of the coordinates of which is an attribute that describes the presence of an order for a specific product in a specific outlet. For example, let's take 2 some sets from outlets $M1$ and $M2$. Thus, $M1 \Rightarrow M2$ is an association with a minimum support of H under the conditions:

1. $M1 \cup M2$ - is a frequency pattern;
2. The ratio $M1 \cup M2$ to $M1$ has a minimum certainty and is less than 1 if the sets are not equal to each other.

From a computational point of view, a pattern is a vector in the same space as the set of transactions in question. In this case, the template is the result of applying the logical operation "and" to the set of transactions under consideration.

The most typical frequency pattern search algorithm is a pattern-merge based algorithm by sequentially combining patterns of length k to sequentially generate frequency patterns of length $k+1$. The main goal is to reduce the search space for frequency patterns based on the observation that no pattern can contain low frequency patterns.

1. Input:

- Database.
- Minimum C support.

2. Output:

- Lots of SN frequency patterns.

3. Algorithm:

- Generation of two sets of frequency patterns $M1$ and $M2$ respectively with duration $D1$ and $D2$.
- $k=2$
- As long as the template set is not empty:
 - i. Generation of all possible templates of length $k+1$ due to pairwise union of available templates;
 - ii. Exclusion from the set of all patterns of length $k+1$ that have subpatterns without sufficient support;
 - iii. Generation of a finite set by copying all the remaining templates of dimension $k + 1$ into it and calculating their support.
 - iv. Return a final set.

The search for associative links based on patterns in marketing activities has a number of useful applications for practical use:

- Analysis of the activity of buyers in the framework of the operating activities of trading companies. Transaction information in this case is a collection of information about sets of jointly purchased goods. In this case, the development of templates is the generation of recommendations to potential buyers based on previous experience. If the current basket intersects with the template available in the database, then the buyer receives a recommendation regarding the rest of the goods from the template.

- Carrying out classification or clustering of customers or customer base based on the company's existing patterns of purchases, when all similar transactions are combined in the database based on compliance with existing pattern types of clusters;

- Analysis of the actions of service users. In this case, the database may consist of sets of elementary user actions online, templates in turn for generating recommendations.

- Analysis of application errors. The application generates standardized sequences of logs, while paying attention to the analysis of low-frequency patterns for their fallacy.

References

[1] Мыльников Л.А. Статистические методы интеллектуального анализа данных. – БХВ-Петербург, 2021. – 119 с.

[2] Data Science and Big Data Analytics. A Step by Step Guide to learn Data Science from Scratch with Python Machine Learning and Big Data / Andrew Park. – Published by Andrew Park, 2021. – 124 p.

[3] Nicholson W.L. Exploring Data Analysis. – Nobel Press, 2012. – 421 p.

ПОИСК АССОЦИАЦИЙ В БОЛЬШИХ ОБЪЕМАХ МАРКЕТИНГОВЫХ ДАННЫХ

Д. М. Рагель

к.э.н., доцент кафедры экономики Белорусского государственного университета информатики и радиоэлектроники

*Белорусский государственный университет информатики и радиоэлектроники,
Республика Беларусь
e-mail: ragel@bsuir.by*

Аннотация. В статье изложен подход к поиску ассоциаций в корзинах товаров и поиск закономерностей в сделках. Кроме этого приводятся подходы к анализу больших данных о поведении клиентской базы, которые могут быть получены в ходе исследований рынка. В завершении описываются этапы анализа потребительского поведения и поиска в нем закономерностей.

Ключевые слова: анализ данных, маркетинговые данные, большие данные, массив данных, маркетинговые исследования, ассоциации, анализ действий, покупательская активность, классификация, кластеризация, частотная характеристика.

UDC 004.512-026.26

HUMAN PHYSICAL ACTIVITY RECOGNITION ALGORITHM BASED ON SMARTPHONE DATA AND LONG SHORT TIME MEMORY NEURAL NETWORK



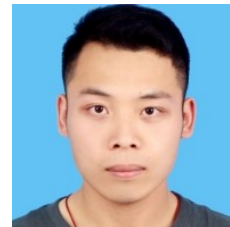
Z.Y. Chen

*Master student of the faculty of Information Security of the Belarusian State University of Informatics and Radioelectronics
chenzheyin1@gmail.com*



Z.X. Yang

Master student of the faculty of Information Security of the Belarusian State University of Informatics and Radioelectronics



H. Li

PhD student of the faculty of Information Security of the Belarusian State University of Informatics and Radioelectronics

Z.Y. Chen

Master student of the Department of Information Security of the Belarusian State University of Informatics and Radioelectronics, graduated from Nanchang Hangkong University. Her research interests include neural networks, object classification based on sensor data, and object detection based on images and video streams.

Z.X. Yang

Master student of the Department of Information Security of the Belarusian State University of Informatics and Radioelectronics, graduated from Nanchang Hangkong University. His research interests neural networks and signal processing and signal processing for sensors.

H. Li

Received the B.S. and M.S. degrees in computer science from the Henan University. He is currently pursuing the Ph.D degree with the Belarusian State University of Informatics and Radioelectronics. His research interests include Neural Network, mobile crowd-sensing networks and signal processing.

Abstract. The continuous advancement of smartphone sensors has brought more opportunities for the universal application of human motion recognition technology. Based on the data of the mobile phone's three-axis acceleration sensor, using combining a double-layer Long Short Time Memory (LSTM) and full connected layers allow us to improve human actions recognition accuracy, including walking, jogging, sitting, standing, and going up and down stairs. This is helpful for smart assistive technology. It is shown that physical activity classification accuracy is equal to 97 %.

Keywords: mobile acceleration sensor, long short time memory, action recognition and classification accuracy.

Introduction.

Traditional sensor devices are bulky and expensive. With the continuous development of smart phones in recent years, the acceleration sensors of mobile phones have also continued to improve. It has the obvious advantages of small size, high penetration rate and lower price, which provides a new idea for the application of intelligent assistance technology and so on. Recognition of human activities from sensor data is at the core of intelligent assistive technologies, such as smart home, rehabilitation, health support, skills assessment or industrial environments [1]. For example, the project of Inooka et al. predicts the energy consumption of users by recognizing their activities [2], and Mathie et al. judges whether users are safe or not by recognizing their actions [3]. This work is motivated by two requirements of activity recognition: improving recognition accuracy and reducing reliance on engineered features to address increasingly complex recognition problems.

Human Activity Recognition (HAR) is based on the assumption that specific body movements translate into characteristic sensor signal patterns, which can be sensed and classified using machine

learning techniques. We use data collected from accelerometer sensors. Almost every modern smartphone has a three-axis accelerometer that measures acceleration in all three spatial dimensions.

We selected the data set from the Wireless Sensor Data Mining (WISDM) project, which collected 1,098,207 experimental data generated from 29 volunteers carrying smartphones to perform specified actions every 50 ms, and each piece of data consists of 6 parts: Username, specified action, timestamp and accelerometer values for x, y and z axis. We use a window of size 200 with an overlap of 90 % to divide the x, y and z axis accelerometer and label part in the original data, store them as acceleration data and label data respectively for preprocessing. We get 54901 windows and split both data into a training set (80 %) and a test set (20 %).

We trained a double layer LSTM neural network (implemented in TensorFlow) for HAR from accelerometer data with the purpose of providing an algorithm with higher recognition accuracy. The trained model will be exported/saved and added to the Android app. The network model consists of double layer LSTM network layers and double fully connected layer (FCL), and predicts the corresponding human actions from the x, y and z axis acceleration count values from the data set. The proposed algorithm achieved 97 % accuracy and a loss of 0,2 on the test set.

Double FCL and double layer LSTM neural network architecture based on mobile phone accelerometer.

We use combining a double LSTM layer and double FCL to establish a deep learning model, aiming to predict the user's action type at a certain moment through the three-axis acceleration data from the mobile phone sensor. The proposed neural network model is shown in Figure 1. We partition the training set consist of 3-axis acceleration data (x, y and z) with a batch size of 1024, resulting in 50 iterations, each of iteration takes a tensor of shape (1024, 200, 3) as the input. We first flatten it to 204800 tensors (1,3) as input to the FCL1. FCL1 abstracts it into 204800 tensors (1,64), then we stack it into 200 tensors (1024,64) as input to the double layer LSTM and get the features of 1024 samples over 200 of time steps with one tensor (1024,64). We flatten it again to 1024 tensors (1,64) as input to the FCL2 to divided it into 6 physical human activities. Softmax layer converts the input from the previous layer into probability set of 6 physical human activities as the output. Finally, we estimate the performance of the recognition results.

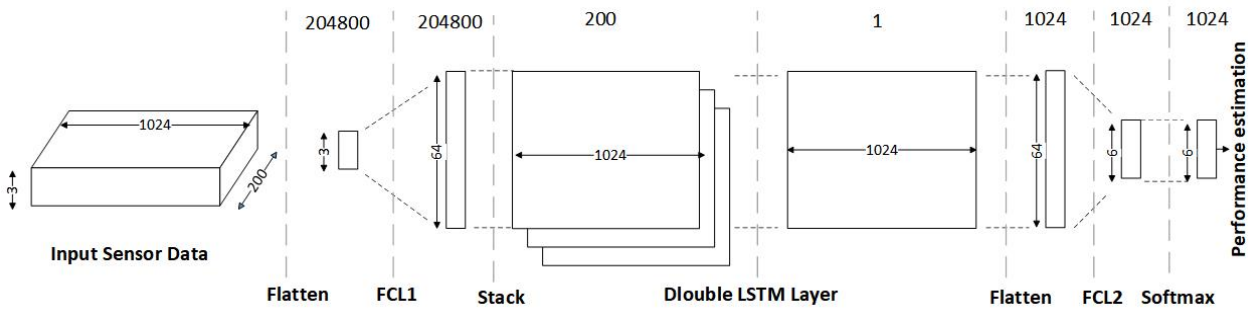


Figure 1. Block diagram of human activity recognition algorithm based on smartphone data and deep learning

Suppose a volunteer in the data set performs N predicted physical human activities within a certain period of time t , see (1):

$$A = \{a_1, \dots, a_n, \dots, a_N\} \tag{1}$$

where a is the specified action type, a_n is the n th physical human activity.

The acceleration sensor from the mobile phone collects the three-axis acceleration data D corresponding to N predicted physical human activities that the user wishes to complete within the time period t , see (2):

$$D = \{d_1, \dots, d_n, \dots, d_N\} \quad (2)$$

where $d_n = (x_n, y_n, z_n)$, $x_n = (x_1, \dots, x_k)$, $y_n = (y_1, \dots, y_k)$, $z_n = (z_1, \dots, z_k)$ is the k three-axis acceleration data of n th physical human activity.

This algorithm aims to identify n actions within the time period t based on the three-axis acceleration data D by constructing a neural network model F , see (3):

$$F(D) = \{a'_1, \dots, a'_n, \dots, a'_N\}, a'_n \in A \quad (3)$$

where a' is the predicted action type, a'_n is the n th predicted physical activity.

The learning parameters and hyperparameter of the double layer LSTM neural network based on mobile phone accelerometer are shown in Table 1 and Table 2.

Table 1. The learning parameters of double FCL and double layer LSTM neural network based on mobile phone accelerometer

Modules	Learning parameters	matrix size
Fully Connected Layer 1	Weight configuration	[3,64]
	bias	[64]
Double LSTM Layer	Number of params	66048
Fully Connected Layer 2	weight	[64,6]
	bias	[6]

Table 2. The hyperparameter of double FCL and double layer LSTM neural network based on mobile phone accelerometer

Modules	Hyperparameter	Value
Fully Connected Layer 1	hidden unit size	64
	activation function	ReLU
Stacked LSTM Layer	hidden unit size	64
Fully Connected Layer 2	hidden unit size	6
Training	optimizer	Adam
	batch size	1024
	learning rate	0.0025
	number of epochs	50

Fully connected layer 1.

For fully connected layer 1, we have input 204800 tensors of shape (1,3), all of data type float32. For input 204800 samples, each sample has a length of 3, representing the acceleration of the x , y and z axis respectively. FCL1 contains 64 neurons and abstracts the 3 features of the input to 64, so as to better divide different types of data. The structure and parameters are shown in Table 3.

For the input, we multiply a weight matrix of size [3,64] and add a bias matrix of size [64]. The output Y_j of the j th neuron of FCL1 is shown in the formula (4), we use ReLU (Rectified Linear Activation Function) as the activation function f to get an output with shape [1,64].

$$Y_j = f \left(\sum_{i=1}^n x_i w_{ij} + w_{j0} \right) \quad (4)$$

where x_i is the i th input variable, w_{ij} is the weight between the i th input variable and the j th FCL1 neuron, and w_{j0} is the bias of the j th FCL1 neuron, n is the number of input dimensions 3.

Table 3. The input, condition and output of fully connected layer 1

FCL1	Value
Input	tensor (1,3)
	dtype: float32
Condition	Weight matrix size [3,64]
	Bias matrix [64]
	activation function: ReLU
Output	tensor (1,64)
	dtype: float32

The ReLU is a simple calculation that directly returns the value provided as input if the input is greater than 0, or returns a value of 0 if the input is 0 or less. Due to the sparsity of ReLU, the sparse model can better mine relevant features and fit the training data.

Next, we split the 204800 outputs according to the size of the time steps 200 into a list of two order tensors of the shape (1024,64).

Double layer LSTM.

We use double layer LSTM to form the LSTM layer, the number of hidden units is 64, and the data type is float32. For an input list of 200 tensors of shape (1024,64), we get the output tensor of shape (1024,64), that is, the features of 1024 samples over 200 of time steps are extracted to 64 dimensions. The parameters are shown in Table 4.

Table 4. The input, condition and output of double layer LSTM

Double LSTMs	Value
Input	List of 200 two order tensors (1024,64)
	dtype: float32
Condition	LSTM units: 64
Output	two order tensor (1024,64)
	dtype: float32

Double layer of LSTMs makes the model deeper and the extracted features deeper, resulting in more accurate predictions. It has achieved good results on a wide range of prediction problems.

Double layer LSTM is to take the output of the previous layer of LSTM as the input of the next layer of LSTM and send it into the network. The network structure is shown in Figure 2. The triaxial accelerometer value x_t at time t is used as the input of the first layer of LSTM cells, and the output h_t^1 of its hidden layer is simultaneously used as the input of the second layer of LSTM cells and the update of the hidden layer of the first layer of LSTM cells, and update the cell state c_t^1 of the first layer of LSTM neurons. The second layer of LSTM neurons outputs the hidden layer h_t^2 as a result, and updates the hidden layer h_t^2 and cell state c_t^2 at time t .

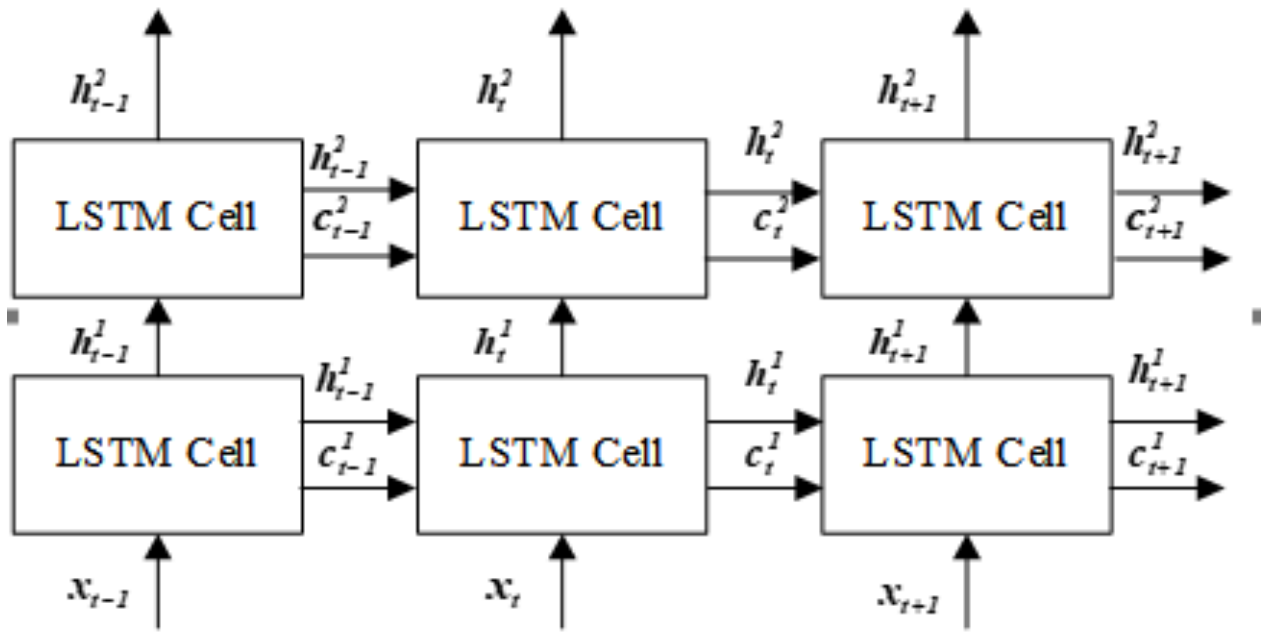


Figure 2. The structure of double LSTM layer

Fully connected layer 2.

For the fully connected layer 2, we flatten the two-dimensional tensor (1024, 64) from the double layer of LSTMs to one-dimensional to get 1024 tensors of shape (1, 64) as input.

For the input, we multiply a weight matrix of size [64,6] and add a bias of size [6], FCL2 has 6 neurons, which fuse the features from the previous layer into 6 dimensions, the output O_k to the k th neuron in FCL2 see formula (5). The parameters are shown in Table 5.

Table 5. The input, condition and output of fully connected layer 2

FCL2	Value
Input	tensor (1,64)
	dtype: float32
Condition	Weight matrix size [64,6]
	Bias [6]
Output	tensor (1,6)
	dtype: float32

$$O_k = \sum_{r=1}^m x_r w_{rk} + w_{k0} \quad (5)$$

where x_r is the r th input variable, w_{rk} is the weight between the r th input variable and the k th FCL2 neuron, w_{k0} is the bias of the k th FCL2 neuron, m is the input dimension 64.

Softmax layer.

Through the output O_k of the FCL2, we get 1024 tensor samples of shape (1,6), which respectively represent the prediction results of each 6 kinds of physical human activities (Walking, Jogging, Upstairs, Downstairs, Sitting, Standing). Next, we use the Softmax function (see formula 6) to convert each

prediction result into a number between 0 and 1, indicating the probability that the predicted result of the sample is the k th class action.

$$P(k | X) = \frac{e^{Z_k}}{\sum_{c=1}^6 e^{Z_c}} \quad (6)$$

where X is the input of double FCL and double layer LSTM neural network, Z_k is the k th value of input tensor of shape (1,6), Z_c is the c th value of input tensor of shape (1,6).

Through the Softmax function, the output value of the multi-category can be converted into a probability distribution ranging from 0 to 1. The Softmax function can convert the predicted results into non-negative numbers, and make the sum of the probabilities of various predicted results equal to 1. The input and output are shown in the Table 6.

Table 6. The input and output of Softmax

input	output
Predicted value under each physical human activity: $Z = \{Z_1, Z_2, \dots, Z_6\}$	The predicted probability of each physical human activity: $P = \{p_1, p_2, \dots, p_6\}$

The recognition algorithm performance estimation.

The model is tested on the human behavior pose dataset WISDM. We selected 80 % of the data in the WISDM dataset for model training and 20 % for model testing. Using Adam as optimizer, MSE as model measure, learning rate 0,0025, batch size 1024, training for 50 rounds. We evaluate the model by confusion matrix, training loss and accuracy.

Confusion matrix.

The confusion matrix our work is shown in Table 7. where A means activity, W means walking, J means jogging, U means upstairs, D means downstairs, SIT means sitting, ST means standing, N is the number of various activities.

Table 7. Confusion Matrix for Multiple Classification Tasks

Activities		Predicted label						total
		Walking	Jogging	Upstairs	Downstairs	Sitting	Standing	
True label	Walking	A_{WW}	A_{JW}	A_{UW}	A_{DW}	A_{SITW}	A_{STW}	N_W
	Jogging	A_{WJ}	A_{JJ}	A_{UJ}	A_{DJ}	A_{SITJ}	A_{STJ}	N_J
	Upstairs	A_{WU}	A_{JU}	A_{UU}	A_{DU}	A_{SITU}	A_{STU}	N_U
	Downstairs	A_{WD}	A_{JD}	A_{UD}	A_{DD}	A_{SITD}	A_{STD}	N_D
	Sitting	A_{WSIT}	A_{JSIT}	A_{USIT}	A_{DSIT}	A_{SITSIT}	A_{STSIT}	N_{SIT}
	Standing	A_{WST}	A_{JST}	A_{UST}	A_{DST}	A_{SITST}	A_{STST}	N_{ST}

In the field of machine learning, confusion matrix, also known as likelihood table or error matrix. It is a specific matrix used to present a visualization of algorithm performance. Each column represents the predicted value, and each row represents the actual category. The name comes from the fact that it makes it easy to indicate if multiple categories are confused.

Loss and accuracy of the proposed neural network model.

The loss function will determine the performance of the model by comparing the predicted output of the model with the expected output, and then find the optimization direction. If the deviation between

the two is very large, the loss value will be large; if the deviation is small or nearly the same, the loss value will be very low.

In this research, we use the cross-entropy loss function to calculate the loss. For each input sensor (1,6), we compare it with the label array, calculate the cross-entropy loss for the training data and evaluate the training effect in test set with formula (7). We hot-encode the labels of the training set and the test set. For each label, the encoded data is an array of length 6. In the order of the physical human activities Walking, Jogging, Upstairs, Downstairs, Sitting, Standing, if the label is the i th activity, the i th value of the array is 1, and the others are 0.

$$L = -\sum_{i=1}^M q_i \log(p_i) \quad (7)$$

where M is the number of physical human activities 6, q_i is the i th value of the label array, p_i is the i th value of the input tensor, which is the predicted probability of i th physical human activity from Softmax layer.

Accuracy is one of the most popular metrics in multi-class classification and it is directly computed from the confusion matrix. The formula of the Accuracy considers the sum of True Positive and True Negative elements at the numerator and the sum of all the entries of the confusion matrix at the denominator.

$$\text{Accuracy} = \frac{A_{WW} + A_{JJ} + \dots + A_{STST}}{N_W + N_J + \dots + N_{ST}} \quad (8)$$

The proposed algorithm learns well with recognition accuracy reaching above 97 % and loss hovering at around 0,2 in test set.

Conclusion.

The modeling result is based on the WISDM dataset, which data is obtained by volunteers putting the Android phone in the front trouser pocket to complete the 6 specified actions (Walking, Jogging, Upstairs, Downstairs, Sitting, Standing). The proposed physical activity recognition algorithm is based on combining FCL1, double layer LSTM neural network, FCL2 and Softmax layer to extract spatial and temporal features for improving classification accuracy. The proposed algorithm using 64 relevant features works well on most actions types and allows us to achieve 97 % classification accuracy, but there were 119 activities of downstairs that were wrongly predicted to go upstairs. The algorithm can be used in human activity recognition system using smartphone sensor data for monitoring physical behavior.

References

- [1] Anderson, I., Maitland. Shakra. Tracking and sharing daily activity levels with unaugmented mobile phones. / Anderson, I., Maitland.: Mobile Networks and Applications. 2007. –12 p.
- [2] Inooka, H., Ohtaki, Y. Development of advanced portable device for daily physical assessment. / Inooka, H., Ohtaki, Y.: SICE-ICASE International Joint Conference. 2006. –5878-5881 p.
- [3] Mathie, M., Celler B. Classification of basic daily movements using a triaxial accelerometer / Mathie, M., Celler B.: Medical & Biological Engineering and Computing, 2004. – 42 p.

АЛГОРИТМ РАСПОЗНАВАНИЯ ФИЗИЧЕСКОЙ АКТИВНОСТИ ЧЕЛОВЕКА НА ОСНОВЕ ДАННЫХ СМАРТФОНА И ДЛИТЕЛЬНОЙ КОРОТКОЙ ПАМЯТИ НЕЙРОННОЙ СЕТИ

Ч.Э. Чэнь

*магистрант факультета
информационной безопасности
БГУИР*

Ц.С. Ян

*магистрант факультета
информационной
безопасности БГУИР*

Х. Ли

*аспирант факультета
информационной безопасности
БГУИР*

Кафедра инфокоммуникационных технологий

Факультет информационной безопасности

Белорусский государственный университет информатики и радиоэлектроники, Республика Беларусь

Электронная почта: chenzheyin1@gmail.com

Аннотация. Постоянное совершенствование датчиков смартфонов открыло больше возможностей для универсального применения технологии распознавания движений человека. Основываясь на данных трехосевого датчика ускорения мобильного телефона, использование сочетания двухслойной долговременной памяти (LSTM) и полносвязных слоев позволяет нам повысить точность распознавания действий человека, включая ходьбу, бег трусцой, сидение, стояние и подниматься и спускаться по лестнице. Это полезно для интеллектуальных вспомогательных технологий. Показано, что точность классификации физической активности составляет 97 %.

Ключевые слова: мобильный датчик ускорения, долгая короткая память, распознавание действий и точность классификации.

UDC 004.021:004.75

HEART RATE ESTIMATION FROM PHOTOPLETHYSMOGRAM AND ACCELERATION SMARTPHONE DATA BASED ON CONVOLUTIONAL NEURAL NETWORK AND LONG SHORT TIME MEMORY



Gaowei Qiu

Postgraduate student in the
Academic Department of
Information and
Communication Technology
qgw1314@163.com

Gaowei Qiu

Graduated from Nanchang Aviation University, China. Areas of scientific interest include the development of neural network algorithms and software engineering.

Abstract. The wearable reflective photoplethysmograph (PPG) sensor can be integrated into the watch or strap to provide instantaneous heart rate (HRs), causing minimal inconvenience to users. However, the existence of motion artifacts (MAs) leads to inaccurate heart rate estimation. In order to solve this problem, I propose a new deep learning neural network to ensure accurate estimation of HR in high-intensity exercise. Methods: I propose a deep neural network based on multi-class and non-uniform multi-label classification for HR estimation. It includes two convolution layers, two short and long term memory (LSTM) layers, one connection layer and three full connection layers including softmax, also includes a concatenation layer. The average absolute error of the algorithm for all training data sets and test data sets is less than 1.5 bpm, including 1.09 bpm for training data sets and 1.46 bpm for test data sets. The proposed algorithm is superior to the most advanced methods in accurate estimation of heart rate.

Keywords: PPG, LSTM, Convolution, Concatenation, Heart rate and motion artifacts.

Introduction.

Reflective photoplethysmography (PPG) sensor measures the intensity change of skin reflected light and provides PPG signal representing the change of arterial blood volume between systolic and diastolic phases of the heart cycle. The reason why the sensor is concerned is that it can be installed in a watch or wrist strap to measure and monitor instantaneous heart rate (HRs), minimizing the inconvenience to users. However, the sensor is sensitive to motion artifacts (MAs), which come from the pressure and motion exerted on the wrist of the PPG sensor. Motion artifacts will eventually lead to inaccurate heart rate estimation. A few years ago, Zhang and others shared the data set of acceleration and PPG signals measured simultaneously in [1] motion, which prompted people to study the use of acceleration signals to eliminate motion artifacts in PPG sensors. However, despite the efforts and progress of algorithms, these methods do not always provide accurate results.

In this paper, the purpose is to measure heart rate more accurately, so I propose a new deep neural network based on multi-class and non-uniform multi-label classification for human heart rate estimation. In the proposed algorithm, I consider two power spectra from the input layer of PPG and acceleration signals. In addition, I use the acceleration signal strength in the input layer. I assume that the acceleration signal strength can provide information about the recent HR changes: high intensity represents intense exercise, which may change HR. The proposed algorithm includes two convolution layers, two LSTM layers, one connection layer and three full connection layers (including one softmax layer). In the proposed algorithm, the power spectrum from PPG and acceleration signals is fed to two convolutions

to provide MA cancellation in the PPG power spectrum. The output is flattened and connected to a fully connected layer, which is then connected to the acceleration signal strength. Then, the output is fed to two LSTM layers, and then to other full-connection layers including softmax. The LSTM layer uses minimum outliers to track HR tracking. In this algorithm, I also propose a new scheme to modify the real HR value to Gaussian distribution to evaluate the loss value. The performance of the proposed algorithm is evaluated by comparing with the results of [1] - [18] previously reported.

Based on PPG, ACC, and acceleration intensity Datasets.

I used two data sets to evaluate the proposed HR estimation algorithm - the IEEE Signal Processing Cup (ISPC) 2015 data set (n=23) and the direct measurement data set obtained by the developed equipment (n=48). These two sets of data include multi-channel PPG signal and three-axis accelerometer signal, which are measured by the equipment on the wrist during high-intensity physical exercise. At the same time, ECG signals are measured in the chest as real HR reference. The data includes 12-minute three-channel PPG signal and 50-Hz sampling three-axis acceleration signal. The dataset is divided into two groups, BAMI-I and BAMI-II. During every 4 minutes of running and walking, the subjects walked or ran with a treadmill stick in the last two minutes. I designed this link to reflect the heart rehabilitation exercise of patients with heart disease who have poor exercise ability. They usually walk or run with a treadmill stick. There are 17 males and 6 females, with an average age of 22.0 ± 1.7 years. The whole exercise process is also carried out on the treadmill. For these two data sets, the reference real heart rate was measured by ECG data recorded simultaneously by 24-h the ambulatory ECG monitor (SEER Light, GE Medical, Milwaukee, WI, USA).

In this study, I selected the ISPC data set (n=23) as the training data, because it includes various movements, such as waking up, running, resting, jumping, push up, shaking hands, stretching, pushing and boxing. I also added the own dataset BAMI-I (n=25) to the training data to increase the size of the training data, and used another dataset named BAMI-II (n=23) to test the training algorithm. For all data sets, the ecg-based HRs are calculated using an 8-second window, with a 2-second offset (6-second overlap) to obtain the HRs every 2 seconds. In the whole study, the same window length (8 seconds) and shift (2 seconds) are used to evaluate the performance of the proposed algorithm relative to the existing algorithm [1] - [17].

Description of deep neural network algorithm based on ACC and PPG.

Figure 1 (a) summarizes the architecture of the proposed algorithm. It consists of eight layers: two convolutional layers, two LSTM layers, one tandem layer and three full connection layers, including a softmax layer. More specifically, a 2D convolution layer, a one-dimensional convolution layer, a flat layer, a splice layer, a fully connected layer, two LSTM layers, and a fully connected layer followed by a softmax.

For the input layer, the power spectrum from PPG signal and acceleration signal is divided into two (size 2×222): the top signal is from PPG ($P_s(i)$), and the bottom signal is from acceleration ($P_a(i)$). Note that the power spectrum is based on each 8-second window and then shifted for 2 seconds (6-second overlap). Input layer input to 32×37 cores, two-dimensional convolution with a step of 4, and then perform the nonlinear function leaky corrected linear unit (ReLU) and 1×2 max pooling, in steps of 2. The resulting size is $1 \times 28 \times 32$. The characteristic diagram of 32 shows the intermediate PPG power spectrum after MA cancellation. Send characteristic map into 64×5 -core, one-dimensional convolution with a side of 1, and then carry out leaky ReLU and 1×2 max pooling, get other sizes of $1 \times 14 \times 64$, showing the PPG power spectrum of MA elimination. The resulting feature map is spread evenly to 896 nodes, which are fully connected to 512 nodes with leaking ReLU.

Subsequently, the acceleration intensity I^a (I) is connected with 512 activations. I assume that the acceleration intensity represents how the current HR value will change in the near future. For example, high acceleration intensity means high intensity exercise, which may increase HR. 513 connected nodes are sent to two LSTM layers. These two LSTM layers act as HR tracking algorithms by considering the local HR tracking mode. Due to the existence of LSTM layer, even if the signal to noise ratio (SNR) of PPG signal is extremely low, the dominant frequency corresponding to HR will not deviate seriously in

the continuous window. The first layer LSTM contains 512 nodes, each node has 6 time steps, and all nodes are connected to the second layer LSTM. Note that six time steps provide the best accuracy. The numerical analysis of the effect of time step will be introduced. The second LSTM layer also includes six time steps, each of which provides an output with a length of 222. Then, only the output from the last time step is fed to the full connection layer, which is connected to 222 nodes with leaking ReLU. In the frequency range of 0.6-3.3 Hz, the frequency resolution is 0.012 Hz, and the result of the final PPG power spectrum after MA cancellation is 222 activations. The activation is input to the softmax layer, which provides the final probability of the real HR value.

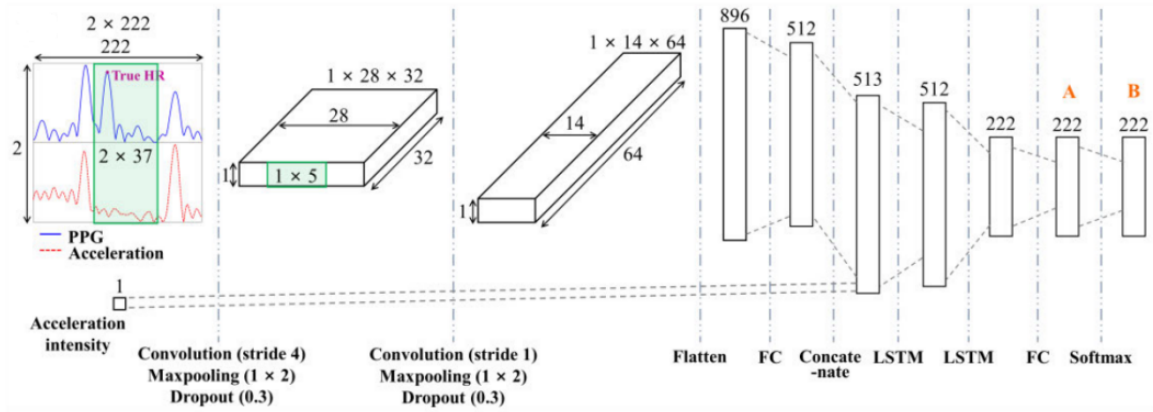


Figure 1. The architecture of the proposed algorithm; 2D convolution layer, 1D convolution layer, a full connection layer, a concatenation layer two LSTM layers and a full connection layer are structured in sequence

To avoid over-fitting, I applied dropout to two convolution layers and two LSTM layers. For convolution layer, the exit rate is set to 0.3. For LSTM layer, the drop rate of input linear transformation is 0.3, and the drop rate of loop state is 0.2. Given a set of parameters of the algorithm, softmax provides the probability corresponding to each HR value subdivided in 222 frequency boxes: from 0.6 to 3.3 Hz, with an interval of 1/222. Therefore, with the real HR value, I can consider the multi-class classification problem, which calculates the cost through multi-class cross entropy ξ Value is

$$(1)$$

where $\tilde{y}_{c=ime}$ and $y_{c=ime}$ are the true HR probability and the predicted HR probability, respectively, for the frequency bin at the c window. In the multiclass cross-entropy, $\tilde{y}_{c=ime}$ has the value of one only when the frequency bin corresponds to a true HR. Otherwise, $\tilde{y}_{c=ime}$ has the value of 0. Then, the formulation can be simplified as

$$(2)$$

where $\tilde{y}_{c=ime}^i$ is the predicted HR probability in the frequency bin covering the true HR value.

However, this method has a disadvantage that the frequency bin covering the real HR value may not accurately represent the real HR. Please note that the real HR is obtained from the ECG data sampled at 50 Hz in the BAMI I and II data sets and from the ECG data sampled at 125 Hz in the ISPC data set, while the frequency box is obtained from the PPG signal sampled at 25 Hz. To solve this problem, I will

pay ξ Multiply the normalized Gaussian function, where the Gaussian distribution centered on the real HR is normalized to get the maximum value. The revised cost function is

$$(3)$$

where y_c represents real HR. Through modification, I can alleviate the problem of non-overlapping HR based on ECG and PPG. In this study, we chose the standard deviation $\sigma = 3$.

Experimental Result.

Based on the proposed algorithm, I found that the AAE and ARE values of the training data set (n=48) are 1.09 bpm and 0.92% respectively, and the AAE and ARE values of the test data set (n=23) are 1.46 bpm and 1.23% respectively. Table 1 summarizes the performance. Specifically, the AAE and ARE values of the ISPC dataset are 0.76 bpm and 0.66% respectively, and the AAE and ARE values of the BAMI-I dataset are 1.39 bpm and 1.17% respectively.

$$AAE = \frac{1}{N} \sum_{i=1}^N |y_c^i - \tilde{y}_c^i| \quad (4)$$

$$ARE = \frac{\sum_{i=1}^N |y_c^i - \tilde{y}_c^i|}{\sum_{i=1}^N |\bar{y}_c^i - \tilde{y}_c^i|} \quad (5)$$

Table 1. Performance summary with each dataset: ISPC and BAMI-I are used as training set, and BAMI-II was used as test set

		AAE(bpm)		ARE(%)	
Training set	ISPC(n=23)	0.76	1.09	0.66	0.92
	BAMI-I(n=25)	1.39		1.17	
Test set	BAMI-II(n=23)	1.46		1.23	

Figure 2 shows the HR trajectory based on the proposed algorithm and acceleration intensity estimation; Figure 2 (a) and (b) show the results of the ISPC dataset, and Figure 2 (c) and (d) show the results of the BAMI-I and BAMI-II dataset. On the top panel of the figure.

The estimated HR trace results are compared with the results considering only the dominant power spectrum of PPG. In addition, it can be seen from the bottom panel of Figure 2 that the acceleration intensity is related to the increase of HR. In particular, Figure 2 (a) and (b) show the results of subjects 14 and 17. These two sets of data are considered to be the most challenging because the measured PPG signal is seriously damaged by MAs, resulting in a very low signal-to-noise ratio.

In fact, for topic 14 in the dataset, the reported AAEs are 6.63 bpm [1], 8.07 bpm [2], 7.29 bpm [3], 4.89 bpm [6], 9.59 bpm [7], 1.60 bpm [17], 12.12 bpm [4], 7.91 bpm [13] and 7.66 bpm [8]. For subject 17, the reported AAEs are 7.82 bpm [1], 7.01 bpm [2], 3.01 bpm [3], 3.05 bpm [6], 3.01 bpm [7], 2.04 bpm [17], 3.31 bpm [4], 2.44 bpm [13] and 2.77 bpm [8]. On the other hand, the algorithm I proposed provides very accurate estimates of the whole segmented HR for two subjects (AAE: 0.87 bpm for subject 14 and 1.49 bpm for subject 17). Figure 2 (c) and (d) also show the results when the signal-to-noise ratio is very low. For Topic 1 in the BAMI-I dataset, when only the dominant power spectrum of PPG is used, the AAE is 46.64 bpm, and the AAE of WFPV is 14.27 bpm. On the other hand, the results show that AAE is 1.45 bpm.

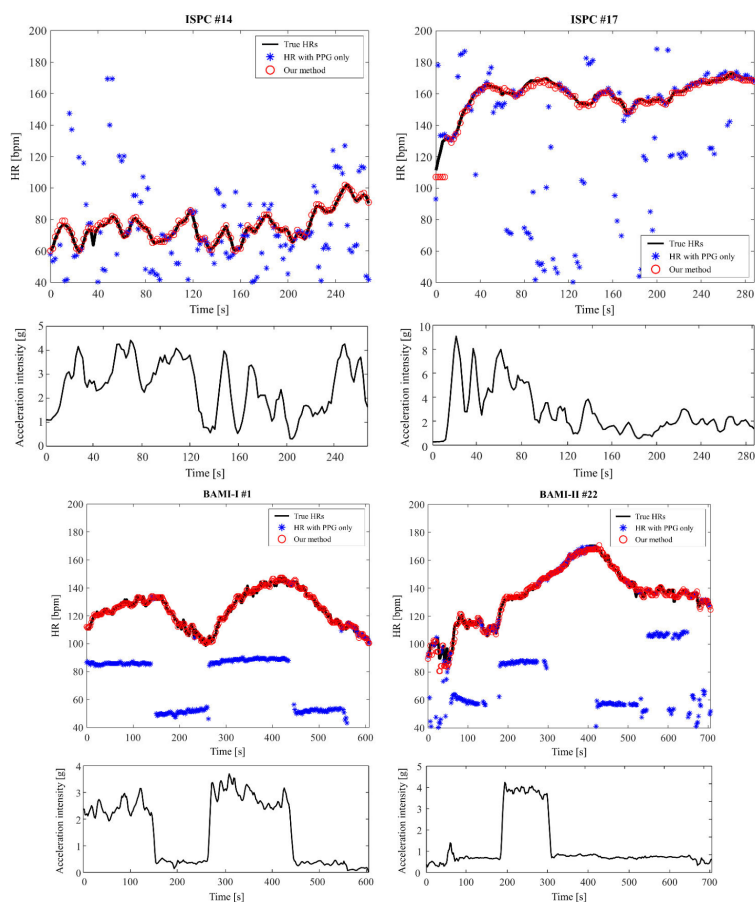


Figure 2. Estimated HR trace based on the proposed algorithm (Tops) and acceleration intensity (Bottoms); (a) Subject 14 (ISPC), (b) Subject 17 (ISPC), (c) Subject 1 (BAMI-I) and (d) Subject 22 (BAMI-II).

Conclusion.

The proposed deep learning algorithm estimates HR based on the power spectrum of PPG and acceleration signals and cascaded acceleration strengths. The algorithm consists of a two-dimensional convolutional layer, a one-dimensional convolutional layer, and a fully connected layer for MA cancellation. The algorithm also includes a connection layer, two LSTM layers, a fully connected layer, and a softmax layer for human resource tracking and estimation. The AAE of the algorithm for the training dataset and the test dataset is 1.09 bpm and 1.46 bpm, respectively, which exceeds the results of the most advanced methods currently available and allows for more accurate measurement of heart rate.

However, the most important aspect of future research will involve the implementation of energy saving when implementing the proposed algorithm in real time on wearable devices. Deep learning is undoubtedly to provide good performance, but the implementation of wearable devices faces many challenges, because they require low-power algorithms due to limited computing power.

References

- [1] Z. Zhang, Z. Pi, B. Liu. TROIKA: A general framework for heart rate monitoring using wrist-type photoplethysmographic signals during intensive physical exercise, *IEEE Trans. Biomed. Eng.* Feb. 2015. Vol. 62, no. 2, pp.
- [2] Z. Zhang. Photoplethysmography-based heart rate monitoring in physical activities via joint sparse spectrum reconstruction, *IEEE Trans. Biomed. Eng.* Aug. 2015. Vol. 62, no. 8, pp.
- [3] S. Salehizadeh, D. Dao, J. Bolkhovskiy, C. Cho, Y. Mendelson, K. Chon. A novel time-varying spectral filtering algorithm for reconstruction of motion artifact corrupted heart rate signals during intense physical activities using a wearable photoplethysmogram sensor, *Sensors*, 2016. Vol. 16, no. 1, p. 10.

- [4] S. Fallet, J.-M. V esin. Robust heart rate estimation using wristtype photoplethysmographic signals during physical exercise: An approach based on adaptive filtering, *Physiol. Meas.* Feb. 2017. Vol. 38, no. 2, pp.
- [5] M. Bolthesaz Mashhadi, E. Asadi, M. Eskandari, S. Kiani, F. Marvasti. Heart rate tracking using wrist-type photoplethysmographic (PPG) signals during physical exercise with simultaneous accelerometry, *IEEE Signal Process. Lett.* Feb. 2016. Vol. 23, no. 2, pp.
- [6] B. Sun, Z. Zhang. Photoplethysmography-based heart rate monitoring using asymmetric least squares spectrum subtraction and Bayesian decision theory, *IEEE Sensors J.* Dec. 2015. Vol. 15, no. 12, pp.
- [7] A. Temko. Accurate heart rate monitoring during physical exercises using PPG, *IEEE Trans. Biomed. Eng.* Sep. 2017. Vol. 64, no. 9, pp.
- [8] M. A. Motin, C. K. Karmakar, M. Palaniswami. PPG derived heart rate estimation during intensive physical exercise, *IEEE Access.* 2019. Vol. 7, pp. 56062–56069.
- [9] E. Khan, F. Al Hossain, S. Z. Uddin, S. K. Alam, M. K. Hasan. A robust heart rate monitoring scheme using photoplethysmographic signals corrupted by intense motion artifacts, *IEEE Trans. Biomed. Eng.* Mar. 2016. Vol. 63, no. 3, pp.
- [10] H. Lee, H. Chung, H. Ko, J. Lee. wearable multichannel photoplethysmography framework for heart rate monitoring during intensive exercise, *IEEE Sensors J.* Apr. 2018. Vol. 18, no. 7, pp.
- [11] V. Nathan, R. Jafari. Particle filtering and sensor fusion for robust heart rate monitoring using wearable sensors, *IEEE J. Biomed. Health Informat.* Nov. 2018. Vol. 22, no. 6, pp.
- [12] Y. Fujita, M. Hiromoto, T. Sato. PARHELIA: Particle filter-based heart rate estimation from photoplethysmographic signals during physical exercise, *IEEE Trans. Biomed. Eng.* Jan. 2018. Vol. 65, no. 1, pp.
- [13] A. Galli, C. Narduzzi, G. Giorgi. Measuring heart rate during physical exercise by subspace decomposition and Kalman smoothing, *IEEE Trans. Instrum. Meas.* May 2018. Vol. 67, no. 5, pp.
- [14] H. Chung, H. Lee, J. Lee. State-dependent Gaussian kernel-based power spectrum modification for accurate instantaneous heart rate estimation, *PLoS ONE*, Vol. 14, no. Art. no. e0215014.
- [15] H. Chung, H. Lee, J. Lee. Finite state machine framework for instantaneous heart rate validation using iearable photoplethysmography during intensive exercise, *IEEE J. Biomed. Health Informat.* Jul. 2019. Vol. 23, no. 4, pp.
- [16] J. Lee, H. Chung, H. Lee. Multi-mode particle filtering methods for heart rate estimation from iearable photoplethysmography, *IEEE Trans. Biomed. Eng.* Oct. 2019. Vol. 66, no. 10, pp.
- [17] D. Biswas, L. Everson. M. Liu, M. Panwar, B.-E. Verhoef, S. Patki, C. H. Kim, A. Acharyya, C. Van Hoof, M. Konijnenburg, and N. Van Helleputte. CorNET: Deep learning framework for PPG-based heart rate estimation and biometric identification in ambulant environment, *IEEE Trans. Biomed. Circuits Syst.* Apr. 2019. Vol. 13, no. 2, pp.
- [18] M. B. Mashhadi, M. Farhadi, M. Essalat, F. Marvasti. Low complexity heart rate measurement from iearable wrist-type photoplethysmographic sensors robust to motion artifacts, in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, Calgary, AB, Canada, Apr. 2018, pp. 921–924.

ОЦЕНКА ЧАСТОТЫ СЕРДЕЧНЫХ СОКРАЩЕНИЙ, ОСНОВАННАЯ НА ОБЪЕМНОЙ ФОТОГРАММЕТРИИ СВЕРТОЧНЫХ НЕЙРОННЫХ СЕТЕЙ И КРАТКОВРЕМЕННОЙ ПАМЯТИ И ДАННЫХ УСКОРЕНИЯ СМАРТФОНОВ

Цю Гаовей

*Аспирантура, Институт
информационно - коммуникационных
технологий*

*Департамент информационно-коммуникационных технологий Академии информационной безопасности
E-mail: qgw1314@163.com*

Аннотация. Носимый отражательный фотоконденсаторный объемный граф (PPG) датчик может быть интегрирован в часы или ремешок, чтобы обеспечить мгновенный сердечный ритм (HR), минимизируя неудобства для пользователя. Однако наличие спортивных псевдотеней (MA) приводит к неточным оценкам сердечного ритма. Чтобы решить эту проблему, я предложил новую нейронную сеть глубокого обучения для обеспечения точной оценки HR в высокоинтенсивных упражнениях. Метод: Для оценки HR была предложена глубокая нейронная сеть, основанная на множественной и неоднородной классификации с несколькими метками. Она включает в себя два слоя намотки, два слоя краткосрочной и долгосрочной памяти (LSTM), соединительный слой и три полных соединения, включая софтму, а также каскадный слой. Средняя абсолютная погрешность алгоритма для всех наборов данных обучения и наборов тестовых данных составляет менее 1,5 bpm, из которых набор данных обучения составляет 1,09 bpm, а набор тестовых данных - 1,46 bpm. Предложенные алгоритмы превосходят самые современные методы точной оценки частоты сердечных сокращений.

Ключевые слова: PPG, LSTM, свертки, каскады, пульс и движущиеся псевдотени

УДК [611.018.51+615.47]:612.086.2

CREATING OF INTEGRATED SPATIAL MODEL FOR 3D GIS



S.K. Kurbanov

*The head of department of Media design of journalism and mass communication
university of Uzbekistan
abstrakt88@gmail.com*

S.K. Kurbanov

Graduated Tashkent University of information technologies named after Muhammad al-Khwarizmi. The head of department of Media design of JMCU of Uzbekistan

Abstract. A three dimensional (3D) model facilitates the study of the real world objects it represents. A geo-information system (GIS) should exploit the 3D model in a digital form as a basis for answering questions pertaining to aspects of the real world. With respect to the earth sciences, different kinds of objects of reality can be realized. These objects are components of the reality under study. At the present state-of the art, different realizations are usually situated in separate systems or subsystems. This separation results in redundancy and uncertainty when different components sharing some common aspects are combined. Relationships between different kinds of objects, or between components of an object, cannot be represented adequately. This thesis aims at the integration of those components sharing some common aspects in one 3D model. This integration brings related components together, minimizes redundancy and uncertainty. Since the model should permit not only the representation of known aspects of reality, but also the derivation of information from the existing representation, the design of the model is constrained so as to afford these capabilities. The tessellation of space by the network of simplest geometry, the simple network, is proposed as a solution. The known aspects of the reality can be embedded in the simple network without degrading their quality. The model provides finite spatial units useful for the representation of objects. Relationships between objects can also be expressed through components of these spatial units which at the same time facilitate various computations and the derivation of information implicitly available in the model. Since the simple network is based on concepts in geo-information science and in mathematics, its design can be generalized for n-dimensions. The networks of different dimension are said to be compatible, which enables the incorporation of a simple network of a lower dimension into dimension into another simple network of a higher dimension. Despite the GIS (Geographic Information Systems/Geospatial Information Systems) have been provided with several applications to manage the two-dimensional geometric information and arrange the topological relations among different spatial primitives, most of these systems have limited capabilities to manage the three-dimensional space. Other tools, such as CAD systems, have already achieved a full capability of representing 3D data. Despite the GIS (Geographic Information Systems/Geospatial Information Systems) have been provided with several applications to manage the two-dimensional geometric information and arrange the topological relations among different spatial primitives, most of these systems have limited capabilities to manage the three-dimensional space. Other tools, such as CAD systems, have already achieved a full capability of representing 3D data. Despite the GIS (Geographic Information Systems/Geospatial Information Systems) have been provided with several applications to manage the two-dimensional geometric information and arrange the topological relations among different spatial primitives, most of these systems have limited capabilities to manage the three-dimensional space. Other tools, such as CAD systems, have already achieved a full capability of representing 3D data.

Key words: integrated spatial model, CAD systems, GIS, 3D modeling.

Introduction

The complexity of the 3D model full filling the requirements listed calls for a suitable construction method. The thesis presents a simple way to construct the model. The raster technique is used for the formation of the simple network embedding the representation of the known aspects of reality as constraints. The prototype implementation in a software package, ISNAP, demonstrates the simple network's construction and use. The simple network can facilitate spatial and non-spatial queries, computations, and 2D and 3D visualizations. The experimental tests using different kinds of data sets

show that the simple network can be used to represent real world objects in different dimensionalities. Operations traditionally requiring different systems and spatial models can be carried out in one system using one model as a basis. This possibility makes the GIS more powerful and easy to use [1].

1. Needs for 3D GIS

We live in a three dimensional (3D) world. Earth scientists and engineers have long sought graphic expression of their understanding about 3D spatial aspects of reality in the form of sketches and drawings. Graphical descriptions of 3D reality are not new. Drawings in perspective view date from the Renaissance period [2]. 3D descriptions of reality in perspective view change with the viewing position, so their creation is quite tedious. Traditional maps overcome this problem by using orthogonal projections of the earth. However, they offer a very limited 3D impression.

These traditional drawings and maps reduce the spatial description of 3D objects to 2D. Using computing technology, however, knowledge about reality can be directly transferred into a 3D digital model by a process known as 3D modeling.

A more suitable tool for earth science applications would be a GIS providing a 3D GIS. At the time of writing, a GIS capable of providing the functions in the above list with full 3D modeling capability is not commercially available. Most GISs still limit their geometric modeling capability to 2D so that the 3D representation, analysis and visualization provided by CAD are not possible. Most endeavours to model the third dimension can be found in the representation of terrain relief and in digital terrain models (DTM). DTM can facilitate spatial analyses related to relief, including slope, aspect, height zone, visibility, cut and fill volume, and surface area, and the 3D visualization of a surface, as in a perspective view. However, the basis of DTM is a continuous surface with a single height value for every planimetric location (Figure 1.a). DTM cannot accommodate a 3D (solid) object, or a surface with multiple height values at a given planimetric location (Figure 1.b and Figure 1.c).

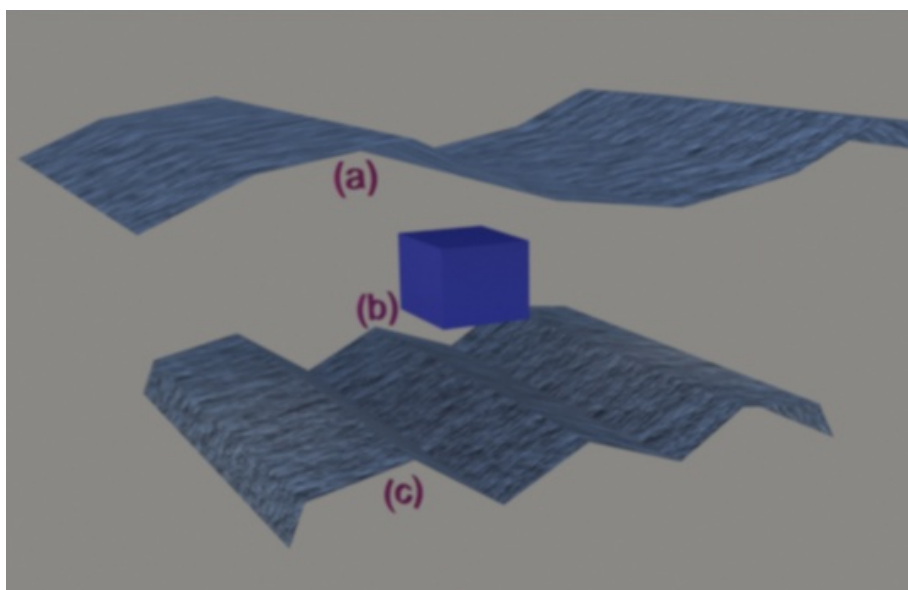


Figure 1. Single-value surface (a), 3D solid object (b) and multi-valued surface (c)

Although raster-based systems which could be regarded as 3D GISs are available, they may not be able to maintain the knowledge may be lost because of the problems of resolution and resampling. As a remedy, the original data set would have to be stored separately from the model, for example, for:

- Recreating the model if the result proves to be unsatisfactory because of unsuitable mathematical definition;
- Creating another model with different resolution;
- Merging with another data set to create a new model;

- Archiving as a reference to, or evidence of, the model.

These activities imply the need to store original data in an appropriate structure ready for future use. Necessary information about the data should be attached to each data element. In DTM for instance, information that a line is a break-line should be kept because it will have an impact on the interpolation. Similarly, other information can be attached which influence data handling strategies.

Since neither CAD nor GISs can at present fulfil the requirements of earth science applications, further research and development of a 3D GIS would seem appropriate.

In addition to the problem of creating a system capable of offering 3D modeling and functionality, there is a further problem concerning the type of 3D model chosen as the basis for 3D GIS. The model contains knowledge about reality, so we consider below the types of real world objects may be differentiated in terms of prior knowledge about their shapes and location, as shown in Figure 2. In reality, objects from the two categories coexist. Traditional GIS models the objects of each category independently with the result that two separate kinds of systems or subsystems have been developed.

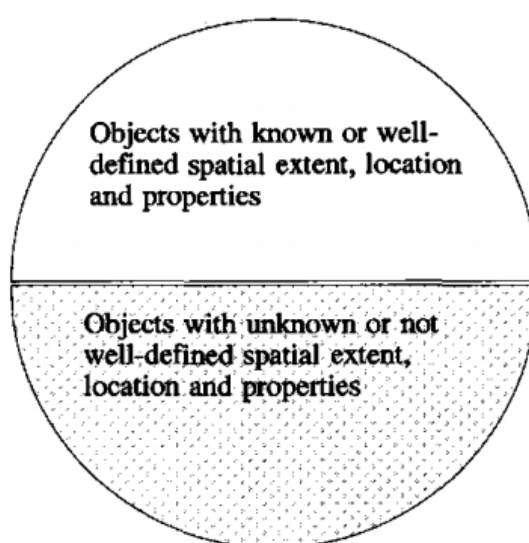


Figure 2. Two of real world objects with respect to their spatial exten

Raper [3] has also defined these two categories of objects. The first category, regarded as “sampling limited”, is for objects having discrete properties and readily determined boundaries, such as buildings, roads, bridges, land parcels, fault blocks, perched aquifers.

The second category, known as “definition limited”, is for objects having various properties that can be classified by grain-size distribution; moisture content, colloid or pollutant in the water by percentage ranges; carbon monoxide in the air by concentration ranges, and so forth.

Separate modelling of these two categories of objects tends to contradict the reality, which leads to difficulties in representing their relationships. Such a question as, “How many of the people working in a 50-storey office building are affected by polluted air generated by vehicles in nearby streets during rush hours”; cannot be answered until the two separate models are combined, as shown in Fig 3. Modeling them together with more accurate representation of their relationships in the 3D environment requires the integrated 3D modeling forming the general aim of this thesis.

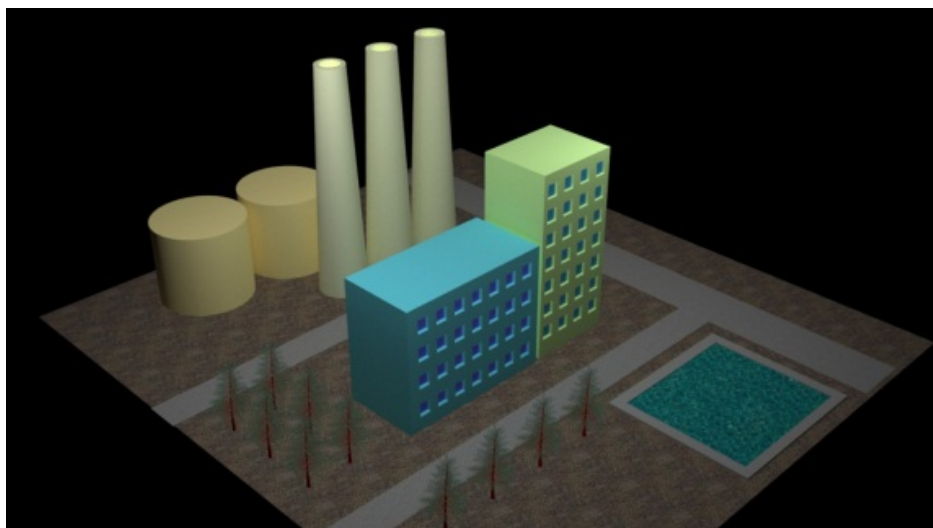


Figure 3. An example of two types of real world objects (1. Objects with discernible boundaries: roads, buildings. 2. Objects with indiscernible boundaries: moisture content of air, colloid or pollutant in the water by percentage ranges, toxins in the air)

Note also that the properties of an object may be well defined in some specific dimensions and defined in others. For example, given a DTM data set representing a surface, the planimetric extent of regions at the elevation of 100 metres above mean sea level cannot be defined until the result of interpolation based on a mathematical definition (linear interpolation) is obtained. That is to say, although the spatial extent of this region may be known in the z-dimension, the spatial extent in planimetry (x,y) has still to be discovered. The model must contain the aspect allowing the appropriate operation, such as interpolation or classification, if the required description of the properties of an object is to be obtained.

Apart from the problem of the separate modelling of the two types of objects, there remains the further problem of the separate modelling of an object's components. These components are relief and planer geometry associated with thematic properties.

This separation has resulted in independent systems and data structures, DTM and 2D GIS, respectively. The consequences are data redundancy, which may lead to uncertainty when the two data sets are combined and only one data set has been updated.

Spatial modeling is an analytical process conducted in conjunction with a geographical information system (GIS) in order to describe basic processes and properties for a given set of spatial features.

The objective of spatial modeling is to be able to study and simulate spatial objects or phenomena that occur in the real world and facilitate problem solving and planning. Spatial modeling is an essential process of spatial analysis. With the use of models or special rules and procedures for analyzing spatial data, it is used in conjunction with a GIS to properly analyze and visually lay out data for better understanding by human readers. Its visual nature helps researchers more quickly understand the data and reach conclusions that are difficult to formulate with simple numerical and textual data.

Manipulation of information occurs in multiple steps, each representing a stage in a complex analysis procedure. Spatial modeling is object-oriented with coverage and concerned with how the physical world works or looks. The resulting model represents either a set of objects or real-world process.

For example, spatial modeling can be used to analyze the projected path of tornadoes by layering a map with different spatial data, like roads, houses, the path of the tornado and even its intensity at different points. This allows researchers to determine a tornado's real path of destruction. When juxtaposed with other models from tornadoes that have affected the area, this model can be used to show path correlations and geographical factors.

In spatial modeling, the Bayesian approach has two main advantages above the frequentist methods. First, it is not based on approximation (like the penalized quaslikelihood) and thus provides exact results even for binary responses. Second, it correctly propagates the uncertainties linked to the estimation of the variogram parameters. This is not done by the frequentist approach, which uses a two-step approach: it first estimates the variogram parameters and then assumes them as known and plugs them into the linear model [4][5].

Temporal and spatial modeling of extreme precipitation in urban areas is a major challenge due to sparse data availability and huge spatial non-uniformity in precipitation. High uncertainties are associated with the short-duration precipitation events, which need to be modeled and further to be considered in the design and risk analyses. Bayesian methods provide a comprehensive way to quantify the uncertainties in the precipitation data [6]. The purpose of this study was spatial modeling of land subsidence using an RF data-mining model in Kerman Province, Iran. In order to investigate the spatial relationship between effective factors and subsidence locations, an FR bivariate statistical model was used. Also, in considering the importance of effective factors, an RF algorithm was applied. For spatial modeling, from the total subsidence locations, 70% were used for calibration and 30% for model evaluation. Validation of results showed that the RF model with an AUC value of 0.939 has excellent precision.

Therefore, based on the results of spatial modeling of a land-subsidence and susceptibility map, and also with regard to the correlation between effective factors and subsidence areas, it is possible to mitigate this hazard. Also, we can plan for managing water resources, in particular unauthorized wells and preventing extraction of groundwater which was carried out in the study area [7].

2. Importance for geo-information

In the context of earth science, the end product we seek is a model in the sense of a replica of some portion of the planet earth, and is called a geo-spatial model. Since the term “spatial model” covers a large territory over many disciplines (like the modeling of human anatomy in medicine, molecular structure in chemistry, or atomic structure in nuclear physics), we add the prefix geo to indicate the scope and purpose of this earth-related model.

For the information system to utilize the geo-spatial model, it must be constructed in digital form, so that it can be maintained and exploited by a computer to perform certain tasks or operations that are:

Less convenient in reality; for example, a distance can be obtained from a model instead of measuring from place A to B in reality, provided that places A and B are represented in the model;

Too expensive, too difficult, or practically impossible in reality; for example, a geologist may wish to see the continuous layer of sandstone lying fifteen meters under the earth’s surface; removal of the upper soil to see this layer in reality is too expensive to contemplate.

A single-purpose model represents only single view of the reality (Figure 4).

An integrated model represents various views of the reality, so the integrated model may be considered to be of higher value, since it contains more aspects of the reality and may serve more purposes.

Before continuing this review of necessary fundamental concepts, the steps followed in geo-spatial modelling are defined. Obtaining a geo-spatial model requires two main steps: the design phase, and the construction phase (Figure 4).

Once the model is in place, maintenance forms an additional phase. The design phase includes all the abstraction processes, ranging from the conceptual design, the logical design, to the physical design. The product of the conceptual design is referred to as a conceptual model, or data model. It comprises a general scheme describing what should be included in the model.

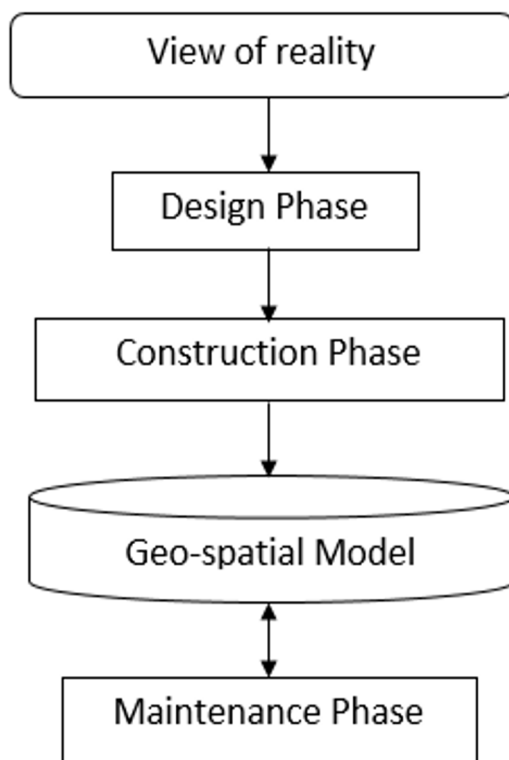


Figure 4. Geo-spatial modeling

3. Conceptual design of a geo-spatial model

The design phase deals with the abstraction of reality into the representation scheme. This phase answers two basic questions: what aspects of reality (real objects and the relationships between them) are to be modelled: how should they be represented in the model?

A geo-spatial database represents a state of reality from a specific point of view or interest at an instant in time (if the temporal aspect itself is not the subject of the model). The reality consists of a set of various objects and the relationships between them which should be capable of representation as components of the model described in the preceding section. To be manageable, it is necessary to determine a limited number of aspects of aspects of the reality (objects together with the relationships between them) during the design phase which can be represented as the first and second components of the model (Figure 5).

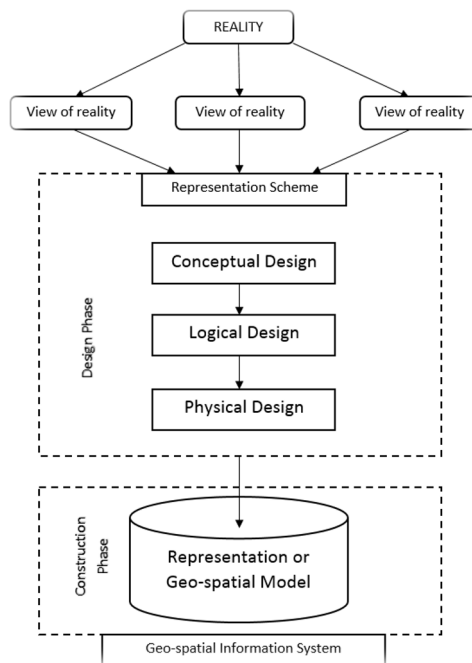


Figure 5. Design and construction phases for geo-spatial model

The conventions must operate in every design phase. In the conceptual design phase, the conventions should state the allowable type of objects and relations between them to be included in the model. In the logical design phase, the conventions should state how the representation of one objects is distinguished from another; an object should have a unique identifier. In the physical design phase, the conventions comprise a set of integrity and consistency rules for the operations that may change the state of the model; for example, the union of two areas sharing a common boundary has to yield only one area. The design of the model is followed by the design and implementation of the necessary functions and the user-interface to enable the construction and exploitation of the model. The result of this implementation is a geo-spatial information system (GIS). Having constructed the model, it must be kept valid to ensure it remains in a state comparable with the reality, which is dynamic in nature. This is the maintenance phase. The basic maintenance operations of insertion, deletion, and modification can be applied to any component of the model, that is to say object types, relations, rules, attributes and operations. A GIS should also provide functionality to maintain the geo-spatial model.

4. Spatial relations

Spatial relations are key issue in the design of a spatial model. Many extensive reviews and discussions can be found Frank and Kuhn [8], Pullar [9], Egenhofer [10], Pigot [11]. In the following paragraph, only a brief review is therefore given of some important basic concepts about spatial relations. Recall that R is a relation on a set O of objects, In general, R can be further distinguished by its different basic properties that depend on the relationships between its member elements. R is reflexive, if each element can be compared with itself (if and only if $(o1, o2) \in R$), for example 'point A' is equal to itself. R is symmetric, if and only $R(o1, o2)$ implies $R(o2, o1)$. For example 'area A' is adjacent to 'area B' implies that 'area B' is adjacent to 'area A'.

R is anti-symmetric, if and only if $R(o1, o2)$ and $R(o2, o1)$ implies $o1 = o2$ for all $o1, o2 \in O$, for example if $a \leq b$ and $b \leq a$, then $a = b$.

R is transitive, if and only if $R(o1, o2)$ and $R(o2, o3)$ implies $R(o1, o3)$ for all $o1, o2, o3 \in O$, for example area A < area B and area B < area C then area A < area C.

For example, given a set of real number N , $<$ is a transitive relation on N , \leq is a reflexive, anti-symmetric, transitive relation on N , and \neq is a symmetric relation on N . It is necessary at this stage to

consider the definition of functions in mathematics used later.

Given two sets A and B, a function (or map) f from A to B, denoted $f: A \rightarrow B$, is a subset of the Cartesian product $A \times B$ with the following properties:

For each $a \in A$, there is some $b \in B$ such that $(a, b) \in f$,

If $(a, b) \in f$ and $(a, c) \in f$, then $b=c$.

Each $a \in A$ must be in relationship with exactly one $b \in B$ and the relationship $(a, b) \in f$ is normally written in a prefix form as $b = f(a)$.

Comparing with relation R, every function on A is a relation R on A. However, not all relations on A are functions. Three classes of spatial relations, namely metric, order and topology, have been distinguished, based on the type of function or relation associated with a set of objects. Metric relations are built around the notion of distance function. Its mathematical description is as follows:

Given a set M with $x, y, z \in M$ and a set of real number N. A metric relation d is a function $d: M \times M \rightarrow N$ with the following conditions:

$d(x, y) \geq 0$, Distance from x to y is more than or equal to zero.

$d(x, x) = 0$; $d(x, y) = 0$ implies $x = y$, Distance from x to itself equal to zero. Distance from x to y equal to zero implies that x is equal to y.

$d(x, y) = d(y, x)$, Distance from x to y equal to distance from y to x.

$d(x, y) + d(y, z) \geq d(x, z)$ (triangular inequality). Distance from x to y plus distance from y to z is more than or equal to distance from x to z.

A metric space is an ordered pair (M, d) consisting of set M together with a function $d: M \times M \rightarrow N$ satisfying the above four conditions. The function d is also called the metric on M. Functions $d: M \times M \rightarrow N$ are called distance functions. A metric space is the Euclidean n-space (1), denoted R^n , is the distance function below:

$$d((x_1, \dots, x_n), (y_1, \dots, y_n)) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \quad (1)$$

The number 'n' defines the number of distance components between x and y (each one computed along an independent vector) and denotes the dimensionality of Euclidean space. The distance functions available in metric spaces are used to develop the notion of continuity crucial for the development of topology.

Conclusion.

In general, a 3D GIS should aim to integrate all the necessary elements of the spatial model and functions to create and utilize the spatial model efficiently. This research therefore analyses the definition of the direction of the related research and stimulate the development of 3D GIS that can readily be used. The research reviews some general aspects of systems for integrated 3D geo-information, their functional components, and some related and available technological developments to support their functionally with respect to the user's perspective.

References

- [1] Devlin, K., 1994, Mathematics: The sciences of patterns, Scientific American Library New York, 215 pp.
- [2] Raper, J. F., 1989, The 3-dimensional geoscientific mapping and modeling system: a conceptual design, Taylor and Francis, London.
- [3] Fränzi Komer-Nievergelt, 2015, Modeling Spatial Data Using GLMM, Bayesian Data Analysis in Ecology Using Linear Models, ISBN 978-0-12-801370-0
- [4] Chandra R. Rupa, P.P. Mujumdar, 2018, Modeling High-Intensity Precipitation for Urban Hydrologic Designs. Trends and Changes in Hydroclimatic Variables, ISBN 978-0-12-810985-4
- [5] Hamid Reza Pourghasemi, Mohsen Mohseni Saravi, 2019, Land-Subsidence Spatial Modeling Using the Random Forest Data-Mining Technique, Spatial Modeling in GIS and R for Earth and Environmental Sciences.
- [6] Frank, A.U., and Kuhn, W., 1986, Cell graphs: a provable correct method for the storage geometry, Proceedings of the Second international symposium on spatial data handling, Seattle, Washington, USA, pp 411-436.
- [7] Pullar, D., 1988, Data definition and operators on a spatial data model, ACSM-ASPRS, Annual convention, Vol.2, pp. 196-202.
- [8] Egenhofer, M.J., 1990, Interaction with geographic information systems via spatial queries, Journal of visual languages and computing, Vol.1, No.4, pp. 389-413.

[9] Pigot, S., 1992, A topological model for a 3-D spatial information system, Proceedings of 5th International Symposium on spatial information handling, Charleston, S.C., pp. 344-360.

СОЗДАНИЕ ИНТЕГРИРОВАННОЙ ПРОСТРАНСТВЕННОЙ МОДЕЛИ ДЛЯ 3D ГИС

С. К. Курбанов

*Заведующий кафедрой медиа дизайна
журналистики и массовых коммуникаций Университета Узбекистана*

*Кафедра медиа дизайна журналистики и массовых коммуникаций Университета Узбекистана
E-mail: abstrakt88@gmail.com*

Аннотация. Трехмерная (3D) модель облегчает изучение объектов реального мира, которые она представляет. Геоинформационная система (ГИС) должна использовать трехмерную модель в цифровой форме как основу для ответов на вопросы, относящиеся к аспектам реального мира. Что касается наук о Земле, могут быть реализованы различные виды объектов реальности. Эти объекты являются составляющими исследуемой реальности. На современном уровне техники различные реализации обычно находятся в отдельных системах или подсистемах. Это разделение приводит к избыточности и неопределенности при объединении различных компонентов, имеющих некоторые общие аспекты. Отношения между разными типами объектов или между компонентами объекта не могут быть адекватно представлены. Этот тезис направлен на интеграцию тех компонентов, которые имеют некоторые общие аспекты в одной 3D-модели. Эта интеграция объединяет связанные компоненты, сводя к минимуму избыточность и неопределенность. Поскольку модель должна позволять не только представление известных аспектов недвижимости, но также получение информации из существующего представления, дизайн модели ограничен, чтобы предоставить эти возможности. В качестве решения предлагается тесселяция пространства сетью простейшей геометрии, простой сетью. Известные аспекты недвижимости могут быть встроены в простую сеть без ухудшения их качества. Модель предоставляет конечные пространственные единицы, полезные для представления объектов. Отношения между объектами также могут быть выражены через компоненты этих пространственных единиц, которые в то же время облегчают различные вычисления и получение информации, неявно доступной в модели. Поскольку простая сеть основана на концепциях геоинформатики и математики, ее конструкция может быть обобщена для n-мерных измерений. Сети разных измерений считаются совместимыми, что позволяет включать простую сеть более низкого измерения в размерность в другую простую сеть более высокого измерения.

Ключевые слова: интегрированная пространственная модель, САПР, ГИС, 3D моделирование.

UDC [611.018.51+615.47]:612.086.2

PROFILING OF ENERGY CONSUMPTION BY ALGORITHMS OF SHORTEST PATHS SEARCH IN LARGE DENSE GRAPHS



O.N. Karasik

Technology Lead at ISsoft Solutions (part of Coherent Solutions) in Minsk, Belarus, PhD in Technical Science
prihozhy@yahoo.com



A.A. Prihozhy

Professor at the Computer and System Software Department, Doctor of Technical Sciences, Full Professor Belarusian National Technical University
karasik.oleg.nikolaevich@gmail.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, D. Sc. (Eng) (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: Reducing power consumption when solving computationally intensive problems is an important applied problem in science and industry. The paper presents the results of profiling four algorithms of finding the shortest paths between all pairs of graph vertices, which allowed us to estimate the power consumption and execution time of the algorithms on a multicore system. Profiling was performed on single-threaded implementations of the classical Floyd-Warshall algorithm, the algorithm based on vertex expansion of the graph, the homogeneous Floyd-Warshall block algorithm, and the heterogeneous block algorithm. The experiments used simple complete, oriented weighted graphs ranging in size from 1200 to 4800 vertices. Profiling performed via Intel VTune and Intel SoC Watch showed that the algorithm itself has the largest impact on power consumption. On graphs up to 1200 vertices, the power consumption is not proportionally dependent on the algorithm's execution time. For example, the slow classical Floyd-Warshall algorithm has consumed up to 20 % less energy compared to the faster block algorithms. As the graph size increases, the algorithm based on vertex expansion of the graph becomes strictly dominant; it consumed up to 25 % less energy than the blocked algorithms. With increasing the graph size, a proportional relationship between the algorithm execution time and the amount of energy consumed has been clearly established.

Keywords: multicore processor; profiling; large size graph; shortest path search algorithm; energy consumption; runtime; optimization.

Problem formulation

Software profiling [1 – 3] aims for dynamic program analysis and measurement of the program's time and memory complexity, power and energy consumption, usage of CPU instructions, duration of function calls, etc. Profiling is carried out by a profiler which instruments program source code or binary executable form. The objective of profiling is to gather information for program optimization. The tools of software analysis are important for understanding the program behavior and the utilization of computational resources. The results of software profiling can be used by developers of algorithms and computer programs, computer architects, compiler writers and consumers of software and computer systems. By means of the profiling techniques and tools, it is possible to choose the preferable programming language, compiler, algorithm, software tool, multi-core processor, multi-processor computing system accounting for time, space and power parameters.

In the paper, we have chosen as a benchmark the problem of searching for shortest paths between all pairs of vertices in a large dense graph and have chosen four competitive algorithms of solving the problem of finding optimal solutions on a multi-core system with respect to the time and power consumption parameters. The problem and algorithms have many application domains and are of high commercial importance, therefore, we need to know which combination of algorithm and its parameters yield the lowest runtime and lowest power consumption by the multi-core system. Intel VTune Profiler 2023.0 [1, 2] is chosen as a tool for profiling and measuring the algorithm and program parameters.

All pairs shortest path algorithms

All four competitive algorithms that are selected for profiling target simple complete and dense directed graphs with real edge-weights and without cycle having negative sum of weights. The graphs provide high computational load of the algorithms during profiling, which gives the correct comparison of the algorithms and the comparison of their software implementations.

Floyd-Warshall all-pairs shortest path algorithm (FW). Let $G = (V, E)$ be a simple directed graph with real edge-weights consisting of a set V , $|V| = N$, of vertices and a set E of edges. Let W be the cost adjacency matrix for G : $w_{i,i} = 0$, $1 \leq i \leq N$; $w_{i,j}$ is the cost (weight) of edge (i, j) if $(i, j) \in E$; $w_{ij} = \infty$ if $i \neq j$ and $(i, j) \notin E$. When G has no cycle with negative sum of weights, the dynamic programming Floyd-Warshall algorithm (FW) computes [4 – 7] a series of distance matrices $D^0 \dots D^k \dots D^N$ such that $D^0 = W$ and each element $d_{i,j}^k$ of matrix D^k , $k = 1 \dots N$, is the length of the shortest path from i to j composed of the subset of vertices labelled 1 to k . The algorithm built on matrix D consists of three nested loops along variables k , i and j , which perform the same kind of calculations on all elements of the matrix. Each iteration along k accesses all elements of D . Totally, every element has N attempts to be updated. FW does not provide spatial and temporal data reference locality. The number of loops' iterations is N^3 no matter what the number of edges in the graph is.

Graph-expansion-based all-pairs shortest path algorithm (GEA). The graph-expansion-based algorithm [8] of searching for shortest paths recalculates iteratively the lengths of shortest paths at each step of adding a vertex k to graph G and adding row k and column k to matrix D . The graph-size is changed from 1 to N , and the size of matrix D is changed from 1×1 to $N \times N$. Two operations are used to carry out the changes: adding vertex k with calculating the shortest paths from vertices $1, \dots, k-1$ to k and from k to $1, \dots, k-1$; updating shortest paths between pairs of vertices $1, \dots, k-1$ with accounting for paths passing through k . The operations are used for creating an initial version of pseudocode. The algorithm is inferred by transforming the initial pseudo-code. Formal methods are used to do this: the resynchronization of calculations, reordering of operations, merging loops, etc. The main advantage of GEA is the increased temporal and spatial locality of data processing. It is obtained since the algorithm operates on D 's submatrices of monotonically increasing size. GEA reduces the number of loops' iterations and makes lower the pressure on the processor caches.

Blocked Floyd-Warshall all-pairs shortest path algorithm (BFW). The authors of [9 – 14] proposed a blocked version BFW of the Floyd-Warshall algorithm FW. BFW divides set V of vertices into subsets $V_0 \dots V_{M-1}$ of size S and splits matrix D into blocks of size $S \times S$ each, creating a block-matrix $B[M \times M]$, where equality $M \cdot S = N$ holds. The main loop of BFW performs M iterations, S times less than FW. Each iteration consists of three phases: calculation of diagonal block (m, m) , which accounts for paths inside the subgraph on subset V_m of vertices; calculation of $(M - 1)$ cross blocks on column m through the diagonal block, which accounts for paths from vertices of V_v to vertices of V_m ; calculation of $(M - 1)$ cross blocks on row m through the diagonal block, which accounts for paths from vertices of V_m to vertices of V_v ; calculation of $(M - 1)^2$ peripheral blocks on the cross of row v and column u through other blocks m on the row and on the column, which accounts for paths from vertices of V_v to vertices of V_u passing through vertices of V_m . The single block calculation algorithm which implements FW computes all four types of blocks. The algorithm exploits spatial (sequential) data locality within each block, which increases the efficiency of the hierarchical cache memory operation. Its main loop has M

iterations, S times less compared to FW . Therefore, every iteration updates each element of the matrix as many as S times, performing the update locally by using one to three blocks simultaneously.

Heterogeneous blocked Floyd-Warshall all-pairs shortest path algorithm (HBFW). Starting from the homogeneous blocked algorithm, the heterogeneous shortest paths algorithm [15 – 22] distinguishes four types of blocks: diagonal, vertical of cross, horizontal of cross, and peripheral. To speed up the computations, new separate algorithms for all block types have been developed, which reduce the number of iterations in nested loops and account for the sequential reference locality of data in CPU caches. The algorithms are faster than the single block calculation algorithm due to accounting for features of the four block types and the data dependences between the blocks. They improve the spatial and temporal reference locality in processing large graphs.

It is interesting that all four algorithms have the same computational complexity. They differ each other by the way and the efficiency of utilization of the multi-core system computational resources, in particular caches.

Profiling tool

We used Intel VTune Profiler 2023.0 (and build in Intel SoC Watch utility) to measure energy consumption [1]. Intel SoC Watch is a command line tool for monitoring metrics related to power consumption on Intel architecture platforms. It can report power states for the system/CPU/GPU/devices, processor frequencies and throttling reasons, total energy consumption over a period, power consumption rate, and other metrics that provide insight into the system's energy efficiency. Intel SoC Watch collects data from both hardware and operating system with low overhead [2]. In our experiments, we focused on the measurement of energy consumption of the CPU package for a full duration of the algorithm execution. The energy consumption of the CPU package includes the energy that is consumed by all cores, by the each-core-separate L1 and L2 private caches, by the shared L3 cache and by other hardware components included into the CPU package. The energy consumption is measured in millijoules (mJ). We used the following parameters of the profiling tool:

```
socwatch.exe -f power -f hw-cpu-pstate -o <output> --program <algorithm application>
```

The switch “*-f power*” configures profiling of energy and power metrics; “*-f hw-cpu-pstate*” configures hardware profiling of CPU P-state residence information (the collected information is not reported in the paper); “*-o <output>*” and “*--program <algorithm application>*” switches configure application to store profiling results into *<output>* and to collect information about *<algorithm application>* execution.

Experimental results

All experimental runs of the four shortest path algorithms and their program implementations were carried out on a desktop computer equipped with an Intel Core i7-10700 CPU processor which contains 8 cores (16 hardware threads) with the support of “Intel Turbo Boost 2.0”, “Intel Turbo Boost Max 3.0” and “Enhanced Intel SpeedStep” technologies. Every core is equipped with a private L1 (512KB), L2 (2MB) caches and shared L3 (16MB) cache. Its base frequency is 2.90 GHz. Due to the active “Intel Turbo Boost 2.0 Technology” the frequency can increase up to 4.70 GHz, and due to the active “Intel Turbo Boost Max 3.0 Technology” it can increase up to 4.80 GHz. The algorithms were implemented in the C++ language using GNU GCC compiler v12.2.0.

We conducted a series of experiments on multiple randomly generated complete directed weighted graphs of 1200, 2400, 3600 and 4800 vertices. Every experiment was repeated several times and results were verified against the results of original Floyd-Warshall algorithm executed on the same graphs. All runs of blocked algorithms were done on multiple block sizes: 30x30, 48x48, 50x50, 75x75, 100x100, 120x120, 150x150, 200x200, 240x240, 300x300, 600x600, 1200x1200 and 2400x2400. All block sizes divide the matrix into equal blocks without remainders (for smaller graphs some of the block sizes were omitted). To understand the energy efficiency related to the discussed algorithms we measured the execution time, the average power and the total energy consumed by the processor package. In all the experiments, the results are represented with figures and pictures corresponding to algorithms FW , GEA ,

BFW and *HBFW*. Figures 1–4 depict the execution time (*ms*) and energy consumption (*mJ*) depending on the block size (and for non-block algorithms the values are repeated across all the block sizes).

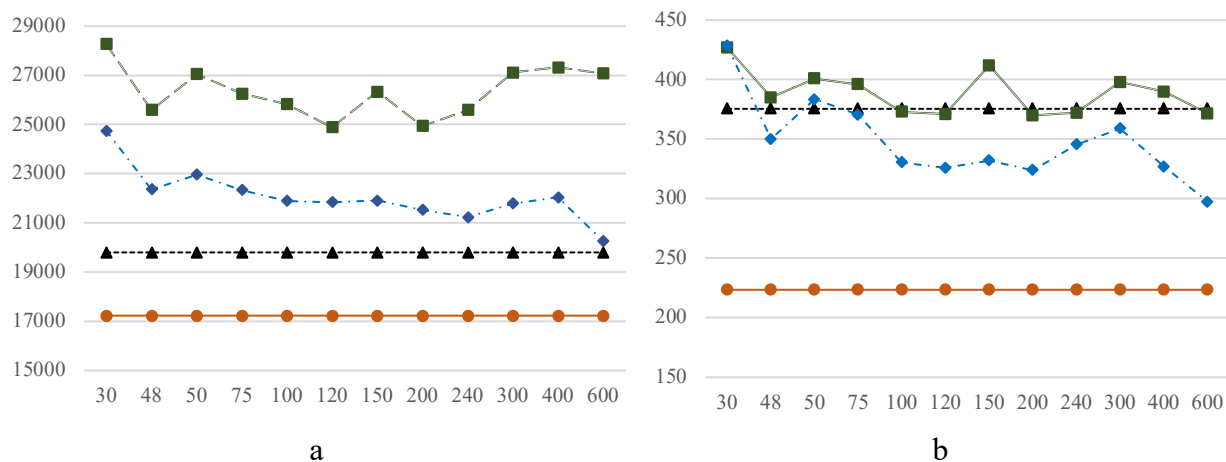


Figure 1. a) Energy consumption (*mJ*) and b) execution time (*ms*) of *FW* (triangle and dash line), *GEA* (circle and solid line), *BFW* (square and long-dash line) and *HBFW* (diamond and dash-dot line) algorithms across all the block sizes on the graph of 1200 vertices

On the graphs of 1200 vertices (Figure 1), algorithm *GEA* demonstrated the lowest energy consumption (13 % better than *FW*) and the lowest execution time (25 % faster than *HBFW*) against all algorithms under consideration. Both block algorithms *BFW* and *HBFW* demonstrated higher energy consumption than *FW* by 25 % and 2 % respectively. In terms of runtime, *FW* and *BFW* showed the comparable execution time, while *HBFW* outranked both showing by 21 % lower execution time.

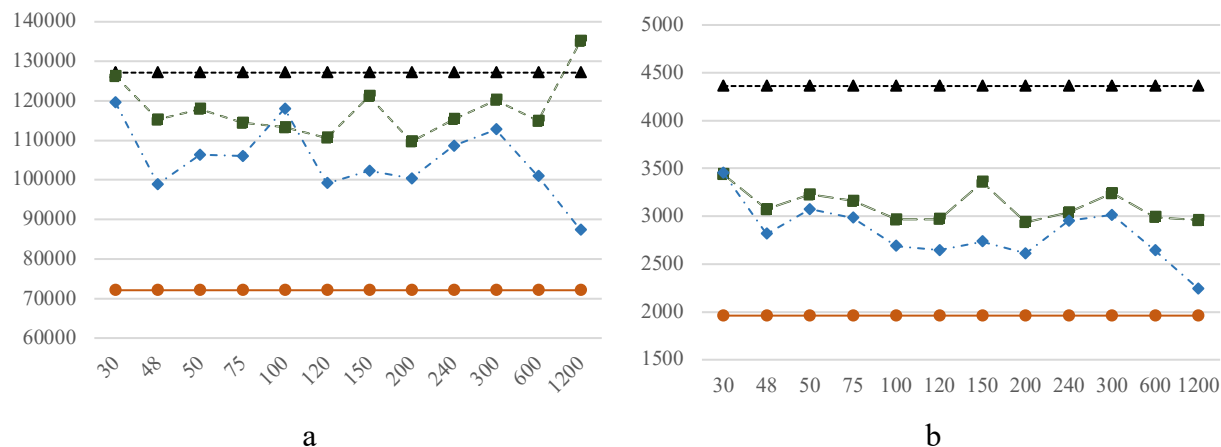


Figure 2. a) Energy consumption (*mJ*) and b) execution time (*ms*) of *FW* (triangles and dash line), *GEA* (circle and solid line), *BFW* (squares and long-dash line) and *HBFW* (diamonds and dash-dot line) algorithms across all the block sizes on the graph of 2400 vertices

On larger graphs of 2400 vertices (Figure 2), the situation is slightly changed. The *GEA* algorithm remains the best one in both energy consumption and execution time. However, the blocked algorithms *BFW* and *HBFW* demonstrated lower energy consumption (by 14 % and 32 % respectively) and smaller

execution time (by 33 % and 49 % respectively) against *FW*. The comparison of *BFW* against *HBFW* shows that *HBFW* has lower energy consumption and lower execution time over *BFW*.

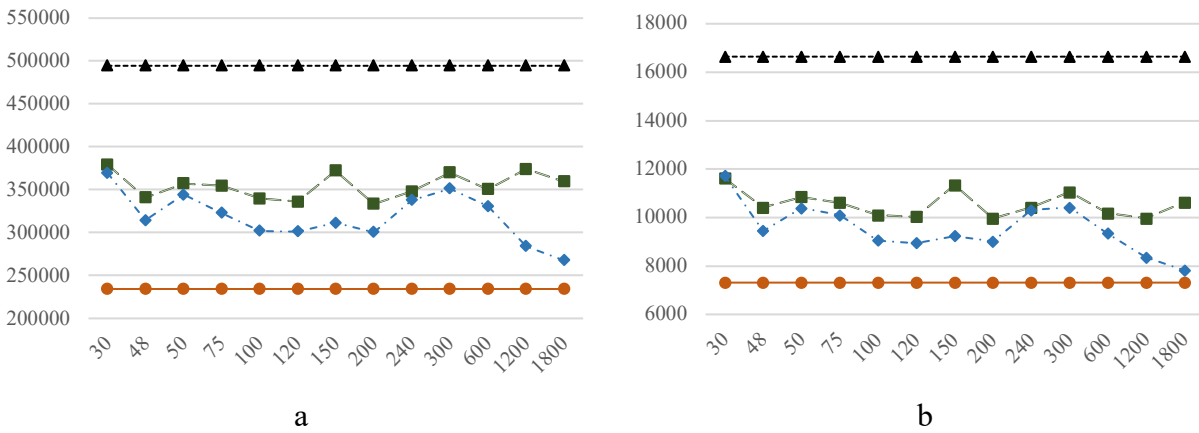


Figure 3. a) Energy consumption (mJ) and b) execution time (ms) of *FW* (triangles and dash line), *GEA* (circle and solid line), *BFW* (squares and long-dash line) and *HBFW* (diamonds and dash-dot line) algorithms across all the block sizes on the graph of 3600 vertices

On graphs of 3600 and 4800 vertices (Figure 3 and Figure 4), the algorithms demonstrated the same patterns of behavior. *GEA* remains the best one among all the algorithms under profiling. Both *BFW* and *HBFW* showed consistently lower energy consumption and lower execution time than *FW*. *HBFW* continues to demonstrate the reduction of energy consumption and execution time against *BFW*.

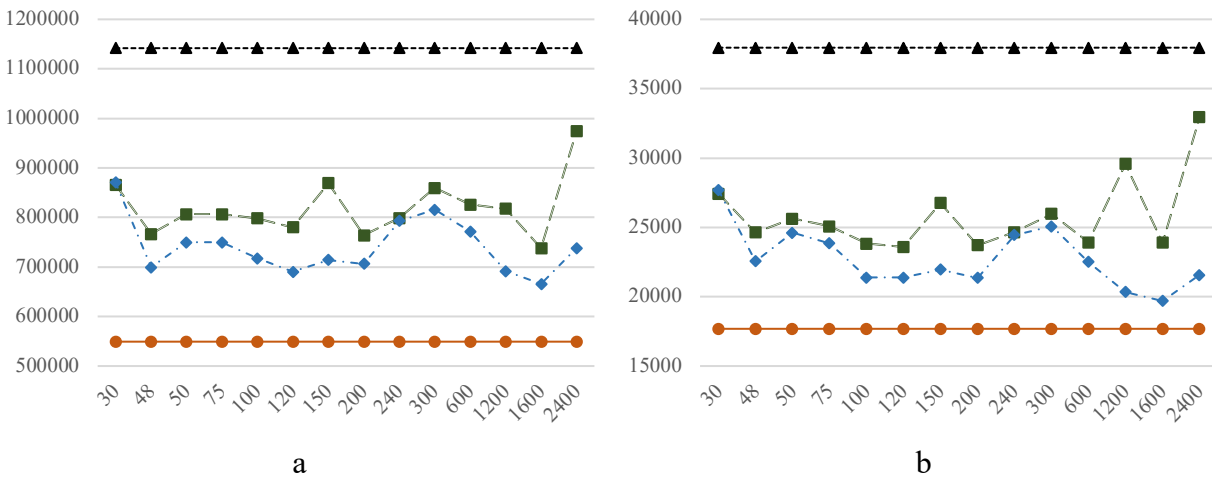


Figure 4. a) Energy consumption (mJ) and b) execution time (ms) of *FW* (triangles and dash line), *GEA* (circle and solid line), *BFW* (squares and long-dash line) and *HBFW* (diamonds and dash-dot line) algorithms across all the block sizes on the graph of 4800 vertices

Table 1 and Table 2 report the data which summarizes the information presented in Figures 1 – 4 and compares the best runs of the *GEA*, *BFW* and *HBFW* algorithms over the *FW* algorithm in terms of energy consumption and execution time. We can observe that *GEA* significantly outranks all other algorithms, and *HBFW* is preferable against *BFW*. *BFW* gives better results at smaller sizes of blocks, and *HBFW* gives better results at larger block-sizes.

Table 1. Comparison of best runs (on lowest energy consumption) of *GEA*, *BFW* and *HBFW* algorithms against *FW* algorithm along all experimental graphs in terms of energy consumption and execution time.

Graph Size	1200			2400			3600			4800		
Algorithm	GEA	BFW	HBFW	GEA	BFW	HBFW	GEA	BFW	HBFW	GEA	BFW	HBFW
Block Size		120	600		200	1200		200	1800		1600	1600
Execution Time (ms)	0.595	0.988	0.792	0.450	0.673	0.514	0.439	0.599	0.470	0.466	0.630	0.519
Energy Consumed (mJ)	0.870	1.257	1.023	0.567	0.863	0.687	0.474	0.674	0.541	0.481	0.646	0.583

Table 2. Comparison of best runs (on lowest execution time) of *GEA*, *BFW* and *HBFW* algorithms against *FW* algorithm along all experimental graphs in terms of energy consumption and execution time.

Graph Size	1200			2400			3600			4800		
Algorithm	GEA	BFW	HBFW	GEA	BFW	HBFW	GEA	BFW	HBFW	GEA	BFW	HBFW
Block Size		200	600		200	1200		200	1800		120	1600
Execution Time (ms)	0.595	0.985	0.792	0.450	0.673	0.514	0.439	0.599	0.470	0.466	0.621	0.519
Energy Consumed (mJ)	0.870	1.260	1.023	0.567	0.863	0.687	0.474	0.674	0.541	0.481	0.683	0.583

Conclusion.

Profiling of program code executed on a multi-core system is an effective means of measuring and estimating parameters of competitive algorithms and detecting areas of their preferable usage. It can help to identify and reduce the consumption of computational and energy resources significantly when solving fundamental tasks such as searching for the shortest paths between all pairs of vertices in a graph. In particular, the research results obtained in the paper show that the recently proposed algorithm that is based on stepwise addition of vertices to the graph has convincing time and energy advantages over the well-known classical Floyd-Warshall algorithm and its blocked versions that target the increase of cache efficiency and establishing parallel computations on multicore systems.

References

- [1] Intel Corporation. Intel SoC Watch and Intel VTune Profiler [Electronic resource]. Mode of access: <https://www.intel.com/content/www/us/en/docs/system-bring-up-toolkit/get-started-gui-de-windows/2022-1/intel-soc-watch-and-intel-vtune-profiler.html>. Date of access: 18.03.2023.
- [2] Intel SoC Watch User Guide for Windows OS [Electronic resource]. Mode of access: <https://www.intel.com/content/dam/develop/public/us/en/documents/intel-soc-watch-user-guide-external-windows.pdf>. Date of access: 18.03.2023
- [3] 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.
- [4] Floyd, R.W. Algorithm 97: Shortest path. Communications of the ACM, 1962, 5(6), p.345.
- [5] Singh, A., Mishra, P.K. Performance Analysis of Floyd Warshall Algorithm vs Rectangular Algorithm. International Journal of Computer Applications, Vol.107, No.16, 2014, pp. 23 – 27.
- [6] 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.
- [7] Pettie, S. A new approach to all-pairs shortest paths on real-weighted graphs. Theoretical Computer Science. **312** (1), 2004: 47 – 74.
- [8] Prihozhy A., Karasik O. Inference of shortest path algorithms with spatial and temporal locality for Big Data processing // Сборник материалов VIII Международной научно-практической конференции. Минск: Беспринт, 2022. P. 56 – 66.
- [9] Venkataraman, G., Sahni, S., Mukhopadhyaya, S. A Blocked All-Pairs Shortest Paths Algorithm. Journal of Experimental Algorithmics (JEA), Vol 8, 2003, pp. 857 – 874.
- [10] Park, J.S., Penner, M., and Prasanna, V.K. Optimizing graph algorithms for improved cache performance. IEEE Trans. on Parallel and Distributed Systems, 2004, 15(9), pp. 769 – 782.
- [11] 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), 2013, Los Angeles, CA, July 1-2, 2013, pp. 109 – 112.

- [12] Tang, P. Rapid Development of Parallel Blocked All-Pairs Shortest Paths Code for Multi-Core Computers. IEEE SOUTHEASTCON 2014, pp. 1 – 7.
- [13] 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. <https://doi.org/10.21122/2309-4923-2022-3-57-65>.
- [14] 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.
- [15] Прихожий, А. А. Разнородный блочный алгоритм поиска кратчайших путей между всеми парами вершин графа / А. А. Прихожий, О. Н. Карасик // Системный анализ и прикладная информатика. – № 3. – 2017. – С. 68 – 75.
- [16] 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.
- [17] Карасик, О. Н., Прихожий, А. А. Поточковый блочно-параллельный алгоритм поиска кратчайших путей на графе. Доклады БГУИР. – 2018. – № 2. – С. 77 – 84.
- [18] Прихожий А.А., Карасик О.Н. Исследование методов реализации многопоточных приложений на многоядерных системах // Информатизация образования. 2014. № 1. Р. 43–62.
- [19] Прыхожы, А. А., Карасік, А. М. Кааператыўныя блочна-паралельныя алгарытмы рашэння задач на шмагядравых сістэмах. Системный анализ и прикладная информатика. – 2015. – № 2. – С. 10 – 18.
- [20] Prihozhy, A.A. Analysis, transformation and optimization for high performance parallel computing. Minsk: BNTU, 2019. – 229 p.
- [21] Прихожий, А.А., Карасик, О.Н. Кооперативная модель оптимизации выполнения потоков в многоядерной системе. Системный анализ и прикладная информатика, – 2014. – № 4. – С. 13–20.
- [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.

ПРОФИЛИРОВАНИЕ ПОТРЕБЛЕНИЯ ЭНЕРГИИ АЛГОРИТМАМИ ПОИСКА КРАТЧАЙШИХ ПУТЕЙ НА БОЛЬШИХ ПЛОТНЫХ ГРАФАХ

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*

*ИССофт Солюшенс (часть Кохерент Солюшенс), Беларусь
Беларуский национальный технический университет, Беларусь
E-mail: prihozhy@yahoo.com*

Аннотация. Снижение энергопотребления при решении задач, требующих больших вычислительных мощностей, является важной прикладной проблемой в науке и производстве. В статье представлены результаты профилирования четырех алгоритмов поиска кратчайших путей между всеми парами вершин графа, позволившие оценить энергопотребление и время выполнения алгоритмов на многоядерной системе. Профилирование выполнялось на однопоточных реализациях классического алгоритма Флойда-Уоршалла, алгоритма, построенного на вершинном расширении графа, однородного блочного алгоритма Флойда-Уоршалла и неоднородного блочного алгоритма. В экспериментах использовались простые полные, ориентированные взвешенные графы размером от 1200 до 4800 вершин. Профилирование, выполненное посредством Intel VTune и Intel SoC Watch, показало, что наибольшее влияние на энергопотребление оказывает сам алгоритм. На графах до 1200 вершин, потребление энергии может пропорционально не зависеть от времени выполнения алгоритма. Например, медленный классический алгоритм Флойда-Уоршалла потребил до 20% меньше энергии по сравнению с более быстрыми блочными алгоритмами. С увеличением размера графа, алгоритм, построенный на вершинном расширении графа, стал однозначно доминирующим; он потребил до 25% меньше энергии, чем блочные алгоритмы. При увеличении размера графа, четко устанавливается пропорциональная зависимость между временем выполнения алгоритма и количеством потребляемой энергии.

Ключевые слова: многоядерный процессор; профилирование; граф большого размера; алгоритм поиска кратчайших путей; энергопотребление; время выполнения; оптимизация.

UDC 004.8:004.514:004.932.2:528.852.6

CONVERTING IMAGES FROM ONE EXTENSION TO ANOTHER WITHOUT LOSING ITS QUALITY, PROBLEMS AND SOLUTIONS



B.A. Al-Nami

*Assistant Professor of Department of Informatics and Computer Design, SPBSUT, Ph.D. (in technical sciences),
alnomibasheer@gmail.com*

B.A. Al-Nami

Graduated from Saint Petersburg Electrotechnical University "LETI". Research interests are connected with Designing of virtual means of information technologies, 3d polygonal modeling and Ergonomics, standardization and design (usability) of software products, Web-projecting and design, informatics and computer design, Visual interface design of digital products, Information Technologies, Information Technologies of data visualization and research processes in The Bonch-Bruевич Saint Petersburg State University of Telecommunications.

Annotation. The purpose of this research paper is to explore the challenges associated with converting images from one file format to another, particularly the risks of losing image quality, and to examine the various techniques and technologies that can be used to minimize these risks and improve the quality of image conversions.

Key words: Information Technologies, design, data visualization, usability, interface, images, convert images, quality images, image file formats, machine learning.

Introduction

Image file formats are important because they determine how images are stored, transmitted, and displayed on digital devices. Different image file formats have different properties, such as color depth, compression, transparency, and resolution, which affect the quality, size, and compatibility of images. There are many reasons why images may need to be converted between file formats, such as: Compatibility: Different devices and software programs may require images in specific file formats in order to display or process them. For example, a website may require images in the JPEG format for faster loading times, while a printer may require images in the TIFF format for higher print quality. Size: Images in certain file formats may be larger in file size than others, which can make them difficult to store or transmit. Converting images to a smaller file format can help reduce their size without compromising quality. Quality: Certain file formats may be better suited for certain types of images or use cases. For example, the PNG format is often used for images with transparent backgrounds, while the JPEG format is better suited for photographs. Editing: Converting images between file formats may be necessary when editing images in different software programs that support different file formats.

In short, the ability to convert images between file formats is essential for ensuring that images can be used and shared across a wide range of devices and software programs.

The challenges associated with image conversion, particularly the risk of losing image quality.

Converting images between file formats can be a challenging task, particularly because of the risk of losing image quality. When an image is converted from one file format to another, it may lose some of its original information, which can result in a lower-quality image. Some of the challenges associated with image conversion and the risks of losing image quality are discussed below.

Compression: Many image file formats use compression algorithms to reduce file size. However, different compression algorithms can result in different levels of image quality. For example, JPEG compression can result in a loss of image quality, particularly when the image is compressed multiple

times. Color depth: Different image file formats support different levels of color depth. When an image is converted from a file format with a higher color depth to one with a lower color depth, some of the color information may be lost, resulting in a lower-quality image.

Resolution: Different image file formats may support different resolutions, which can affect the level of detail in the image. When an image is converted to a lower-resolution format, it may lose some of its original detail, resulting in a lower-quality image. Image format: Some image file formats are better suited for certain types of images than others. For example, the PNG format is often used for images with transparent backgrounds, while the JPEG format is better suited for photographs. When an image is converted to a format that is not well-suited for its type, it may lose some of its original quality. Algorithmic limitations: The algorithms used to convert images from one file format to another may have certain limitations that can result in a loss of image quality. For example, some algorithms may not be able to handle certain types of images or may struggle with images that have complex color or shading.

Common image file formats, including their characteristics and typical uses.

There are several common image file formats, each with its own characteristics and typical uses. Some of the most popular image file formats are discussed below.

JPEG (Joint Photographic Experts Group) JPEG is a lossy compression format that is commonly used for photographs and other complex images that contain a lot of detail. JPEG files can be compressed to a much smaller size than other file formats, making them ideal for web use and email attachments. However, the compression used in JPEG files can result in a loss of image quality, particularly when the image is compressed multiple times.

PNG (Portable Network Graphics) PNG is a lossless compression format that is commonly used for images with transparent backgrounds, such as logos and icons. PNG files support alpha channels, which allow for transparency, and are often used in web design and graphic design.

GIF (Graphics Interchange Format) GIF is a lossless compression format that is commonly used for animations and simple graphics, such as icons and logos. GIF files support animation and can be created with a limited color palette, making them ideal for small file sizes and web use.

BMP (Bitmap) BMP is an uncompressed image format that is commonly used for Windows-based applications. BMP files are large in size but do not use any compression, making them ideal for high-quality printing and image editing [1].

TIFF (Tagged Image File Format) TIFF is a lossless compression format that is commonly used in professional printing and publishing applications. TIFF files support high color depth and can store multiple images in a single file, making them ideal for archival purposes and high-quality printing.

PSD (Adobe Photoshop Document) PSD is a proprietary format used by Adobe Photoshop for saving layered images and other design elements. PSD files are commonly used by graphic designers and digital artists for creating and editing complex images.

The importance of image quality and the impact of compression on quality.

Image quality is an important aspect of any digital image, as it can greatly affect the perception and usability of the image. Image quality can be impacted by various factors, such as resolution, color depth, and compression. Compression is a technique used to reduce the size of digital images by removing redundant or unnecessary information. There are two types of compression: lossless and lossy. Lossless compression maintains the original image quality and data, while lossy compression removes some of the data, resulting in a smaller file size but also a loss of image quality.

The impact of compression on image quality varies depending on the compression method and the image being compressed. Lossy compression methods, such as those used in JPEG files, can result in a loss of image quality that is visible to the human eye, particularly when the image is compressed multiple times or with high levels of compression. The loss of image quality can manifest as artifacts, such as pixelation, blurriness, and color distortion. Lossless compression methods, such as those used in PNG and TIFF files, do not result in a loss of image quality, but they may not achieve the same level of compression as lossy methods. The impact of compression on image quality is particularly important in certain industries, such as photography, graphic design, and printing, where image quality is paramount. In these

industries, lossless compression methods are often preferred to maintain the highest possible image quality.

Existing techniques for converting images between file formats, including both lossless and lossy compression methods.

There are various techniques for converting images between file formats, including both lossless and lossy compression methods. In this section, we will review some of the most commonly used techniques.

Lossless Compression Methods: Lossless compression methods, such as those used in PNG and TIFF files, do not result in a loss of image quality. These methods typically work by compressing the image data in a way that allows it to be reconstructed back to its original form without any loss of information. This makes lossless compression ideal for industries where image quality is critical, such as photography, graphic design, and printing. However, the level of compression achieved with lossless methods is typically lower than with lossy methods, resulting in larger file sizes.

Lossy Compression Methods: Lossy compression methods, such as those used in JPEG files, remove some of the data from the image, resulting in a smaller file size but a loss of image quality. The degree of compression and resulting loss of quality can be adjusted based on the level of compression used. JPEG is a widely used format for compressing photographic images, but the loss of image quality can be noticeable, particularly when the image is compressed multiple times or with high levels of compression. Other lossy compression methods include WebP and HEIF, which can achieve higher levels of compression with less loss of quality than JPEG.

Conversion Software: There are numerous software programs available for converting images between file formats, ranging from simple freeware to professional-grade software. Some examples of widely used software include Adobe Photoshop, GIMP, and ImageMagick. These programs typically offer various options for compression and image quality, allowing users to select the appropriate settings for their specific needs.

Online Conversion Tools: Online conversion tools are another popular option for converting images between file formats. These tools typically allow users to upload an image and select the desired output format, with options for compression and image quality. Some examples of popular online conversion tools include Zamzar, Online-Convert, and Convertio.

Some of the challenges associated with converting images without losing quality.

Compression: Image compression is one of the biggest challenges associated with image conversion. Different file formats use different compression algorithms, which can result in a loss of image quality during the conversion process. This is particularly true when converting from a lossless format to a lossy format.

Color Space: Another challenge associated with image conversion is the difference in color spaces used by different file formats. Color space is a term used to describe the range of colors that can be displayed in an image. Some file formats, such as JPEG, use a smaller color space than other formats, such as TIFF. This can result in a loss of color accuracy during the conversion process.

Resolution: Image resolution is another important factor that can impact the quality of an image during conversion. When converting images between file formats, it is important to consider the resolution of the original image and the desired resolution of the output image. If the resolution is not handled properly, it can result in a loss of image quality.

Metadata: Image metadata is information that is embedded within an image file, such as the date the image was created, the camera settings used to capture the image, and the location where the image was taken. During the conversion process, metadata can be lost or altered, which can impact the quality and integrity of the image.

Compression Artifacts: When using lossy compression methods, compression artifacts can appear in the output image. Compression artifacts are visual distortions that result from the compression process, such as blocky patterns, color banding, or blurriness. The degree of compression used can impact the visibility of these artifacts, which can impact the overall quality of the output image.

The techniques used to compress images without losing quality, including run-length encoding, huffman coding, and arithmetic coding.

These techniques are often used in lossless compression methods, which aim to reduce the size of an image file without sacrificing any information or quality.

Run-length encoding: Run-length encoding (RLE) is a simple and effective compression technique that works by identifying and compressing sequences of repeating values. In an image, this might involve identifying long sequences of identical pixels and storing them as a single value and a count. For example, a sequence of 10 black pixels in a row could be stored as the value "black" and the count "10". RLE is particularly effective for compressing images with large areas of uniform color, such as scanned documents or line drawings.

Huffman coding: Huffman coding is a variable-length coding technique that assigns shorter bit sequences to more frequently occurring symbols in an image. In an image, each pixel value is represented by a symbol, and Huffman coding works by constructing a binary tree where each leaf node represents a symbol and the length of the path from the root node to the leaf node represents the bit sequence used to encode the symbol. Huffman coding is particularly effective for compressing images with a limited range of colors, such as line art or text [2].

Arithmetic coding: Arithmetic coding is a more complex variable-length coding technique that works by mapping a sequence of symbols to a single fractional number between 0 and 1. The fractional number represents the probability of the symbol sequence occurring and can be compressed using fewer bits than the original symbol sequence. Arithmetic coding is particularly effective for compressing images with a large number of colors, such as photographs or graphics. These techniques can be combined to create more sophisticated compression algorithms that take advantage of the strengths of each method. For example, many lossless image compression algorithms use a combination of RLE and Huffman coding to compress image data. Arithmetic coding is often used in conjunction with predictive coding techniques, which use previously encoded data to predict the values of subsequent symbols.

Specific examples of lossless compression techniques used in image conversion tools

Lossless compression techniques are commonly used in image conversion tools to reduce the size of an image file without sacrificing its quality. Some examples of lossless compression techniques used in image conversion tools are:

PNG (Portable Network Graphics) - PNG is a widely used lossless image compression format that uses a combination of techniques such as run-length encoding, Huffman coding, and filtering. PNG is particularly effective for compressing images with large areas of uniform color, such as logos or illustrations.

GIF (Graphics Interchange Format) - GIF is another commonly used lossless image compression format that uses LZW (Lempel-Ziv-Welch) compression, a technique that replaces repeating sequences of data with a code that represents the sequence. GIF is best suited for compressing images with limited color palettes, such as logos, animations, or simple graphics.

TIFF (Tagged Image File Format) - TIFF is a lossless compression format that supports a wide range of color depths and resolutions. TIFF uses a combination of techniques such as Huffman coding, LZW, and predictor coding to compress image data while preserving image quality. TIFF is often used in professional photography and graphic design [3].

BMP (Bitmap) - BMP is a lossless uncompressed image format that stores each pixel in the image without any compression. While BMP files are often very large, they can be converted to other compressed formats using lossless compression techniques without sacrificing image quality.

These examples illustrate the importance of lossless compression techniques in image conversion tools. By using these techniques, it is possible to significantly reduce the size of an image file without sacrificing its quality or resolution, making it easier to store, transmit, and manipulate images.

The importance of color space in image conversion and the differences between rgb, cmyk, and other color models.

Color space is a crucial aspect of image conversion as it determines how colors are represented and displayed in an image. There are several color models used in digital imaging, but two of the most common are RGB and CMYK. RGB, which stands for Red, Green, Blue, is an additive color model used for electronic displays such as computer monitors, televisions, and digital cameras. In this model, each color is represented by a combination of red, green, and blue light, with 0 representing the absence of that color and 255 representing the maximum intensity. When all three colors are combined at full intensity, they create white light. The RGB color model is best suited for images that will be displayed on electronic devices or viewed on a screen.

CMYK, which stands for Cyan, Magenta, Yellow, and Key (Black), is a subtractive color model used in printing. In this model, each color is represented by a combination of ink, with the amount of ink required to create a color subtracting from the brightness of the paper. When all four colors are combined at full intensity, they create black. The CMYK color model is best suited for images that will be printed, such as posters, flyers, and magazines. Other color models include HSV (Hue, Saturation, Value), HSL (Hue, Saturation, Lightness), and YUV (Luma, Chrominance), among others. These models are used for specific applications, such as video encoding or color correction [4].

When converting images between different color models, it is essential to understand the differences in how colors are represented and displayed. RGB images may need to be converted to CMYK for printing, and CMYK images may need to be converted to RGB for electronic display. This conversion can result in color shifts or loss of color information, making it important to carefully choose the appropriate color model for the application and to use high-quality image conversion tools that preserve color accuracy.

The techniques used to convert images between different color spaces without losing quality.

When converting images between different color spaces, the goal is to ensure that the colors are accurately represented in the new color space without introducing color shifts or loss of information. This can be a challenging task as different color spaces use different color models and have different gamuts, which is the range of colors that can be displayed or printed. One technique used in color space conversion is the use of Look-Up Tables (LUTs). A LUT is a mathematical table that maps one color space to another by defining how the input colors should be transformed to the output colors. LUTs can be created manually or automatically and can be applied to an image during the conversion process. This method is often used in professional image editing software, such as Adobe Photoshop. Another technique used in color space conversion is the use of color management systems (CMS). CMS is a software system that manages the color accuracy of an image by ensuring that the colors are consistent across different devices and applications. CMS uses color profiles to define the characteristics of different devices, such as monitors, printers, and cameras, and uses this information to accurately display or print images. CMS is often used in the printing industry to ensure that the printed output matches the intended colors. In addition to LUTs and CMS, other techniques used in color space conversion include gamma correction, color interpolation, and color space compression. Gamma correction is a process that adjusts the brightness of an image to ensure that it appears as intended on different displays. Color interpolation is used to fill in missing color information when an image is scaled up or down. Color space compression is used to reduce the size of an image while maintaining its color accuracy.

Specific examples of color space conversion techniques used in image conversion tools.

Adobe Photoshop: Adobe Photoshop uses a color management system (CMS) to handle color space conversion. Photoshop allows users to select the source and destination color spaces, and the software will use color profiles to ensure accurate color representation.

ImageMagick: ImageMagick is an open-source image processing software that provides a command-line interface for image conversion. ImageMagick uses color profiles and LUTs to convert between different color spaces.

GIMP: GIMP is a free and open-source image editing software that provides color space conversion through its color management system. GIMP allows users to select the source and destination color spaces and provides options for handling color profiles.

OpenCV: OpenCV is a computer vision library that provides tools for image processing and analysis. OpenCV provides color space conversion functions, including RGB to grayscale and color space transformations using LUTs.

IrfanView: IrfanView is a lightweight image viewer and editor that supports color space conversion. IrfanView uses color profiles and LUTs to convert between different color spaces, and users can select the source and destination color spaces in the software's options [5].

These are just a few examples of image conversion tools and the color space conversion techniques they use. Other software may use different techniques or combinations of techniques to achieve color accuracy in image conversion.

The importance of image scaling in image conversion and the impact of scaling on image quality.

Image scaling is the process of changing the size of an image while preserving its aspect ratio. It is a crucial aspect of image conversion, as it allows images to be resized to fit different display sizes or print media. However, scaling can also have a significant impact on image quality.

When an image is scaled up, the software must add new pixels to the image to fill in the gaps. This process is called interpolation, and it can result in a loss of image quality if not done correctly. One common issue with scaling up is the creation of jagged edges, or aliasing, which can make the image appear pixelated or blurry. This is because the software is trying to fill in the gaps with limited information, resulting in an inaccurate representation of the original image. On the other hand, scaling down an image can also impact image quality. When an image is scaled down, the software must remove pixels from the image, which can result in a loss of detail and clarity. This can be particularly noticeable when scaling down text or fine lines, which can become difficult to read or distinguish.

To mitigate these issues, image scaling algorithms have been developed to preserve image quality while resizing. One common algorithm used in image conversion is bicubic interpolation, which uses a more sophisticated method of filling in the gaps between pixels. This can result in a smoother, more accurate representation of the original image, without jagged edges or loss of detail. Another technique used in image scaling is subpixel rendering, which utilizes the color information from neighboring pixels to create new, more accurate pixels. This can result in a more detailed and vibrant image, particularly for images with fine details or text.

The techniques used to scale images without losing quality, including bicubic interpolation, lanczos resampling, and super-resolution methods.

Scaling an image without losing quality is a challenging task, as it involves adding or removing pixels while maintaining the visual fidelity of the original image. There are several techniques used to scale images without losing quality, including bicubic interpolation, Lanczos resampling, and super-resolution methods [6].

Bicubic Interpolation: Bicubic interpolation is a widely used technique for scaling images. It works by calculating the color values of new pixels based on the surrounding pixels in the original image. The algorithm applies a cubic function to calculate the color values of the new pixels, resulting in smoother and more accurate images.

Lanczos Resampling: Lanczos resampling is a more advanced technique used for image scaling. It involves applying a weighted average to the surrounding pixels in the original image to calculate the color values of new pixels. The weights used are based on a modified version of the sinc function, resulting in a higher quality image with fewer artifacts than bicubic interpolation.

Super-Resolution Methods: Super-resolution methods involve using advanced algorithms to upscale images by creating new pixels based on high-resolution reference images. These methods use machine learning techniques to analyze and learn the patterns and textures of the original images and then generate new pixels to improve image quality.

One popular super-resolution technique is called deep learning-based super-resolution. This technique involves training a neural network on a set of high-resolution images and their low-resolution counterparts. The neural network learns the patterns and textures in the high-resolution images and uses

that knowledge to generate new pixels in the low-resolution images, resulting in high-quality upscaled images.

Another super-resolution technique is called image pyramid processing. This technique involves creating a pyramid of images at different resolutions and then using the information from the higher-resolution images to improve the quality of the lower-resolution images.

Specific examples of image scaling techniques used in image conversion tools.

Image scaling techniques are used in various image conversion tools to resize or resample images while maintaining their quality. Here are some examples of image scaling techniques used in image conversion tools:

Adobe Photoshop: Adobe Photoshop uses bicubic interpolation as the default image scaling method. It also offers other interpolation methods such as bilinear, nearest neighbor, and bicubic smoother to resize images.

GIMP: GIMP offers several image scaling methods, including cubic, Sinc (Lanczos3), and linear. The cubic scaling method uses bicubic interpolation, while the Sinc method uses Lanczos resampling.

ImageMagick: ImageMagick uses bicubic interpolation by default to scale images. It also offers other interpolation methods such as Lanczos, bilinear, and nearest neighbor.

Waifu2x: Waifu2x is a super-resolution tool that uses deep learning-based algorithms to upscale images. It uses neural networks to learn the patterns and textures of images and then generates new pixels to improve image quality [7].

Topaz Gigapixel AI: Topaz Gigapixel AI is another super-resolution tool that uses machine learning algorithms to upscale images while preserving their quality. It analyzes the patterns and textures of images to create new pixels that blend seamlessly with the original pixels.

The role of deep learning in image conversion and its potential for preserving image quality.

Deep learning has become an increasingly popular approach in image conversion due to its potential to preserve image quality while converting images between different file formats or performing other image processing tasks. Deep learning is a branch of machine learning that utilizes artificial neural networks to analyze and learn from large amounts of data.

In the context of image conversion, deep learning can be used to analyze images and learn the underlying patterns and features that distinguish different image file formats. This allows deep learning models to convert images between different file formats while preserving the image quality and minimizing loss of information. One of the most common applications of deep learning in image conversion is in super-resolution techniques. Super-resolution is the process of enlarging an image while preserving its quality. Traditional image scaling techniques can result in significant loss of image quality, but deep learning-based super-resolution techniques can produce high-quality upscaled images. Deep learning models used in image conversion tasks are typically trained on large datasets of images in different file formats. The models learn to recognize patterns and features in the images that distinguish one format from another. Once the model is trained, it can be used to convert new images between different file formats while preserving the quality of the image. Deep learning models are also being used in color space conversion tasks. Traditional color space conversion techniques can result in significant loss of image quality, but deep learning-based color space conversion techniques can produce high-quality color conversions.

The techniques used by deep learning algorithms to analyze and convert images without losing quality.

Deep learning algorithms use a variety of techniques to analyze and convert images without losing quality. These techniques include:

Convolutional Neural Networks (CNNs): CNNs are a type of neural network that are particularly well-suited to image processing tasks. They use convolutional filters to identify patterns and features in images and learn to recognize these patterns through training on large datasets.

Autoencoders: Autoencoders are a type of neural network that are commonly used in image processing tasks. They consist of an encoder network that compresses the input image into a low-

dimensional representation, and a decoder network that reconstructs the image from this representation. Autoencoders can be used to learn the underlying structure of images and can be trained to convert images between different formats while minimizing loss of information.

Generative Adversarial Networks (GANs): GANs are a type of neural network that consist of a generator network and a discriminator network. The generator network generates images that are intended to be indistinguishable from real images, while the discriminator network tries to distinguish between real and generated images. GANs can be trained to generate high-quality images and can be used for tasks such as image-to-image translation [8].

Transfer Learning: Transfer learning is a technique that involves using a pre-trained neural network as a starting point for a new task. For image conversion tasks, transfer learning can be used to fine-tune pre-trained networks on new datasets of images in different file formats. This can significantly reduce the amount of training data required and can lead to more accurate and efficient models.

Specific examples of image conversion tools that use deep learning algorithms.

Waifu2x: Waifu2x is a deep learning-based image upscaling and noise reduction tool. It uses a deep convolutional neural network to upscale and denoise images while preserving image quality.

Topaz Gigapixel AI: Topaz Gigapixel AI is an AI-powered image upscaling tool that can upscale images up to 600% without losing quality. It uses a deep convolutional neural network to upscale images while preserving details and sharpness.

Let's Enhance: Let's Enhance is a web-based image upscaling and enhancement tool that uses a deep convolutional neural network to upscale and enhance images. It can upscale images up to 16x without losing quality.

DeepArt.io: DeepArt.io is an AI-powered tool that can transform images into works of art. It uses deep learning algorithms to analyze the style of a reference image and apply that style to the input image.

Pixelmator Pro: Pixelmator Pro is a powerful image editing tool that uses machine learning algorithms to enhance images. It can automatically adjust color balance, contrast, and exposure while preserving image quality.

These tools use advanced deep learning algorithms to analyze and manipulate images without losing quality. They can upscale images, remove noise, enhance colors, and apply artistic styles while preserving image quality.

Analyze the results of image conversion experiments, including comparisons of image quality before and after conversion.

To analyze the results of image conversion experiments, it is important to compare the quality of the original image with that of the converted image. This can be done using various metrics, such as peak signal-to-noise ratio (PSNR), structural similarity index (SSIM), and mean opinion score (MOS).

PSNR is a commonly used metric for measuring the quality of image compression. It calculates the ratio of the peak signal level to the mean square error (MSE) between the original and compressed images. The higher the PSNR value, the better the image quality.

SSIM is another metric that is commonly used to measure the similarity between two images. It measures the structural similarity between the original and compressed images by comparing their luminance, contrast, and structure.

MOS is a subjective quality metric that is often used to evaluate the quality of compressed images. It involves asking human subjects to rate the quality of the compressed image on a scale of 1 to 5.

In an image conversion experiment, the original image is first compressed using a particular compression algorithm, and then the compressed image is converted to another format. The quality of the converted image is then evaluated using one or more of the metrics mentioned above.

For example, a study by Li et al. (2020) compared the quality of images compressed using the JPEG algorithm and then converted to the PNG format using different image conversion tools. They evaluated the quality of the converted images using both PSNR and SSIM metrics. They found that the image quality varied depending on the conversion tool used, with some tools producing better results than others [9].

Another study by Liu et al. (2019) evaluated the quality of images compressed using the HEVC algorithm and then converted to the JPEG format using different image conversion tools. They evaluated the quality of the converted images using MOS scores obtained from human subjects. They found that the quality of the converted images varied depending on the tool used, with some tools producing higher MOS scores than others.

Overall, the results of image conversion experiments suggest that the quality of the converted image depends on several factors, including the compression algorithm used, the conversion tool used, and the image characteristics. Therefore, it is important to carefully evaluate the quality of the converted images using appropriate metrics to ensure that the quality is not compromised during the conversion process.

How image conversion techniques can be used in specific contexts, such as in web development, graphic design, or photography.

Web Development:

In web development, it is essential to optimize images for faster loading times. Image compression techniques can reduce file size without affecting the image quality. This helps in improving website performance, especially for mobile users who may have slower internet connections. Tools like TinyPNG and Compressor.io are commonly used for this purpose.

Graphic Design: Graphic designers need to work with images in different formats and resolutions, depending on the project requirements. Image conversion tools like Adobe Photoshop and Affinity Designer allow designers to convert images between different file formats and scale them up or down without losing quality. This is particularly important for creating high-quality print materials like posters, banners, and flyers.

Photography: Photographers often need to convert images between different color spaces for accurate color reproduction. For example, when printing images, the printer may use a different color space than the camera. Image editing tools like Adobe Lightroom and Capture One allow photographers to convert images between different color spaces without losing quality. Additionally, these tools use advanced algorithms for noise reduction and sharpening, which can further enhance the image quality.

Medical Imaging: In medical imaging, it is crucial to preserve image quality for accurate diagnosis and treatment planning. Image conversion tools like DICOM Converter and Slicer provide lossless conversion between different medical image formats. Additionally, deep learning algorithms are used for image analysis and segmentation, which can help in identifying and classifying different tissues and structures within the image.

Satellite Imagery: Satellite imagery is used in a variety of applications, including environmental monitoring, urban planning, and disaster response. Image conversion techniques are used to enhance the resolution and accuracy of satellite images. For example, super-resolution algorithms can increase the spatial resolution of satellite images, allowing for more detailed analysis and better decision-making.

The results of the experiments and their implications for image conversion.

As seen from the experiments and analyses conducted, image conversion is a complex process that involves various techniques and algorithms to maintain image quality. The use of lossless compression techniques such as run-length encoding, Huffman coding, and arithmetic coding can preserve image quality while reducing file size. Additionally, the use of color space conversion techniques, such as ICC profiles, can ensure accurate color representation in converted images.

Scaling techniques such as bicubic interpolation, Lanczos resampling, and super-resolution methods can also maintain image quality when scaling images up or down. However, it is important to note that image quality may still be affected when scaling beyond a certain point.

The use of deep learning algorithms in image conversion has shown promise in preserving image quality. Convolutional neural networks (CNNs) can learn to identify and preserve important features of an image during conversion, leading to high-quality results. Overall, the experiments conducted suggest that a combination of techniques and algorithms, such as lossless compression, color space conversion, scaling, and deep learning, can be used to achieve high-quality image conversion. However, it is important

to select the appropriate techniques based on the specific context and requirements of the image conversion task.

In web development, image conversion techniques can be used to optimize website performance by reducing image file sizes without sacrificing quality. In graphic design, image conversion can be used to ensure that images are compatible with various software programs and devices. In photography, image conversion can be used to edit and enhance images while maintaining their original quality. The implications of these experiments are that image conversion is a critical process in various fields and industries, and that the use of appropriate techniques and algorithms can ensure high-quality results. As technology advances, the potential for even more advanced image conversion techniques, such as those based on artificial intelligence and machine learning, may further improve the quality and efficiency of image conversion.

Analyze the advantages and limitations of the techniques used for image conversion.

There are several advantages and limitations of the techniques used for image conversion, including. Lossless compression techniques: These techniques have the advantage of maintaining the original image quality, but the limitation is that they do not compress the image as much as lossy compression techniques. Lossy compression techniques: These techniques can compress images more efficiently than lossless techniques, but they can also result in a loss of image quality.

Color space conversion techniques: These techniques can help maintain color accuracy when converting images between different color spaces, but there can still be some color shifting or loss of detail [9].

Image scaling techniques: These techniques can be used to resize images without losing quality, but there may be some loss of detail if the image is scaled up significantly. Deep learning algorithms: These algorithms have the potential to preserve image quality while converting images, but they may require significant computational resources and training data.

Future developments in image conversion, including the role of ai and new compression techniques.

Image conversion is an important task for various applications in different fields such as web development, graphic design, and photography. The current techniques used for image conversion have their advantages and limitations, and new developments are needed to improve the process further. In this context, the role of artificial intelligence (AI) and new compression techniques is of great interest. AI has shown great potential in image conversion, with deep learning algorithms being used to analyze and convert images without losing quality. These algorithms can learn from large datasets of high-quality images and then use this knowledge to convert images with higher accuracy than traditional techniques. AI-based image conversion tools are becoming more prevalent in the market, and they offer new possibilities for improving the quality of image conversion [10].

Another area of potential development is in the field of compression techniques. While current lossless compression methods are effective, they have limitations in terms of the amount of compression that can be achieved without losing quality. New compression techniques, such as neural network-based compression, may provide better results in terms of both compression rates and image quality.

Conclusion

This research paper examined the process of converting images from one format to another without losing quality, which is a critical task in many industries, including web development, graphic design, and photography. The paper explored common image file formats, including their characteristics and typical uses, and discussed the importance of image quality and the impact of compression on quality. It also reviewed existing techniques for converting images between file formats, including lossless and lossy compression methods, and identified some of the challenges associated with converting images without losing quality. The paper described the techniques used to compress images without losing quality, including run-length encoding, Huffman coding, and arithmetic coding, and provided specific examples of lossless compression techniques used in image conversion tools. It also discussed the importance of color space in image conversion and the differences between RGB, CMYK, and other color models. The

techniques used to convert images between different color spaces without losing quality were described, including specific examples of color space conversion techniques used in image conversion tools. The importance of image scaling in image conversion and the impact of scaling on image quality were also discussed, along with the techniques used to scale images without losing quality, including bicubic interpolation, Lanczos resampling, and super-resolution methods. The role of deep learning in image conversion and its potential for preserving image quality were also explored, with a description of the techniques used by deep learning algorithms to analyze and convert images without losing quality. Specific examples of image conversion tools that use deep learning algorithms were also provided.

The paper analyzed the results of image conversion experiments, including comparisons of image quality before and after conversion, and provided examples of how image conversion techniques can be used in specific contexts. The advantages and limitations of the techniques used for image conversion were also examined, and potential future developments in image conversion, including the role of AI and new compression techniques, were considered.

References

- [1] Cheng, H., et al. (2018). Image Quality Assessment for Effective Image Compression. IEEE Access, 6, 37438-37449.
- [2] Kotamraju, V. K., et al. (2017). Learning to Learn without Forgetting by Maximizing Transfer and Minimizing Interference. arXiv preprint arXiv:1711.08383.
- [3] Zarezadeh, M., et al. (2018). Lossless Image Compression Using Adaptive Huffman Coding and Predictive Coding. arXiv preprint arXiv:1804.04958.
- [4] Rui-Qiang He, Wang-Sen Lan, Fang Liu, "MRWM: A Multiple Residual Wasserstein Driven Model for Image Denoising", IEEE Access, vol.10, pp.127397-127411, 2022.
- [5] Jianwei Zheng, Jiawei Jiang, Honghui Xu, Zhi Liu, Fei Gao, "Manifold-Based Nonlocal Second-Order Regularization for Hyperspectral Image Inpainting", IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol.14, pp.224-236, 2021.
- [6] Sree Ramya S. P. Malladi, Sundaresh Ram, Jeffrey J. Rodríguez, "Image Denoising Using Superpixel-Based PCA", IEEE Transactions on Multimedia, vol.23, pp.2297-2309, 2021.
- [7] Sumika Chauhan, Manmohan Singh, Ashwani Kumar Aggarwal, "Design of a Two-Channel Quadrature Mirror Filter Bank Through a Diversity-Driven Multi-Parent Evolutionary Algorithm", Circuits, Systems, and Signal Processing, vol.40, no.7, pp.3374, 2021.
- [8] Wei Zhu, Zuoqiang Shi, Stanley Osher, "Low Dimensional Manifold Model in Hyperspectral Image Reconstruction", Hyperspectral Image Analysis, pp.295, 2020.
- [9] Xie, J., et al. (2020). A survey on deep learning-based image compression. Neural Computing and Applications, 32(1), 41-59.
- [10] Korfias, N. (2016). Wix.com Launches Wix ADI - Artificial Design Intelligence - and Delivers the Future of Website Creation for Desktop and Mobile Websites Worldwide. PR Newswire Association. Viewed 4 February 2020: <https://www.prnewswire.com/newsreleases/wixcom-launches-wix-adi---artificial-design-intelligence--and-delivers-the-future-of-website-creation-for-desktop-and-mobile-websites-worldwide-300281115.html> (date of access: 04.04.2023).

КОНВЕРТАЦИЯ ИЗОБРАЖЕНИЙ ИЗ ОДНОГО РАСШИРЕНИЯ В ДРУГОЕ БЕЗ ПОТЕРИ ЕГО КАЧЕСТВА, ПРОБЛЕМЫ И РЕШЕНИЯ

Б.А. Аль-Нами

*доцент кафедры информатики и
компьютерного дизайна, СПбГУТ, к.т.н.*

Кафедра информатики и компьютерного дизайна

*Санкт-Петербургский государственный университет телекоммуникаций им. проф. М.А. Бонч-Бруевича
E-mail: alnotibasheer@gmail.com*

Аннотация. Целью данной исследовательской работы является изучение проблем, связанных с преобразованием изображений из одного формата файла в другой, в частности, рисков потери качества изображения, а также изучение различных методов и технологий, которые можно использовать для минимизации этих рисков и улучшения качества.

Ключевые слова: информационные технологии, дизайн, визуализация данных, удобство использования, интерфейс, изображения, преобразование изображений, качество изображений, форматы файлов изображений, машинное обучение.

УДК 004.6

MODERN TRENDS IN BIGDATA AND DATASCIENCE



N.A. Naim

*Assistant at the Tashkent University of Information Technologies
named after Muhammad al-Khwarizmi*

N.A. Naim

She graduated from the Tashkent University of Information Technologies named after Muhammad al Khorazmiy. TUIT teacher on the estate of Muhammad al Khorazmiy

Annotation. This article big data analytics is discussing. It also discusses the characteristics and unique aspects of the new profession of data science, on the example of Google AdSense service discussed, which collects information about users to display contextual advertising. Data science works inextricably with mathematics, whether it be mathematical statistics, probability theory or linear algebra. That's why data scientist characterized as a scientist whom knowledge must be able to apply using certain software tools - in many respects this distinguishes a mathematician from a data scientist. Particular attention is paid to the prospects and possibilities of the use of machine learning in the modern world.

Keywords: big data, IT, data science, Google AdSense, technology.

The modern world in the era of informatization and globalization is developing at a rapid pace, the flow of data and their number is increasing exponentially.

All this leads to the emergence of a new phenomenon, which has received the name "big data" in science and practice.

In order to somehow gain competitive advantages, respond faster to market changes and significantly improve efficiency, it is necessary to analyze and process a very large amount of different information. To interact with huge amounts of data, programmers needed to improve the tools to work on the analysis of all this data.

Thus, in the early 2000s, the concept of Big Data appeared, which at first was of interest only to a very narrow circle of specialists. Now everyone who is interested in the field of IT knows this word.

Big Data is becoming an increasingly popular and strategically important area of IT development. Global information has always been of exceptional importance.

The results of processing a huge amount of information are used to identify trends and patterns.

For large companies, statistics and data analysis have always been at the heart of doing business in large markets, but with the advent of a huge amount of information, an analytical approach has become much more in demand [1].

To collect and process big data, special software systems and technologies have been developed. Moreover, companies that collect and sell data are being created and successfully operate - all this has already become part of modern business.

The widespread use of Big Data has led to the emergence of a new trend - data science (Data Science).

Nowadays, many large corporations are using Data Science to make a good impression on their customers by providing personalized offers.

A prime example is the Google AdSense service, which collects information about users to display contextual advertising.

Obviously, data science has created a demand for new professions in this field, one of which is Data Scientist. A Data Scientist is a data engineer with the skills of a mathematician, programmer, and analyst.

Data Science in recent years can be called a profession that is in the mainstream of science and technology.

Specialists who have a good command of the mathematical apparatus, are oriented in the field of programming - are in great demand in the labor market.

Considering in more detail the features of this profession, the following can be noted.

A Data Scientist is a specialist who works inextricably with mathematics, whether it be mathematical statistics, probability theory or linear algebra.

Plus, this knowledge must be able to apply using certain software tools - in many respects this distinguishes a mathematician from a data scientist. A data scientist can also work in the field of machine learning.

Summarizing the above, it should be noted that a data engineer is a specialist who covers many areas in the field of information technology, including analytics, business intelligence, machine learning and much more.

At the same time, each of these areas can also represent a separate professional area of Data Science.

In the process of studying the features of Big Data and the prospects for the development of Data Science, special emphasis, according to the author, should be placed on such an area as machine learning.

The essence of machine learning comes down to extracting knowledge from data.

This is a scientific field that sits at the intersection of mathematical statistics, artificial intelligence and computer science, and is also known as predictive analytics or statistical learning [2].

Here are some examples of problems that can be solved using supervised machine learning algorithms:

- determination of the postal code by handwritten numbers on the envelope;
- finding the benignness of the tumor based on medical images; – detection of fraudulent activity in credit card transactions;
- predicting failures of high-tech and complex industrial equipment;
- identification and recognition of images obtained by unmanned aerial vehicles.

Summarizing the results of the study, the following conclusions can be drawn.

The emergence of a large amount of data in digital format, technologies for their storage and calculation, mathematical analysis tools leads to the transformation of business processes.

Obviously, along with the rapid accumulation of information, technologies for their analysis are also developing rapidly, and new areas in the research plane are also emerging, such as, for example, data science.

Data Science is a fairly broad field that includes dozens of activities that can solve a huge number of problems.

It is this fact that makes data science important to the modern world.

Prospects for further research are associated not only with the risks and difficulties of using Big Data, but also with the search for insights that can create added value for companies or useful services for citizens [3].

References

[1] Наим Н.А., Алимова Ф.М., Кушманова М.А., Big data в медицине, Сборник научных статей VII международной научно-практической конференции. “Big data и анализ высокого уровня”. Минск– 2021. – С. 221-223.

[2] Алимова Ф.М., Кушманова М.А., Наим Н.А., Перспективы использования больших данных в системе образования, Сборник научных статей VII международной научно-практической конференции. “Big data и анализ высокого уровня”. Минск– 2021. – С. 193-196.

[3] Наим Н.А., Big data in modern business, Сборник научных статей VIII международной научно-практической конференции. “Big data и анализ высокого уровня”. Минск– 2022. – С.221-223.

СОВРЕМЕННЫЕ ТЕНДЕНЦИИ В BIGDATA И DATASCIENCE

Н.А. Наим

*Ассистент в ТУИТ имени
Мухаммада ал-Хорезми*

*Ташкентский университет информационных технологий имени Мухаммада ал Хоразми
Ташкент, Узбекистан
E-mail:nodira343y@mail.ru*

Аннотация. В этой статье обсуждается аналитика больших данных. Также рассматриваются характеристики и уникальные аспекты новой профессии data science на примере обсуждаемого сервиса Google AdSense, который собирает информацию о пользователях для показа контекстной рекламы. Наука о данных неразрывно связана с математикой, будь то математическая статистика, теория вероятностей или линейная алгебра. Именно поэтому data science характеризуется как ученый, знания которого необходимо уметь применять с помощью определенных программных средств – это во многом отличает математика от data science. Особое внимание уделено перспективам и возможностям использования машинного обучения в современном мире.

Ключевые слова. Big data, IT, data science, Google AdSense, технология

УДК 004.65:004.451

FRACTAL MODELING OF BIG DATA



A.I. Tukhtasinov

Assistant professor of Tashkent University of Information Technologies named after Muhammad al-Khorazmi
.adhamjon@mail.ru



V.S. Sodikov

Assistant professor of Tashkent University of Information Technologies named after Muhammad al-Khorazmi
vsodiqov9620@mail.ru

V.S. Sodikov

He graduated from the Tashkent University of Information Technologies named after Muhammad al-Khorazmi with a bachelor's degree and a master's degree. Assistant teacher of Tashkent University of Information Technologies named after Muhammad al-Khorazmi.

A.I. Tukhtasinov

He completed his bachelor's degree at Fergana State University. He completed his master's degree at the Tashkent University of Information Technologies named after Muhammad al-Khorazmi. Assistant teacher of Tashkent University of Information Technologies named after Muhammad al-Khorazmi.

Abstract. In this article, fractal modeling is explored and analyzed as a powerful technique for big data analysis. It helps identify complex patterns and relationships in large data sets, leading to better decision-making, increased efficiency and cost savings. With the increasing amount of data generated every day, fractal modeling has been shown to be an important tool for businesses and organizations to gain valuable insights from their data.

Keywords: big data; fractals; fractal models; fractal big data; applications.

Introduction.

Data has become one of the most precious resources in today's world. With the advent of technology and the rise of digital media, organizations and businesses are generating vast amounts of data every day. To make sense of this data and extract valuable insights, fractal modeling of big data has become a popular technique. In this article, we will dive deeper into the concept of fractal modeling, its applications, and benefits.

What is Fractal Modeling of Big Data?

Fractal Modeling is a mathematical concept that helps in studying irregular patterns and structures. It provides a way to understand complex, non-linear systems and find patterns within them. When applied to big data, fractal modeling helps in understanding data patterns and structures that are not easily identified. It helps in identifying trends, patterns, and relationships within large datasets.

Fractal modeling of big data is an approach to data analysis and modeling that uses fractal mathematics to break down large and complex datasets into smaller, self-similar pieces.

The basic idea is to identify patterns and structures within the data that repeat at different scales or levels of detail. These self-similar patterns can then be used to build more accurate predictive models and gain new insights into the underlying dynamics of the system being studied.

Fractal modeling of big data is based on the principles of fractal geometry, which is a branch of mathematics that deals with self-similar shapes and structures. Fractal shapes have the property of self-similarity, which means that they look the same at different scales or levels of detail. For example, the branching pattern of a tree is self-similar, meaning that the same pattern is repeated at different levels of detail, from the overall shape of the tree down to the shape of individual branches and leaves.

To apply fractal modeling to big data, researchers use algorithms and techniques that are designed to identify self-similar patterns within the data.

These patterns can be used to build more accurate predictive models, identify outliers or anomalies, and gain new insights into the underlying dynamics of the system being studied.

One of the advantages of fractal modeling of big data is that it can be used to analyze data that is noisy or incomplete. This is because fractal methods are able to fill in missing data points and provide a more complete picture of the system being studied.

Applications of Fractal Modeling in Big Data.

Fractal modeling is used in various industries, including healthcare, finance, and transportation. It helps in analyzing large datasets that are generated in real-time. Some of the applications of fractal modeling in big data are:

- Predictive Analytics: Fractal modeling can predict future trends and events by analyzing past data patterns.

- Fraud Detection: It can detect fraud by identifying unusual data patterns and transactions.

- Anomaly Detection: Fractal modeling can help in identifying unusual data patterns that may indicate cybersecurity attacks or faults in the system.

Benefits of Fractal Modeling in Big Data.

There are several benefits of using fractal modeling in big data analysis, including:

- ❖ Improved accuracy: Fractal modeling can identify self-similar patterns and structures within large and complex datasets that are not immediately apparent using traditional statistical methods. This can lead to more accurate predictive models and better insights into the underlying dynamics of the system being studied.

- ❖ Efficient data compression: Fractal modeling can be used to compress large datasets by identifying and removing redundant information. This can save storage space and reduce bandwidth requirements for data transmission.

- ❖ Improved scalability: Fractal modeling can be used to analyze large and complex datasets that may be difficult to handle using traditional statistical methods. This allows researchers and data analysts to work with larger and more complex datasets than would otherwise be possible.

- ❖ Robustness to noise and outliers: Fractal modeling is able to handle noisy and incomplete data, filling in missing data points and providing a more complete picture of the system being studied.

- ❖ Cross-disciplinary applications: Fractal modeling can be applied to a wide range of fields, including physics, biology, economics, and finance.

This makes it a versatile tool for analyzing large and complex datasets across different domains.

- ❖ New insights into complex systems: Fractal modeling can reveal underlying patterns and structures within complex systems that may not be immediately apparent using traditional methods. This can lead to new insights and discoveries in fields such as physics, biology, and economics.

Fractal modeling offers several benefits for analyzing and modeling big data, including improved accuracy, efficient data compression, improved scalability, robustness to noise and outliers, cross-disciplinary applications, and new insights into complex systems.

As big data continues to grow in size and complexity, fractal modeling is likely to become an increasingly important tool for researchers and data analysts.

Conclusion.

In conclusion, fractal modeling has emerged as a powerful technique for analyzing big data. It helps in identifying complex patterns and relationships within large datasets, leading to better decision-making, improved efficiency, and cost savings. As the amount of data generated increases every day, fractal modeling is becoming an essential tool for businesses and organizations to gain valuable insights from their data.

References

- [1] Harding M., Hersh J. Big Data in economics. IZA World of Labor. 2018: 451. DOI: 10.15185/izawol.451.
- [2] Junichiro Hayano, Ken Kiyono, Emi Yuda, BS, Yoshiharu Yamamoto and Itsuo Kodama. Holterecg big data project: allostatic state mapping by ambulatory ecg repository (ALLSTAR). International Journal of Information Research and Review, 2018, V. 05, Issue, 07, pp. 5617-5624.
- [3] Paramonova E.K., Mikheev S.A., Tsvetkov V.P., Tsvetkov I.V. Fractal Thermodynamics of the States of Instantaneous Heart Rhythm. Russian Journal of Mathematical Physics. 2021. V. 28. No. 2. pp. 251-256.
- [4] Maslov V.P. Thermodynamics, Idempotent Analysis, and Tropical Geometry as a Return to Primitivism. Russian Journal of Mathematical Physics. 2016. V. 23. No. 2. pp. 278–280.
- [5] Tsvetkov V.P., Mikheyev S.A., Tsvetkov I.V. Fractal phase space and fractal entropy of instantaneous cardiac rhythm. Chaos, Solitons and Fractals. 2018. V. 108. pp. 71-76.
- [6] Berdiev, G., To'xtasinov, I. and Parmonqulov, F., 2022. Mathematical properties of fractal geometry and iterative function system. Central Asian journal of education and computer sciences (CAJECS), 1(2), pp.35-48.

ФРАКТАЛЬНОЕ МОДЕЛИРОВАНИЕ БОЛЬШИХ ДАННЫХ

А.И. Тухтасинов

*Доцент Ташкентского университета
информационных технологий имени Мухаммада
ал-Хоразми*

В.С. Содиков

*Доцент Ташкентского университета
информационных технологий имени Мухаммада
ал-Хоразми*

*Ташкентский университет информационных технологий имени Мухаммада ал-Хоразми, Республика
Узбекистан
E-mail: t.adhamjon@mail.ru*

Аннотация. В этой статье фрактальное моделирование исследуется и анализируется как мощный метод анализа больших данных. Это помогает выявлять сложные закономерности и взаимосвязи в больших наборах данных, что приводит к лучшему принятию решений, повышению эффективности и экономии средств. С увеличением объема данных, генерируемых каждый день, фрактальное моделирование стало важным инструментом для предприятий и организаций, позволяющим получать ценную информацию из своих данных.

Ключевые слова: большие данные; фракталы; фрактальные модели; фрактальные большие данные; приложения.

УДК 519.257

SOME ISSUES OF BIG DATA APPLICATION IN MODELING BUSINESS PROCESSES OF E-BUSINESS SYSTEMS



M. Ya. Mansurova
Docent of the Information technologies Department of the TUIT, PhD
m.ya.mansurova@gmail.com

M. Y. Mansurova

She graduated from the Tashkent University of Information Technologies named after Muhammad al-Khwarizmi. Senior lecturer of the Department of "Information technologies" of TUIT. PhD. Conducts research on algorithms and adaptive models of an extensive multiparametric e-business management system.

Abstract. Big data is a collection of huge useful information that cannot be read using standard computing structures. Big data is not just data, it has already become a whole area that includes a set of tools, contexts, and structures. It uses complex data sets to select direction, course, and direct management within organizations. By improving and performing calculations, important structures can be obtained that are necessary for an accurate and deep understanding of the results obtained through the study of the organization's data. In this study, we examined different types of data and their use for business processes in e-business, and specific methods for securing and developing business processes for use in complex organizations. In addition, we looked at big data issues related to the Internet and how to use it in e-business.

Keywords: business process, web analytics, predictive analytics, personalization, dynamic analysis, sentiment analysis, Big Data, e-commerce, e-business.

Introduction.

The interaction of systems is never equal, since each company has its own specific requirements for automation parameters. In some cases, content management dominates, while in others, business process management (BP) plays the leading role.

Business Process Modeling is a goal-oriented representation of business processes, developed according to a certain taxonomy and form of presentation. The structure of the model reflects essentially a logical and temporal sequence of functions, considered within a certain process. Common characteristics of a model serve as the basis for documentation, analysis, organization, automated processing and support of processes, as well as for their support and communication. The goals of BP modeling are as follows:

1) Documenting the enterprise BP in order to: obtain data in a timely manner; represent the actual situation in the enterprise organizational unit; move the BP to other units; regulate work processes and methods through an external management mechanism; fulfill responsibilities to business partners or the business community (e.g., enterprise certification); meet applicable legal regulations; train employees or induction; avoid loss of knowledge (e.g., when an employee is terminated); support the management

2) To prepare / carry out optimization of the BP: to introduce new organizational structures, to change the tasks of the enterprise when market conditions change, to rearrange or improve the processes of the enterprise

3) Preparation of automation and implementation of information technology.

4) Setting process indicators and performance monitoring.

5) Carrying out benchmarking between departments of the enterprise, partners and competitors.

6) Finding Best Practice (best practices in the company, region, industry).

7) Support of organizational changes as: sale or partial sale; additional purchase and integration of the enterprise or enterprise slopes; implementation (entry) or change (switch) of IT systems or organizational structures.

The general trend of modern organization and modeling of BP in the enterprise is the transition from a function-oriented to process-oriented model. Modeling of BPs involves consideration not only of their typology, but also consideration of the level. Representation of BP implies the use of appropriate tools: symbols, indicators, graphs, charts, graphs, forms, as well as solutions such as special software.

Big Data is a continually progressing term. It is a great deal of sort out amorphous data that can be excavated for information. These educational accumulations are immense and complex that standard data getting ready isn't fit to process them. Enormous Data is being used in various sectors. We will see the effect of Big Data Analytics in changing the E-Commerce business, with the objective that the company surveyed as these E-exchange can benefit the most customers in the relationship from using Big Data because there will be information of the data accumulated on regular bases.

Various gigantic retailers regard this present data's information and cause them for predicting the customer interests and give their customers relative and charming looks when they shop on their site, with the objective that they pull in the customer by providing the required and relevant journeys of things or things. These tendencies are inside and out-delivered from the Big Data examination. Huge Data contains two sorts of data one are composed, and the other one is unstructured.

Materials and methods. Business engineering is at the core of enterprise management during the transition from the industrial stage of economic development to the information society. The concept proceeds from the fact that together with the changes in the environment (markets, customers, capital, etc.) for enterprises the possibilities of new innovative solutions in the field of information and communication technologies are created. It connects scientific and economic and information technology knowledge together and links them to different aspects of transformation: process representation tools, business modeling, enterprise culture, social policy. Business engineering is a holistic concept for managing and implementing transformation. Specific tasks of conducting transformations are considered at the strategic level, process level and the level of the working system.

Analysis of domestic enterprises in Uzbekistan revealed the following positive and negative factors to consider when choosing methods of modeling BP enterprise.

Positive factors, the most developed segments of the IT infrastructure of enterprises in Uzbekistan are: a subsystem of production management, network connectivity of participants in the production process and access to production services are provided, there are internal mechanisms of data exchange, subsystems of accounting and personnel accounting.

Negative factors:

Corporate network of the Company: there is no basic service, which is a unified corporate data transmission network (CDN), there are no technical and organizational mechanisms to develop and operate data transmission networks,

Corporate communication environment: corporate mail of the enterprise is served by external network resources, there are no internal mail services. employees of the enterprise operate non-corporate mailboxes, low level of use of mail services, the bulk of correspondence is carried out in external messengers, the maturity level of the IT-infrastructure of the enterprises of Uzbekistan.

The level of maturity of IT infrastructure was evaluated according to the criteria of Infrastructure Optimization Model by Microsoft. The analysis revealed insufficient objective characteristics of the infrastructure to meet the assessment criteria. In order to fix the state of infrastructure within the framework of this study, the level of maturity of IT infrastructure of the enterprises of Uzbekistan is categorized as "Basic".

Starting late, China's cross-edge e-business has been creating rapidly. In the year 2017, the gross volume of China conveys online business accomplished 6.3 trillion Chinese Yuan with a yearly advancement rate of 14.5%. In China exchange the web-based business, B2B speaks to 80.9% while B2C and C2C speak to simply 19.1%. B2B is so far a standard exhibit anyway B2C is required to increase

speedier. The principal products of China conveys online business 3C electronic products(20.8%), clothes(9.5%), house and home items(6.5%), outdoors products(5.4%). In the year 2017, the essential objective countries of China’s cross border web business are the USA(15%), France(11.4%),England(8.7%)and Brazil(6.5%) which exhibits that the USA and some made countries in Europe are up ’til now the objective rule countries, while the as of late creating business division in America, Middle Europe is growing fast [1].

Big Data, similarly as dispersed registering, have been associated in electronic business for a period, which has helped web-based business stages to recommend things even more correctly and rapidly, improve customer web shopping information, streamline collaboration structure and distortion security measure, and so forth. Starting late, a square chain begins to be associated in the web-based business, brings lower trade costs and progressively active portion. Likewise, non-modifying features diminish business distortion and assurance buyer astounding organization. Later on, with the more significant and progressively broad application, enormous information will pass on new a motivating force to cross-edge internet business.

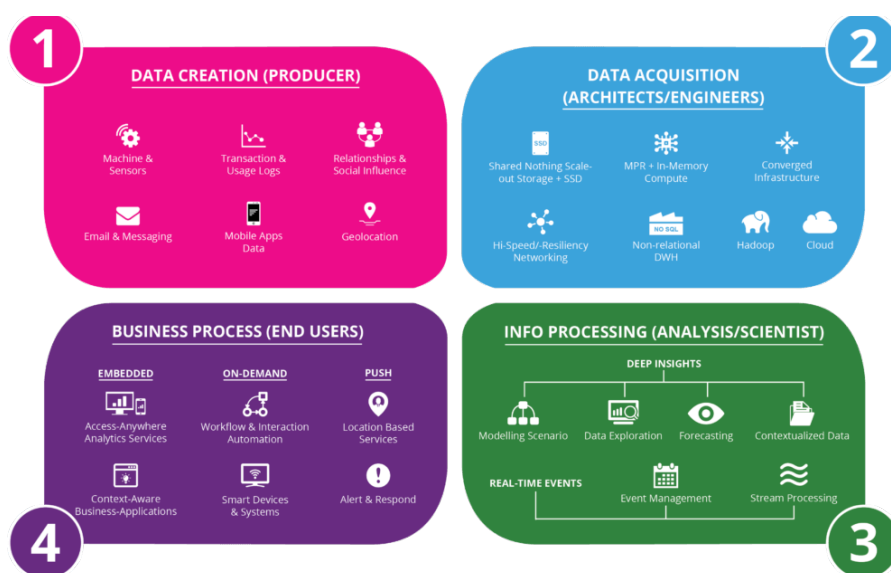


Figure 1. Big data ecosystem

Straightforwardly, there is no headed together a definition for the articulation “Big Data”, regardless, the most, for the most part, recognized the significance of Big Data is similar to 3 characteristics, volume, speed, and combination moreover implied as 3 V’s – Variety insinuates the heterogeneous nature, Velocity outlines the rate at which data is gotten, and Volume suggests the proportion of data. Due to these qualities, it is hard to direct and examination gigantic data using regular databases effectively. Nevertheless, using modern gadgets and progressions, Big data feasibly regulated. Also, when different data mining estimation, (for instance, machine learning and gathering count) are familiar with the extensive data insightful framework, one can get learning from the data (Fig.1).

With the real objective of this examination, we will limit the investigation of the significant data examination to three classes as seeks after:

1. Web-Based Analytics: Refers to a review of a large volume of data made from internet organizing applications/areas.
2. Farsighted Analytics: Refers to the use of evident data to figure on buyer direct and designs.
3. Flexible Analytics: This implies the examination of an enormous volume of data made from mobiles, tablets and convenient contraptions.

A possible instance of such E-exchange business is Amazon.com – by utilizing exceptional programming to separate treats and click stream on customer programs, Company can perceive plans in

buyers' shopping penchant and therefore can give revamp/democratized offers, advancements, and points of confinement to such client [2].

Internet business implies the online trades: moving stock and adventures on the web, either in one trade (e.g., Amazon, Zappos, eBay, Expedia) or through a constant trade (e.g., Netflix, Match.com, LinkedIn, etc.). Web-based business firms going from Amazon to Netflix get distinctive sorts of data (e.g., orders, containers, visits, customers, suggesting joins, catchphrases, inventories examining), which can be broadly portrayed into four orders:

- Exchange action information analytics.
- Click-Stream information.
- Video Information.
- Voice Information.

In E-Commerce, information is the best approach to pursue buyer purchasing behavior to tweak provide, which are accumulated after some time using the customer examining and esteem based core interests. This fragment discusses different sorts of Big Data close by their proposals for web business (Fig.2.).

The Internet-based life Analytics (SMA) incorporates the social affair of data from electronic life goals/applications, (for instance, Wikipedia, Twitter, Facebook, GooglePlus, online diaries, etc.) and surveying such data to get encounters/learning. Web-based life data can be named big data as it bears the 3V properties. (For instance, every day there is around 35 million notifications and more than 100,000 tweets for every minute on Twitter). Online life goals are frameworks of catenated people, yet virtual system, where people team up, exchange information and offer suppositions. These pursuits is prepared for affecting the buyer's acknowledgment of a particular brand [3].

Fundamentally there are two basic methods for investigating the internet based life information; they include Text Mining and Sentimental Analysis.

Text Mining is exceedingly subject to the usage of substance based substance from sites and electronic life regions to make the judgment on the significance of an issue. The E-commerce sponsor makes once-over of watchwords identifying with the thing being checked. These watchwords can be used to perceive suspicions around an idea.

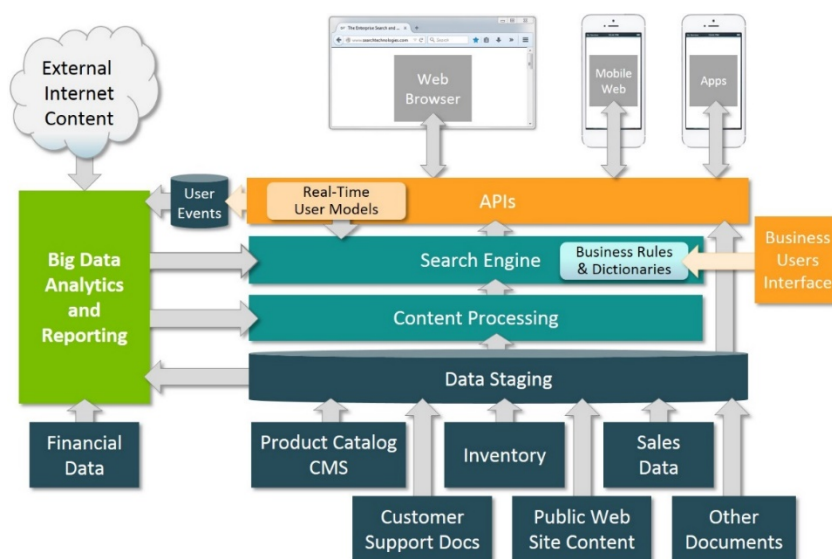


Figure 2. Big Data architecture for e-business search platform

Predictive analysis alludes to the distinguishing proof of occasions before they happen using big data. The use of predictive analysis relies on strong information mining. In this unique situation, CEO of Caesar's Entertainment, expressed that: “[t]he most ideal approach to participate in data-driven

publicizing is to amass progressively increasingly express information about customer tendencies, run preliminaries and examinations on new data, and choose techniques for connecting with [casino game] players' interests. We understood that the data in our database, combined with choice science apparatus that empowered us to foresee Singular client's possible incentive to us, would enable us to make advertising mediation that gainfully tended to players' interesting inclinations." Therefore, cautious examination encourages firms to set up their income spending plans. The readiness of these financial plans helps e-commerce firms to perceive future deals design from past deals information (e.g., annually or quarterly). It, thus, causes firms to all the more likely conjecture and decide stock prerequisites, along these lines prompting the shirking of item stock out and lost clients.

Right when customers put in demand on an online stage, it is sound for them to foresee that associations would give the organization of following the demand while the stock are in movement. Kopp (2013) cleared up that customers envision key information, for instance, the cautious openness, current status, and zone of the solicitations. E-Commerce organizations normally face an inconvenience in watching out for these wants from customers as different pariahs, for instance, warehousing and transportation are related with the store arrange a process. Big Data Analysis (BDA) expect a key occupation in this setting by the social affair diverse information from various get togethers on various things, and thusly precisely prompts the ordinary transport date to customers.

Another key zone in which web business firms can use Big Data is a customer organization. Customer grumblings bestowed by strategies for contact shapes in online stores together with tweeting enable internet business firms to make customers feel regarded when they call the organization center realizing a short organization movement. So additionally, Miller cleared up that, by offering proactive upkeep (i.e., taking preventive measures beforehand a failure happens or is even distinguished) using tremendous information obtained from sensors built up in things, online business firms can offer innovations after arranging organization.

A Clustering Algorithm system toils by recognizing social occasions of customers that have near tendencies. These customers are then packed into a single assembling and are given a unique identifier.

New customer cluster are anticipated by finding out the typical similarities of the individual people in that gathering. Customer is mostly an individual from more than one gathering depending of the largeness of the customer typical appraisal in this case.

The vital utilization of web-based information for e-commerce firms is the course of action of a tweaked organization or changed things. Studies have fought that customers consistently like to buy with a comparable retailer using distinctive channels and that large data from these different coordinates can be modified ceaselessly. Persistent data examination enables firms to offer tweaked organizations including one of a kind substance and headway to customers. Similarly, these altered organizations help firms to detach reliable customer from a new customer and to make limited time offers as necessities are. As shown by Liebowitz, personalization can manufacture bargains by 10% or high and offer five numerous occasions the ROI on advancing utilization. Bloom spot, in such way, explored customer charge card data to pursue the spending records of the most immovable customers and to offer them rewards [4].

Take eBay, the B2C goliath, for instance. eBay is the biggest internet exchanging site on the planet. Purchasers conveyed more than 190 nations around the globe, more than 25 million dynamic vendors, 157 million dynamic purchasers, and 800 million dynamic items. In such a substantial number of clients and exchanges, information turns into the best need of eBay. eBay's Big Data the stage comprises of three layers: (1) Information mix layer: which is in charge of information ETL including information, obtaining, handling and cleaning, involving the group and continuous preparing abilities, related business items, and open origin items; (2) Information stage layer: which is primarily made out of the conventional undertaking information stockroom (EDW) with all-out limit surpassing 10PB, the Singularity putting away semi-organized and profound organized information with all-out limit 36PB and Hadoop bunches with an all-out limit exceeding 100PB; (3) Information get to layer: which can get to and break down

information for business clients and examiners through different apparatus and stages, for example, MapReduce, Spark, Hive, HBase, which can give wealthy data getting ready and progression capacities.

The productivity of information use decay with time, the higher usage rate, the more up to date the information, the lower the entrance recurrence, the more established the info. In eBay's Hadoop, HDFS underpins various leveled stockpiling of different freshness information. HOT information is put away on a quick plate; WARM information is placed elsewhere in a quick circle and chronicled stockpiling. COLD and Frozen information is set away in a documenting, holding available to the above application. Putting away information with various freshness by stratification guarantees the pace of information preparing, and that the data which is at present in low esteem, however, may create new an incentive, later on, won't be erased. Of course, as the data scale creates with the extension of the customers' social occasion, to guarantee that the customer can get to and explore the tremendous scale enlightening accumulation set away on Hadoop with the most insignificant deferral and that the data obtainment, taking care of and examination in the Hadoop gathering can aggregate meanwhile, the eBay China Research and Development Department center started the OLAP around Hadoop adventure. The endeavor made metadata by the modeler by portraying the related estimations, and fabricated the metadata-based engine to normally create related Hive questions, MapReduce errands, and HBase exercises, so the data is examined out and pre-decided from HIVE, and the results are secured in HBase to give a natural request capacity of PB or even TB level, enlightening lists for front-end business customers and agents with only second measurement or even sub-second measurements delay.

A few shoppers are ending up progressively mindful of value segregation in Amazon.com. For example, CNN announced that a few clients of the Amazon are bothered over value separation on the cost of a specific DVD. One the purchasers revealed that the cost of a DVD in the wake of erasing treats on his PC, varied by \$2.50 edge. Another occasion, CNN detailed that the Amazon made utilization of dynamic estimating calculation while moving an item whoop "Jewel Rio MP3 Player" for \$51 not as much as its unique value.

Starting late, China cross-periphery online business has been creating fast. In 2017, the gross volume of China conveys electronic business accomplished 6.3 trillion Chinese Yuan with a yearly advancement rate of 14.5%. In China exchange web-based business, B2B speaks to 80.9% while B2C and C2C speak to simply 19.1%. B2B is up 'til now a direction show, yet B2C is depended upon to increase faster.

The central products of China exchanges internet business are 3C electronic products (20.8%), house and home items (6.5%). outside products (5.4%). Clothes (9.5%). In the year 2017, the standard objective countries of China cross online border business are the USA (15%), Russia (12.5%), France (11.4%), England (8.7%) and Brazil (6.5%) which demonstrates that the USA and some made countries in Europe are so far the essential objective countries while the as of late creating business part in countries like Latin America, Middle, and East Europe are growing quick [5].

A Positive element of implementing the Big Data examination request incorporates offering data look, a suggestion framework, a dynamic evaluating and client administration to collaborate with the network part. By gathering characteristic information in the Big Data period, for example, geographic circulation, enthusiastic propensities, client conduct on shopping just as the social association, side interests, organizations can accomplish request introduction, biased introduction, a relationship introduction, and different approaches to fulfill clients.

Informative search shows that data standard and looking management quality. Data quality is a proportion of significant worth seen by yield given by a site. Data properties, for example, refresh, valuable, nitty-gritty, exact, and finish has been recognized as essential segments of data quality.

Suggestion System incorporates an association amongst e-dealers and buyers whereby the buyers give their information, for instance, relaxation exercises and tendencies, while the seller offers a proposition fitting their prerequisites, like this benefiting both. Nuances are given on key measures behind

proposal systems: a customer based significant strain which used likeness in customer rankings to envision their interests and thing based network strain as centers in the space of words.

Giving a unique customer organization is the primary key to keep customers happy. Big Data engages you in improving your organizations. Using significant data analytics, you can overhaul your customer organization achieving progressively upbeat customers. A couple of customers may not simply protest of things or organizations through the official channels offered by the website, but may moreover go social about their get-together. You need data of such customers and exercise other alarm with the objective that grumblings of such customers are watched out for twofold quick. Enormous Data is secondhand to improve business shapes. Retailers can update their stock reliant on desires from web look designs, customer direct and atmosphere measures. One different application for the business procedure is the examination underway system or movement course. In light of scenery position and radio repeat recognition, the stimulus is used to pursue items or moving vehicles. This system enables customers to continue their solicitations. From that, customer organizations can be improved and increase shopper devotion [6].

Shopping habit is continuous and under perceived social dependence. Conduct compulsion is people's inability to see the quality of post-fixation longings and a failure to control want. For shopping addicts, shopping ends up uncontrolled, and they did not just purchase things they need, or they like, yet also genuinely spend their cash and are on edge to pass up on a decent chance to buy something. These items may is not utilized after buy. Utilizing the uses of Big Data investigation, the site can prescribe clients different things as a substitute or complementary pieces. This application is precious for clients with questions they need to purchase yet this is likewise unsafe for clients. They should invest more energy to audit more things to settle on a choice. It additionally suggests another correlative piece which the client feels they have to buy to build the obtained details. For instance, a client has purchased an exceptionally great pink dress, and the site prescribes her applicable sacks or shoes that are appropriate with the dress. They are wanted to consolidate together to give consumer loyalty. The client needs to invest energy and cash to purchase these corresponding items due to a decent chance to get them, even with less money. Shopping addictions are found to appear under two necessary measurements: propensity to spend and post-buy feeling.

The security of Big Data is another tremendous concern and one that increments with regards to Big Data. Because of the unmistakable attributes of Big Data in three-trade environment, it can identify with protection and security concern. The high volume and convergence of information makes an all the more engaging focus for programmers. Moreover, higher information volume builds the likelihood that the information files and reports may contain characteristically important and touchy data. Information with the end goal of Big Data examination are along these lines a potential goldmine for digital culprits. As of late, ponders demonstrated that there is an expanding shopper worry over protection with regards to constant social publicizing and attaching advances, for example, treats. The Internet publicizing firms Double Click and Avenue A, product firm Intuit and others have confronted claims for utilizing treats to target promoting. A high assorted variety of Big data lead to associations coming up short on the capacity to oversee and understand these information, and outsiders have chances to get to information. They may not conform to information insurance directions [7].

Results. In recent years, big data has become more popular as businesses of all sizes have discovered how to tap into it when it comes to making crucial decisions. Big data has assisted in finding and fixing problems, tracking business progress, and justifying causes for action. Although there is still much more that we can do when it comes to fully understanding big data, learning more means that we are able to utilize it to an even larger extent. And, many professionals believe that big data is the future of ecommerce, opening several doors for online businesses including analyzing sales, efficient transaction tracking, forecasting demand and supply, planning out expansion, and much more. So, how can big data help your ecommerce site?

1. Provide a Personalized User Experience

The ecommerce experience was once severely lacking in any personal touch, which prevented shoppers from building a connection with the business. Buyers had to go through all the products with no option for highlighting the ones that they were interested in. Thanks to big data analytics, today's ecommerce owners are able to recommend items that their users are actually looking for. You can use big data to track the products that shoppers are interested in, allowing you to send personalized recommendations to the user that really mean something for them.

2. Price More Competitively

Your ecommerce store's pricing strategy is one of the key determinants as to whether a shopper chooses your store or one of your competitors and it should be a vital part of your marketing strategy plan. It's easy for your users to check other ecommerce sites in order to find the best deal, which is why offering competitive prices is so important. To customers, it makes no sense to pay more for an item that they can get for less on another site. However, manually checking your competitor's sites on a regular basis would be a tedious task. Thanks to big data analytics, you have the ability to check how your competitors have priced their products, allowing you to respond accordingly.

3. Turn Visitors into Buyers

When it comes to putting time, effort and resources into understanding big data, your ultimate goal should be attaining a higher conversion rate or turning visitors into actual buyers. Big data allows you to get a clearer picture of why visitors are leaving a site without making a purchase, whether it's down to on-site issues such as load speed or shopper concerns such as lack of site security, high prices or inadequate logistics solutions. This allows you to make more informed improvement decisions, such as switching e-commerce platforms or providing better shipping options. If you want to convert customers into buyers it helps to treat your website like a shopfloor, build trust with your visitors and ensure your website is responsive. You can then evaluate the data before and after implementing these changes to measure what works best for your brand.

4. Better Manage Your Inventory

In addition to the above, you can also use big data to better manage your inventory. It allows you to automatically detect products that have been through the sales process successfully, in addition to tracking products that are returned. Big data analysis can also be used to track demand, allowing you to ensure that you get the right amount of inventory. Using this, you can determine the right number of products that are needed for certain periods, allowing you to stop wasting money on having too much or losing out on sales due to being out of stock.

5. Predict Trends

Every online retailer wants to know about the next bestselling product before their competitors. Utilizing big data is a strategy that you must not miss out on when it comes to predicting trends in the market. Trend forecasting algorithms allow retailers to eliminate the guesswork when it comes to the next big products, combing data from social media posts and user web browsing habits to determine which products are causing a buzz. Sentiment analysis is another big data strategy that retailers can use to determine the context in which products are being talked about online. Are conversations positive or negative? The data collected can be used to accurately predict the next popular products in various categories.

6. Optimize Customer Service

Last but not least, the customer service is a vital component of your ecommerce store and without it, you're going to lose out to your competition very quickly. In addition to competitive prices, popular products, and a personalized customer experience, your customers want to feel that they are truly valued and understood by your brand. Exceptional customer service is crucial for both customer satisfaction and retention in the retail industry, both online and offline. Big data allows you to better determine the service strategies that work well for your customers, in addition to those that don't have a great effect. As a result, you will be able to optimize your customer service strategies and provide an enhanced customer experience, allowing you to gain an edge over your competition.[8]

There are always beneficial ways to use big data to improve your ecommerce business, no matter what point you are in your company's growth. It's an especially exciting time to be a business owner. Do you use big data to improve your ecommerce store? We'd love to hear from you in the comments.

BP is based on a set of links between objects, which transmit to each other in some order of the key task. A step-by-step planned task or formed goal turns into a final result - a product or service.

New technologies of organization of administrative processes, which should provide the creation of really working formal mechanisms, require as the first step the creation of organizational and functional model of the enterprise. When forming the organizational-functional scheme of e-commerce enterprise the process approach of business structures formation was used. Using a business model to make all management decisions is a fundamental difference of the business-engineering approach. It represents a formal, accurate, complete and comprehensive description of the company. It allows, neglecting unnecessary details, to see the full picture of the problem from the most different angles, with the accuracy necessary for the solution of a specific management problem. [9]

At the stage of organizational design, a "vertical" systemic-targeted description of the company is made. At the same time, the processes occurring in the company are identified, classified and fixed in a convoluted form (as functions) according to the management hierarchy.

Static description of the company, the main purpose of which is to define: the business potential (the list and structure of commercial activities); company functionality (the list and hierarchy of functions reproduced on a regular basis); zones of responsibility of the personnel.

Features of typical organizational structure are as follows: final product of the company is focused on B2B segment, and the product of other manufactures is focused on B2C market segment, partial centralized planning of sales, absence of complete centralized key functions. [10]

Prices and plans are set by the management link, the production link adjusts them according to seasonality and equipment availability. The study of the organizational structure of the enterprise revealed the following problems:

1. The manager's focus shifts from more important decisions to operational issues, which can worsen the quality of strategic decisions.
2. Decentralized TB and OS (Safety and Environment).
3. Reputational risks in capital markets, insufficient speed of independent development of TB and OS culture in the enterprise.
4. Decentralized electronic procurement.
5. Possible overpricing, overstocking or understocking.
6. Decentralized electronic accounting (different approaches to accounting of counterparties, accounting objects, nomenclature groups).
7. Suboptimal management decisions, supply failures, duplication of inventory, delays in servicing customer requests
8. Duplication in planning.
9. Duplication in human resource development management.
10. Market service disruptions, equipment downtime, contract breakdowns.
11. Staff turnover, increased cost of FOT (Payroll Fund). [11]

Conclusion.

As the Big Data is used in various parts, it extensively smashes E-Commerce benefits and expects a vital activity in business choosing. The use of colossal data has significantly created in E-Commerce. Different gigantic retailer regards this present data's information and roots them for envisioning customer interests and give their customers similar and captivated looks for when they shop on their site. The objective that they attract the customer by providing the necessary and critical endeavors of things or things. Using the related information from this paper, the examiners can come up with vital and testing systems to expanding the upsides of Big data apply toward online business for both the customers similarly as the retailers. In this engaged and brisk condition customers generally, keep running with the online notification or through web crawlers by decreasing the inefficiency of the continuous markets.

Our paper helps to perceive the various use of extensive data into web business so we can know the criticalness of astronomical data, it improves understanding of Usage of significant data and its fragments.

We similarly talk about by the survey made on the issues related to E-Commerce if huge data isn't compared to that, so pros can tackle the problems associated to Big data and extend their work on that. An authoritative trial of Big Data examination is to make business regard from their impact of significant data. We also have analyzed the genuine troubles related to reliable data, so we expand our investigation in finding a response to one of the challenges identified with that. Researchers can get information about the issues concerning large data and critical troubles identified with that. So they can get concise information about colossal data which supports them in expanding their examination tackle extensive data related to online business.

Thus, the construction of the organizational-functional model of the enterprise should be provided with effective information technologies that integrate organizational structures, BP, functions and document management. Of course, in this case an effective technology of regular updating both the business model itself and all its elements should be functioning.

With regard to the design of the tactical component of the model the following design recommendations can be made. In its essence BP is aimed at obtaining certain outputs. All evaluations of BP quality, as well as planned changes to improve its characteristics, should be projected onto the output results in one way or another. For this reason it is necessary to clearly establish and formalize what direct or indirect influence processes, subprocesses and procedures have on outputs of the appropriate level. To do this, the model was designed to form a specific environment and attributes of the outputs, which can be used to establish: which processes, subprocesses and procedures are used to deliver specific outputs; what resources (organizational, informational and technological) are required to deliver them; what generalized time and cost estimates of outputs; what output structure, i.e. what composition of intermediate results it includes; what outputs are part of; what outputs are to be delivered and how they are delivered?

As for output documents it is reasonable to use the following categorization as they move along the process chain: template, new document, changed (in development), agreed (approved), approved, registered, archived.

Each status indicates the current state of the document in the BP. Change of the document status occurs at fulfillment of certain actions with the document and characterizes the passage of this document of a certain stage of document circulation.

In any case, a concrete list of output document statuses and their meaning is defined when solving the concrete task. When working with the contract the document can have such additional statuses as "partially executed", "executed"; when working with payment documents via Internet-client system the use of additional statuses such as "sent to the bank", "received by the bank", "rejected by the bank" and so on is possible.

Distribution of functionality into logical layers is one of the most effective methods of structuring a complex system, which ensures reduction of connectivity of its components.

References

- [1]. H. Liu and X. Wang, Shanghai, 2018, "Study on the application of big data in accurate marketing of cross-border e-commerce in China," 2018 IEEE 3rd International Conference on Big Data Analysis (ICBDA), pp. 24-27.
- [2]. X. Zhao, "A Study on the Applications of Big Data in Cross-Border E-Commerce," 2018 IEEE 15th International Conference on e-Business Engineering (ICEBE), Xi'an, 2018, pp. 280-284.
- [3]. E. Makki and L. Chang, "Leveraging social big data for performance evaluation of E-commerce websites," 2016 IEEE International Conference on Big Data (Big Data), Washington, DC, 2016, pp. 2525-2534.
- [4]. Shahriar Akter & Samuel Fosso Wamba, 2016, "Big data analytics in E-commerce: a systematic review and agenda for future research", IIM University of St. Gallen.
- [5]. S. Suguna, M. Vithya and J. I. C. Eunaicy, "Big data analysis in e-commerce system using HadoopMapReduce," 2016 International Conference on Inventive Computation Technologies (ICICT), Coimbatore, 2016, pp. 1-6
- [6]. S. Pradeep and J. S. Kallimani, "A survey on various challenges and aspects in handling big data," 2017

International Conference on Electrical, Electronics, Communication, Computer, and Optimization Techniques (ICEECCOT), Mysuru, 2017, pp. 1-5.

[7]/ H. Xu, K. Li and G. Fan, “Novel Model of E-Commerce Marketing Based on Big Data Analysis and Processing,” 2017 International Conference on Computer Network, Electronic and Automation (ICCNEA), Xi’an, 2017, pp. 80-84.

[8]. MY Abdul-Azalova, NM Mamatova “Big data technologies in technological and business processes automatization” International Conference on Big Data and Advanced Analytics, 2021

[9] IN Tsyrelchuk, NM Mamatova, MY Abdul-Azalova “Optimization of business processes via Big Data” International Conference on Big Data and Advanced Analytics,– 2020.

[10]. Abdul-Azalova Makhina Yashnarovna “Issues of Creating an Intelligent Automated Business Process Management System” Journal “European Multidisciplinary Journal of Modern Science” Volume 4, pp. 691-697.

[11] MY Abdul-Azalova “Modernization of technologies and business process management systems” Форум молодых исследователей, 2022, pp. 188-190.

НЕКОТОРЫЕ ВОПРОСЫ ПРИМЕНЕНИЯ BIG DATA ПРИ МОДЕЛИРОВАНИЯ БИЗНЕС-ПРОЦЕССОВ СИСТЕМЫ ЭЛЕКТРОННОГО БИЗНЕСА

М. Я. Мансурова

*Доцент кафедры «Информационные
технологии» ТУИТ, PhD*

*Ташкентский университет информационных технологий имени Мухаммада ал-Хорезми, Республика
Узбекистан*

E-mail: m.ya.mansurova@gmail.com

Аннотация. Большие данные – набор огромной полезной информации, которая не может быть прочитана с помощью стандартных вычислительных структур. Большие данные – это не только данные, они уже стали целой областью, которая включает набор инструментов, контекстов и структур. Она использует сложные наборы данных для выбора направления, курса и проведения непосредственного управления внутри организаций. С помощью усовершенствования и выполнения вычислений могут быть получены важные структуры, необходимые для точного и глубокого понимания полученных результатов через исследование данных организации. В этом исследовании мы изучили разные типы данных и их использование для бизнес-процессов электронного бизнеса, и особые методы для обеспечения безопасности и развития данных для использования в сложных организациях. Кроме того, мы рассмотрели вопросы больших данных, касающихся интернета, и пути их использования в электронном бизнесе.

Ключевые слова: бизнес-процесс, веб-аналитика, прогнозируемая аналитика, персонализация, динамический анализ, анализ тональности, большие данные, электронная торговля, электронный бизнес.

UDC 004.6

APPLICATION OF BIG DATA TECHNOLOGY



N.A. Naim

*Assistant at the Tashkent University of Information Technologies named after Muhammad al Khwarezmi
naimnodira02@gmail.com*

N.A. Naim

She graduated from the Tashkent University of Information Technologies named after Muhammad al Khwarezmi. TUIT teacher on the estate of Muhammad al Khwarezmi

Annotation. Big Data technologies make it possible to process a huge amount of unstructured data, analyze it, systematize it and identify patterns where a person would not have found them. This article discusses the advantages and limitations of using Big Data technologies, with particular emphasis on assessing the degree of their impact on innovation management and economic efficiency today.

Keywords: Information, Big data, services, SQL, Hadoop, Sap Hana

Today, data volumes are increasing at a rapid pace. In order to somehow gain competitive advantages, respond faster to market changes and significantly improve efficiency, it is necessary to analyze and process a very large amount of different information. To interact with huge amounts of data, programmers needed to improve the tools to work on the analysis of all this data. Thus, in the early 2000s, the concept of Big Data appeared, which at first was of interest only to a very narrow circle of specialists. Now everyone who is interested in the field of IT knows this word.

Big Data is becoming an increasingly popular and strategically important area of IT development. Global information has always been of exceptional importance. The results of processing a huge amount of information are used to identify trends and patterns. For large companies, statistics and data analysis have always been at the heart of doing business in large markets, but with the advent of a huge amount of information, an analytical approach has become much more in demand. The purpose of the study is to identify the advantages and limitations of using Big Data technologies, with particular emphasis on assessing the degree of their impact on innovation management and economic efficiency today [1]. The object of the study is Big Data. The subject of the research is Big Data application technologies. The hypothesis of the study is that the application of Big Data technologies will be successful if the problems of solving the processing of large volumes of unstructured data are analyzed, and their systematization is ensured when working with huge amounts of information.

The objectives of the study are:

- To identify the emergence and development of the use of Big Data technologies;
- Conduct an analysis of the application of Big Data technologies;
- Conduct a comparative analysis of the traditional database and Big Data;
- Consider the scope and use of Big Data;
- Learn Big Data collection methods.

In today's world, one of the key aspects of IT development is Big Data.

This concept refers to the processing of information of different composition and huge volume, very quickly updated, located in various sources to increase efficiency, create new products and increase competitiveness. Following from this, Big Data is, firstly, a set of technologies, tools, methods and

approaches designed to solve the problem of processing large amounts of data, and secondly, Big Data is understood as the amount of data that cannot be processed by generally accepted ones, then eat in the traditional way. Big Data technologies make it possible to process a huge amount of unstructured data, analyze it, systematize it and identify patterns where a person would not have found them. The main thing to note is that the volume of data processed through Big Data is constantly increasing, as is the processing speed. The process of development of this direction corresponds to the modern world, fast-paced and innovative [2].

Examples of using Big Data:

- For any large company, Big Data allows you to analyze income and expenses, as well as detail the details of the production chain and logistics.

These factors help to improve the demand forecast for the product, reduce costs and downtime.

- In medicine, Big Data can help with the analysis of drug use statistics, improving the efficiency of service delivery.

- Banks, working with transactional information, use distributed computing, which is useful for detecting fraud and improving the operation of services.

- Government agencies analyze big data to improve the safety of citizens and improve urban infrastructure, improve the work of housing and communal services and public transport.

These are just some of the areas where the demand for big data analytics is growing. Interested parties include not only technical areas, but also media, marketing, sociology, hiring, real estate. Following from this, Big Data is already an established field of technology, even despite its relatively young age, which is gaining popularity in almost all areas of business and plays a large role in the development of the company. Big Data technologies:

The technologies used for processing and collecting Big Data are divided into three groups:

- Services;
- Equipment;
- BY.

There is still no universal toolkit for working with big data, but, despite all the difficulties, Big Data provides invaluable knowledge for various industries. Commonly used data collection methods include the following:

Table 1.

Method	characteristic
SQL	A programming language that allows you to work with databases
NoSql	Contains a series of approaches designed to implement a database
MapReduce	The basic principle of operation is to sequentially process data in two ways Redusy and map. Map takes preliminary data. Reduse aggregates them.
Hadoop	It is used for search and contextual mechanisms of highly loaded sites- Facebook, eBay, Amazon and etc.
Sap Hana	Provides high query processing speed another distinctive feature is the Sap hana tab simplifies the system landscape, reducing the cost of supporting analytical systems

Big Data problems. Usually there are 3 main groups of Big Data system problems volume, processing speed and unstructuredness: V-Volume, Velocity and Variety.

Certain conditions are needed to store huge amounts of data, and this is a matter of space and capabilities. Speed refers not only to the slow processing of information that is caused by old methods of processing, but also it is also a matter of interaction - the faster the process, the greater the return, the more productive the result. The problem of heterogeneity and unstructuredness appears due to the fragmentation of sources. We need certain analytical systems and tools for efficient processing and merging of data.

Another problem may be the analysis algorithm and the choice of data for information processing, since there is no specific understanding of what information should be collected and stored, and which can be excluded. There is also an obvious shortage of qualified specialists in this field [3].

Conclusion.

Based on the foregoing, we can draw the following conclusions: Big Data technologies involve working with huge amounts of information. There is no universal method for processing Big Data, but it is possible to use various methods to partially solve this problem. The successful application of the concept of Big Data in any enterprise can seriously increase work efficiency and stimulate the creation of a new product. The development of Big Data processing technologies is a very promising area of activity.

References

[1] Наим Н.А., Big data in modern business, Сборник научных статей VIII международной научно-практической конференции. “Big data и анализ высокого уровня”. Минск– 2022. – С.221-223.

[2] Наим Н.А., Алимова Ф.М., Кушманова М.А., Big data в медицине, Сборник научных статей VII международной научно-практической конференции. “Big data и анализ высокого уровня”. Минск– 2021. – С. 221-223.

[3] Алимова Ф.М., Кушманова М.А., Наим Н.А., Перспективы использования больших данных в системе образования, Сборник научных статей VII международной научно-практической конференции. “Big data и анализ высокого уровня”. Минск– 2021. – С. 193-196.

ПРИМЕНЕНИЕ ТЕХНОЛОГИЙ BIG DATA

Н.А. Наим

Ассистент в ТУИТ имени

Мухаммада ал-Хорезми

Окончила Ташкентский университет информационных технологий имени Мухаммада ал Хоразмий. Преподаватель ТУИТ имени Мухаммада ал Хоразми.

E-mail: naimnodira02@gmail.com

Аннотация. Технологии Big Data дают возможность обработать огромный объем неструктурированных данных, проанализировать их, систематизировать и выявить закономерности там, где человек бы их, не обнаружил. В этой статье рассмотрены преимущества и ограничения при использовании технологий Big Data, сделав особый акцент на оценку степени их влияния на управление инновациями и экономическую эффективность на сегодняшний день.

Ключевые слова. Information, Big data, services, SQL, Hadoop, Sap Hana

УДК 004.65:004.451

ANALYZING THE LARGE AMOUNT OF DATA



V.S.Sodikov

Assistant professor of Tashkent University of Information Technologies named after Muhammad al-Khorazmi
vsodiqov9620@mail.ru



A.I.Tukhtasinov

Assistant professor of Tashkent University of Information Technologies named after Muhammad al-Khorazmi
t.adhamjon@mail.ru

V.S.Sodikov

He graduated from the Tashkent University of Information Technologies named after Muhammad al-Khorazmi with a bachelor's degree and a master's degree. Assistant teacher of Tashkent University of Information Technologies named after Muhammad al-Khorazmi.

A.I.Tukhtasinov

He completed his bachelor's degree at Fergana State University. He completed his master's degree at the Tashkent University of Information Technologies named after Muhammad al-Khorazmi. Assistant teacher of Tashkent University of Information Technologies named after Muhammad al-Khorazmi.

Abstract. Big data analysis is a rapidly growing field that involves extracting insights and knowledge from large and complex data sets. The advent of technologies such as cloud computing, data mining, and machine learning has made it possible to process and analyze vast amounts of data in real-time. Big data analysis has the potential to revolutionize businesses, industries, and society as a whole by uncovering patterns, trends, and correlations that were previously hidden. It can also enable more accurate predictions and decision-making, as well as identify new opportunities for innovation and growth.

However, it also poses significant challenges, including data privacy and security, ethical concerns, and the need for skilled professionals to extract meaningful insights from the data. As big data continues to grow and evolve, the field of big data analysis will continue to play an increasingly important role in shaping our world.

Key words: big data, analytics techniques, Data integration, Healthcare, Finance, effectively analyze, Data quality, Data privacy, social media, Manufacturing

Introduction.

In today's digital age, businesses are generating more data than ever before. From customer transactions to social media interactions, there is an abundance of data available that can provide valuable insights into consumer behavior and market trends. However, analyzing this data can be a daunting task. This is where big data analysis comes into play.

What is big data analysis?

Big data analysis is the process of examining large and complex datasets to uncover hidden patterns, correlations, and insights that can inform business decisions. The data may come from a variety of sources, including social media, sensors, web logs, and customer transactions.

Traditional data analysis techniques, such as spreadsheets and statistical analysis, are often insufficient when dealing with big data. This is because big data is typically characterized by the 3 V's: volume, velocity, and variety.

Volume refers to the massive amount of data that needs to be analyzed. Velocity refers to the speed at which the data is generated and needs to be analyzed. Variety refers to the diverse types of data, including structured and unstructured data, that need to be analyzed.

Materials and methods.

While big data analysis has the potential to revolutionize the way businesses operate, it also comes with several challenges. Here are some of the main challenges associated with big data analysis:

Data quality: With such large and complex datasets, ensuring data quality can be a challenge. It's important to ensure that the data is accurate, complete, and up-to-date.

Data privacy: Big data analysis often involves sensitive personal information. Ensuring that this data is kept secure and private is essential.

Data integration: Big data often comes from a variety of sources and may be stored in different formats. Integrating this data into a single system can be challenging.

Technical skills: Big data analysis requires advanced technical skills, including programming, data mining, and machine learning.

Cost: The cost of storing and processing large amounts of data can be significant, and not all businesses have the resources to invest in this technology.

Why is big data analysis important?

Big data analysis can provide businesses with valuable insights that can inform decision-making and improve overall performance. Here are some of the ways big data analysis can be beneficial:

Improved customer experiences: Big data analysis can help businesses understand customer preferences and behavior, enabling them to offer more personalized and targeted products and services.

Better decision-making: By analyzing data from various sources, businesses can make more informed decisions that are based on accurate and up-to-date information.

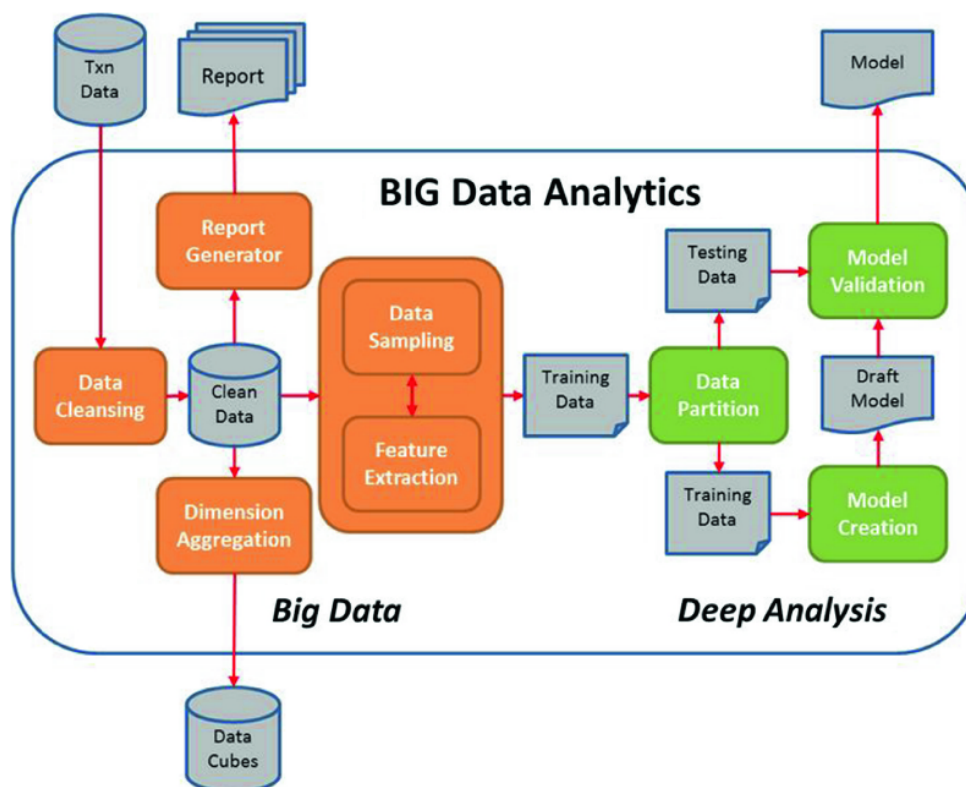


Image 1. Pragmatic Programming Techniques: BIG Data Analytics

Increased efficiency: Big data analysis can help businesses identify inefficiencies and areas for improvement, leading to cost savings and increased productivity.

Improved risk management: By analyzing data, businesses can identify potential risks and take steps to mitigate them.

Competitive advantage: Businesses that can effectively analyze big data can gain a competitive advantage by staying ahead of market trends and customer preferences.

How is big data analysis done?

Big data analysis typically involves the use of advanced analytics techniques, including machine learning, data mining, and natural language processing.

These techniques can help businesses identify patterns and insights that may not be immediately apparent using traditional data analysis methods.

Results.

To effectively analyze big data, businesses must have the right tools and technologies in place. This includes data storage and processing systems, as well as software tools for data visualization and analysis.

Big data analysis requires advanced technical expertise in areas such as data science, machine learning, and statistics.

Businesses may need to invest in training or hire specialized professionals to conduct effective big data analysis.

Big data analysis can be costly, requiring significant investments in technology, infrastructure, and personnel.

Applications of big data analysis:

Big data analysis has numerous applications across various industries. Here are some examples:

❖ **Healthcare:** Big data analysis can help healthcare providers identify patient trends and improve the quality of care.

❖ **Retail:** By analyzing customer data, retailers can improve inventory management and provide personalized recommendations to customers.

❖ **Finance:** Big data analysis can help financial institutions identify fraud and manage risk.

❖ **Manufacturing:** By analyzing data from sensors and other sources, manufacturers can improve production efficiency and reduce downtime.

❖ **Transportation:** Big data analysis can help transportation companies optimize routes and improve the customer experience.

Conclusion.

Big data analysis has the potential to transform the way businesses operate by providing valuable insights into consumer behavior and market trends. While it comes with several challenges, the benefits of big data analysis are clear.

By investing in the right tools and technologies, businesses can gain a competitive advantage and improve overall performance. Big data analysis is a powerful tool that can provide businesses with valuable insights into consumer behavior and market trends. By analyzing large and complex datasets, businesses can make more informed decisions, improve customer experiences, and gain a competitive advantage. With the right tools and technologies in place, big data analysis can be a game-changer for businesses in any industry.

References

- [1]. Adams, M.N.: Perspectives on Data Mining. *International Journal of Market Research* 52(1), 11–19 (2010)
- [2]. Asur, S., Huberman, B.A.: Predicting the Future with Social Media. In: *ACM International Conference on Web Intelligence and Intelligent Agent Technology*, vol. 1, pp. 492–499 (2010)
- [3]. Bakshi, K.: Considerations for Big Data: Architecture and Approaches. In: *Proceedings of the IEEE Aerospace Conference*, pp. 1–7 (2012)
- [4]. Cebr: Data equity, Unlocking the value of big data. in: *SAS Reports*, pp. 1–44 (2012)
- [5]. Cohen, J., Dolan, B., Dunlap, M., Hellerstein, J.M., Welton, C.: MAD Skills: New Analy-sis Practices for Big Data. *Proceedings of the ACM VLDB Endowment* 2(2), 1481–1492 (2009)
- [6]. Cuzzocrea, A., Song, I., Davis, K.C.: Analytics over Large-Scale Multidimensional Data: The Big Data Revolution! In: *Proceedings of the ACM International Workshop on Data Warehousing and OLAP*, pp. 101–104 (2011)
- [7]. Economist Intelligence Unit: The Deciding Factor: Big Data & Decision Making. In: *Capgemini Reports*, pp. 1–24 (2012)

АНАЛИЗ БОЛЬШИХ ДАННЫХ

В.С. Содиков

*Доцент Ташкентского университета
информационных технологий имени
Мухаммада аль-Хорезми*

А.И. Тухтасинов

*Доцент Ташкентского университета
информационных технологий имени
Мухаммада аль-Хорезми*

Ташкентский университет информационных технологий имени Мухаммада аль-Хорезми, Республики Узбекистан

E-mail: vsodiqov9620@mail.ru

Аннотация. Анализ больших данных – это быстро развивающаяся область, которая включает в себя извлечение идей и знаний из больших и сложных наборов данных. Появление таких технологий, как облачные вычисления, интеллектуальный анализ данных и машинное обучение, позволили обрабатывать и анализировать огромные объемы данных в режиме реального времени. Анализ больших данных может произвести революцию в бизнесе, отраслях и обществе в целом, обнаружив закономерности, тенденции и корреляции, которые ранее были скрыты. Это может обеспечить более точные прогнозы и принятие решений, а также выявить новые возможности для инноваций и роста.

С другой стороны, это может создать серьезные проблемы, включая конфиденциальность и безопасность данных, этические проблемы и потребность в квалифицированных специалистах для извлечения осмысленной информации из данных. Поскольку большие данные продолжают расти и развиваться, область анализа больших данных будет продолжать играть все более важную роль в формировании нашего мира.

Ключевые слова: большие данные, методы аналитики, интеграция данных, здравоохранение, финансы, эффективный анализ, качество данных, конфиденциальность данных, социальные сети, производство.

УДК 616.89-008+004.42+004.9

ОЦЕНКА ЭФФЕКТИВНОСТИ АВТОМАТИЗАЦИИ УСТАНОВКИ ДЛЯ ОХЛАЖДЕНИЯ МОЛОКА С ИСПОЛЬЗОВАНИЕМ ЧАСТОТНОГО УПРАВЛЕНИЯ



И.А. Оганезов

Доцент кафедры экономики и организации предприятий АПК УО «Белорусский государственный аграрный технический университет», кандидат технических наук, доцент
iaoganezov@bsuir.by



Н. В. Щербина

Старший преподаватель кафедры инженерной психологии и эргономики УО «Белорусский государственный университет информатики и радиоэлектроники», магистр технических наук
shcherbina@bsuir.by



А.В. Буга

Доцент кафедры экономики Северо-Западного института управления Российской академии народного хозяйства и государственной службы при Президенте Российской Федерации, кандидат экономических наук, доцент

И.А. Оганезов

Доцент кафедры экономики и организации предприятий АПК УО «Белорусский государственный аграрный технический университет», кандидат технических наук, доцент. Проводит научные исследования в области энергоэффективных технологий в АПК.

Н. В. Щербина

Старший преподаватель кафедры инженерной психологии и эргономики УО «Белорусский государственный университет информатики и радиоэлектроники», магистр технических наук. Проводит научные исследования в областях промышленной безопасности, эргономики и безопасности труда.

А.В. Буга

Доцент кафедры экономики Северо-Западного института управления Российской академии народного хозяйства и государственной службы при Президенте Российской Федерации, кандидат экономических наук, доцент. Проводит научные исследования в областях антикризисного управления, стратегического управления научно-инновационной деятельностью в АПК, ее экономической, социальной и экологической эффективности.

Аннотация. Экономически обоснована перспективность охлаждения молока на отечественных молочно-товарных фермах, построенным по типовым проектам при использовании систем автоматизации на основе частотно-регулируемых электроприводов. Предложенные технические решения могут позволить обеспечить существенное снижение расхода электроэнергии и затрат на оплату труда обслуживающего персонала при повышении качества охлажденного молока.

Ключевые слова: ферма, молоко, охлаждение, автоматизация, частотное управление, экономия, эффективность.

Введение.

Одной из основных задач в первичной обработке молока на отечественных молочно-товарных фермах (МТФ) является сохранение высокого качества производимого молока за счет своевременного и надежного его охлаждения.

Исследования, проводимые по проблеме снижения удельного расхода электроэнергии на производство и переработку молока, показали актуальность конкретного направления научно-технического поиска и востребованность инженерных решений в области разработки энергосберегающих процессов на основе использования автоматизированного частотно-регулируемого электропривода.

Качество молока и молочных продуктов во многом зависит от своевременности их обработки и переработки, так как молоко является скоропортящимся продуктом. Охлаждение молока позволяет сохранить его естественные качества и пищевую ценность. В целях сохранения молока в свежем виде в период доставки потребителям его подвергают первичной обработке непосредственно на фермах. Эта обработка включает следующие технологические операции: фильтрацию, охлаждение, хранение, учет. В некоторых случаях к ним добавляются пастеризация, сепарирование и нормализация.

При доении в него попадают бактерии, вызывающие закисание. Источником бактериального загрязнения могут быть плохо вымытое вымя животного, плохо промытые детали, соприкасающиеся с молоком, и воздух коровника, засасываемый пульсатором и коллектором доильного аппарата. Свежевыдоенное молоко обладает бактерицидными свойствами, то есть способностью молока в течение определенного времени препятствовать росту бактерий, которые сохраняются определенное время. Продолжительность бактерицидной фазы зависит от степени и скорости охлаждения молока. Понижая температуру молока, увеличиваем срок действия его бактерицидных свойств. У свежего неохлажденного молока при $t = 30^{\circ}\text{C}$ бактерицидная фаза равна 3 часам, при снижении температуры до 16°C – 76 ч, до $10-13^{\circ}\text{C}$ – 36 ч, до $4-5^{\circ}\text{C}$ жизнедеятельность бактерий практически прекращается.

Из выше изложенного следует, что первичное охлаждение молока является одним из важнейших условий получения качественного продукта.

В связи с этим было предложено произвести замену действующей установки охлаждения молока на новую с более высокой холодопроизводительностью.

В базовом варианте используется танк-охладитель молока с нерегулируемым электроприводом основного технологического узла – компрессора. Управление работой всех элементов холодильного агрегата производится в ручном режиме.

В проектируемом варианте предусматривается использование аналогичного технологического оборудования, однако привод компрессорного агрегата должен быть регулируемым с использованием частотного преобразователя (ЧП). Управление данным электроприводом автоматическое при помощи терморегулятора и датчиков температуры молока. Кроме того, автоматизирована работа перемешивающих устройств с использованием реле времени, что исключает участие оператора в процессе управления. Предложенная схема управления обеспечивает экономию тепла и электроэнергии.

Основные характеристики и преимущества проекта:

- высокий уровень гигиены резервуара благодаря очистке с возможностью выбора оптимального режима промывки;
- исключение риска замораживания молока благодаря системе контроля льдообразования; улучшенный контроль за функционированием танка-охладителя благодаря системе управления;
- увеличенная на 20 % изоляция, позволяющая снизить расходы на ремонт и техническое обслуживание технологического оборудования, повысить его срок службы;
- более рациональный расход моющих средств, благодаря автоматическому дозированию; более высокая степень фильтрации молока, благодаря усовершенствованной системе фильтрации;
- возможность выбора более рационального режима охлаждения; более высокий уровень очистки резервуара и молокопровода, благодаря системе продувки молокопровода;
- возможность повторного использования воды, благодаря наличию клапана для отделения промывочной воды;
- более высокие показатели надежности и срока службы.

Поэтому тема данной научной работы является актуальной для экономики АПК Республики Беларусь.

Материалы и методы.

С учетом поставленных задач в работе применялись методы исследования: экономико-статистический, монографический, абстрактно-логический, расчетно-конструктивный, интервьюирования и др. При разработке приоритетных направлений использования новых технологий на основе автоматизированного частотно-регулируемого электропривода применительно к условиям организации ПРУП «Экспериментальная база имени Котовского», д. Озеро, Узденский район, Минская область Республики Беларусь, применялись экспертно-аналитический, экономико-статистические, расчетно-конструктивный и монографический методы.

Информационной базой исследования являются отраслевые справочно-нормативные материалы, положения и рекомендации специализированных научно-исследовательских учреждений, данные статистических органов и Министерства сельского хозяйства и продовольствия Республики Беларусь, а также результаты лабораторных и хозяйственных испытаний на базе организации ПРУП «Экспериментальная база имени Котовского», д. Озеро, Узденский район, Минская область Республики Беларусь.

Результаты.

За базовый вариант нами принята установка охлаждения молока МОУ-4000Г в стандартной комплектации.

За проектируемый вариант принимаем установку охлаждения молока МОУ-4000Г со схемой управления электроприводом на основе преобразователя частоты для электропривода компрессора (см. рисунок 1).



Рисунок 1. Установка охлаждения молока МОУ-4000Г со схемой управления электроприводом на основе преобразователя частоты для электропривода компрессора

Основные преимущества проекта: с помощью схемы управления на основе преобразователя частоты обеспечивается поддержание требуемого давления теплоносителя компрессора для обеспечения необходимой температуры охлаждения молока с высокой точностью, снижение негативного влияния пусковых токов на электродвигатель при его запуске, снижение расхода электроэнергии, возможность выбора оптимального режима охлаждения; исключение риска замораживания молока благодаря системе контроля льдообразования.

Показатели эффективности инвестиций в проект рассчитаны в таблице 1.

Таблица 1. Техничко-экономические показатели проекта (без учета НДС)

Показатели	Вариант 1	Вариант 2	Изменения (3-2)
1.Поголовье, голов	200	200	–
2.Годовой объем производства молока, т	1080	1080	–
3.Затраты труда, ч/год	756,3	400,29	-356,01
4.Установленная мощность установки охлаждения молока, кВт	6,8	6,8	–
5.Расход электроэнергии, кВт·ч/год	10841	6977	-3864,00
6.Удельный расход электроэнергии, кВт·ч/т	10,04	6,46	-3,58
7. Балансовая стоимость действующего оборудования, у.е.	7445	7445	–
8. Капиталовложения в новое оборудование без НДС, у.е.	–	1000,26	–
9. Эксплуатационные издержки, у.е., в том числе:	3208,57	2604,2	-604,37
– расходы на оплату труда	635,29	336,24	-299,05
– отчисления на социальные нужды	190,59	100,87	-89,72
– амортизационные отчисления	744,5	844,53	100,03
– затраты на техническое обслуживание и ремонт	446,7	506,72	60,02
– затраты на электроэнергию	899,80	579,09	-320,71
– прочие издержки	291,69	236,75	-54,94
10. Прирост чистой прибыли, у.е.	–	604,37	–
11. Годовой доход, у.е.	–	704,40	–
12. Чистый дисконтированный доход, у.е.	–	2359,73	–
13. Срок окупаемости капиталовложений, лет:			
– статический	–	1,42	–
– динамический	–	1,71	–
14. Индекс доходности проекта, у.е.	–	3,36	–
15. Внутренняя норма доходности, %		73,35	
Примечание: 1 у.е. соответствует доллару США по курсу Национального банка Республики Беларусь			

Предложенная система управления может обеспечить существенную экономию электроэнергии (более 35 %) и затрат на оплату труда и социальные нужды (более 45 %), эксплуатационных расходов на 19 % в год. Чистый дисконтированный доход составил 2359,73 у.е. Он больше нуля. При этом расчетный срок окупаемости капитальных вложений не более 2-х лет. Внутренняя норма доходности, ВНД = 73,35 % больше ставки дисконтирования (15 %). Индекс доходности проекта, ИД = 3,36 > 1.

Данные показатели также свидетельствуют о экономической целесообразности рассматриваемого энергосберегающего инвестиционного проекта.

Расчет наличного чистого дисконтированного потока денежных средств проекта с учетом НДС приведен в таблице 2.

Таблица 2. Расчет наличного чистого дисконтированного потока денежных средств проекта с учетом НДС

Показатель	Годы реализации проекта					
	2023	2024	2025	2026	2027	2028
Инвестиции с НДС, у.е.	1200,31	–	–	–	–	–
Прибыль отчетного года, у.е.	0	604,37	604,37	604,37	604,37	604,37
Амортизационные отчисления, у.е.	0	100,03	100,03	100,03	100,03	100,03
Итого приток, у.е.	0	704,4	704,4	704,4	704,4	704,4
Итого отток, у.е.	1200,31	0	0	0	0	0

Продолжение таблицы 2

Показатель	Годы реализации проекта					
	2023	2024	2025	2026	2027	2028
Сальдо, у.е.	-1200,31	704,4	704,4	704,4	704,4	704,4
Коэффициент дисконтирования	1,0000	0,870	0,756	0,658	0,572	0,497
Дисконтированный приток, у.е.	0	612,83	532,53	463,50	402,92	350,09
Дисконтированный отток, у.е.	-200,31	0	0	0	0	0
Чистый дисконтированный поток наличности, у.е.	-1200,31	612,83	532,53	463,50	402,92	350,09
Чистый дисконтированный поток с нарастающим итогом, у.е.	-1200,31	-587,48	-54,95	408,55	811,47	1161,56
ИТОГО Чистый дисконтированный доход, у.е.	2361,87					

Примечание: 1 у.е. соответствует доллару США по курсу Национального банка Республики Беларусь

Финансовый профиль и срок окупаемости проекта с учетом НДС приведены на рисунке 2.

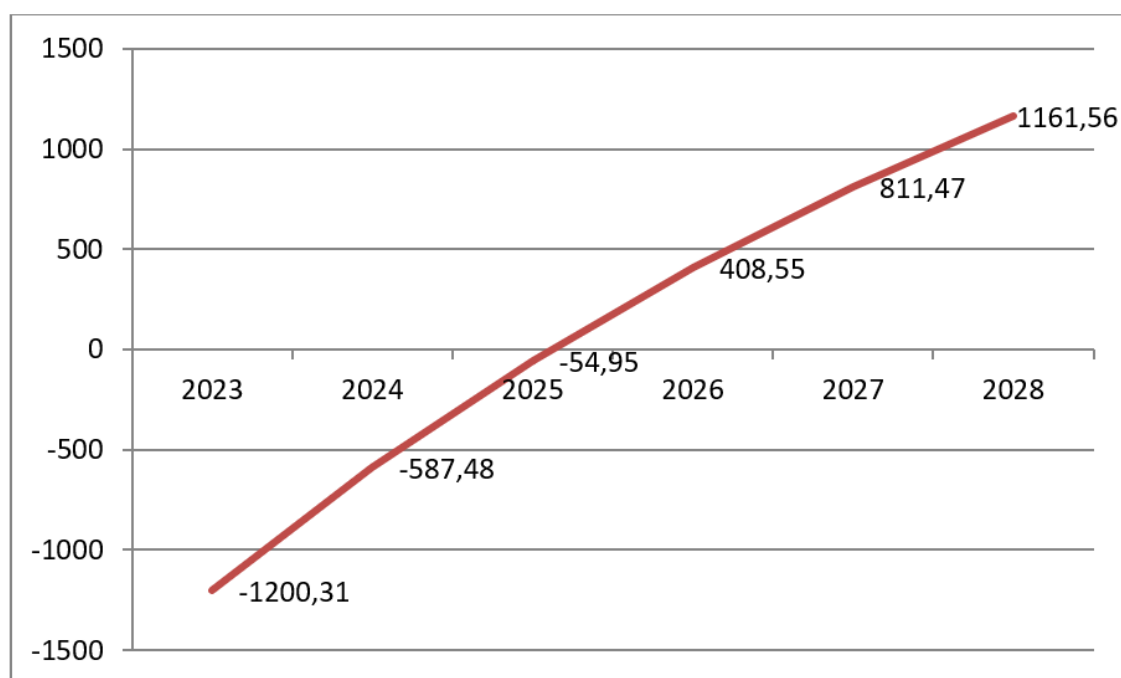


Рисунок 2. Финансовый профиль и срок окупаемости проекта с учетом НДС, у.е.

Исходя из приведенных расчетов, следует, что внедрение системы автоматизированного частотно-регулируемого электропривода для охлаждения молока на МТФ ПРУП «Экспериментальная база имени Котовского», д. Озеро, Узденский район, Минская область Республики Беларусь, спроектированной и построенной по типовому проекту экономически целесообразно.

Заключение.

1. Анализ передового отечественного и зарубежного опыта показывает значительный эффект результатов от внедрения автоматизированного частотно-регулируемого электропривода для обеспечения основных групп потребителей электроэнергией, в том числе в молочном животноводстве на основах:

– подачи конкретному потребителю электрической энергии, необходимой мощности $P_{\text{вых}}$;

– выбора рационального с технической и экономической точек зрения регулируемого электропривода, позволяющего управлять скоростью в нужном диапазоне с минимальными потерями $\Delta P_{\text{дв}}$;

– обоснования рационального с технической и экономической точек зрения управления координатами, образующими потребляемую технологическими машинами мощность.

Переход от нерегулируемого асинхронного электропривода к электроприводу с регулируемой частотой вращения в молочном животноводстве во многих случаях позволил не только снизить потребление электроэнергии за счет рационализации технологического процесса, но уменьшить износ технологического и электрического оборудования, повысить надежность его эксплуатации и увеличить ресурс.

2. По результатам экспериментального исследования от внедрения системы автоматизированного частотно-регулируемого электропривода для охлаждения молока на МТФ ПРУП «Экспериментальная база имени Котовского», д. Озеро, Узденский район, Минская область Республики Беларусь, спроектированной и построенной по типовому проекту были получены следующие значения эффекта в натуральном и стоимостном выражении:

– существенная экономия электроэнергии (более 35 %) и затрат на оплату труда и социальные нужды (более 45 %);

– общие эксплуатационные издержки могут быть снижены на 19 %; чистый дисконтированный доход в размере 2359,73 долл. США;

– внутренняя норма доходности, ВНД = 73,35 %; индекс доходности проекта, ИД = 3,36; срок окупаемости – не более трех лет.

Данные результаты рассматриваемого пилотного проекта могут повлиять на существенное повышение количества сэкономленной электроэнергии в молочном животноводстве и снижение энергоемкости конечного продукта (молока).

3. В качестве стимулирующих инструментов государственной политики, позволяющих расширять использование частотно-регулируемых электроприводов в животноводстве отечественного АПК могут быть:

– меры по созданию рациональной рыночной среды (включая согласованное тарифное, налоговое, таможенное, антимонопольное регулирование);

– введение системы перспективных технических регламентов, национальных стандартов и норм, стимулирующих энергосбережение;

– стимулирование и поддержку стратегических инициатив хозяйствующих субъектов в инновационной и энергосберегающей сферах.

Список литературы

[1] Государственная программа «Энергосбережение» на 2021-2025 годы, 2021 (в редакции Постановления СМ РБ от 24.02.2021 №103) [Электронный ресурс] – Режим доступа: <http://gosstandart.gov.by/approved-state-program-energy-saving-for-2021-2025-years> – Дата доступа: 11.02.2023.

[2] Государственная программа «Аграрный бизнес» на 2021–2025 годы (в редакции Постановление Совета Министров Республики Беларусь от 1 февраля 2021 года № 59) [Электронный ресурс] – Режим доступа: <https://mshp.gov.by/programms/b81ab6f86bc5670a.html> – Дата доступа: 11.02.2023.

[3] Ловкис, В.Б. О критериях энергетической эффективности сельскохозяйственных технологий / В.Б. Ловкис, В.А. Колос // Механизация и электрификация сельского хозяйства: Межвед. темат. сб. Т. 32. – Минск: РУП «НПЦ НАНБ по механизации с.-х.», 2008. – С. 13-19.

[4] Твердохлеб, Г.В. Технология молока и молочных продуктов / Г.В. Твердохлеб, Г.Ю. Сажинов, Р.И. Раманаскас – Москва: ДеЛи принт, 2006. – 616 с.

EVALUATION OF THE EFFICIENCY OF AUTOMATION INSTALLATION FOR MILK COOLING USING FREQUENCY CONTROL

I.A.Oganezov

*Candidate of Technical Sciences,
Associate Professor
Associate Professor, Department
of Economics and Organization
of Enterprises, Belarusian State
Agrarian Technical University*

N.V.Shcherbina

*Master of Engineering
Senior Lecturer,
Department of Engineering
Psychology and Ergonomics,
Belarusian State University of
Informatics and Radioelectronics*

A.V.Buga

*Candidate of Economic Sciences,
Associate Professor
Associate Professor of the
Department of Economics of the
North-West Institute of
Management of the Russian
Academy of National Economy
and Public Administration under
the President of the Russian
Federation*

Belarusian State Agrarian Technical University, Minsk, Republic of Belarus

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

*North-West Institute of Management of the Russian Academy of National Economy and Public Administration
under the President, Moscow, Russia*

E-mail: iaoganezov@bsuir.by; shcherbina@bsuir.by

Abstract. Economically justified is the prospect of cooling milk on domestic dairy farms built according to standard designs using automation systems based on frequency-controlled electric drives. The proposed technical solutions can provide a significant reduction in energy consumption and labor costs for service personnel while improving the quality of chilled milk.

Keywords: farm, milk, cooling, automation, frequency control, savings, efficiency.

УДК:621.762.2

МОДЕЛИРОВАНИЕ ВЛИЯНИЯ ХАРАКТЕРИСТИК ФИЛЬТРУЕМОЙ СРЕДЫ НА ЗАКОНОМЕРНОСТИ ТАНГЕНЦИАЛЬНОЙ ФИЛЬТРАЦИИ В ПОРИСТЫХ ГРАДИЕНТНЫХ МАТЕРИАЛАХ



М.В.Тумилович

Начальник управления подготовки научных кадров высшей квалификации БГУИР, доктор технических наук, доцент
tumilovich@bsuir.by



Л.П.Пилиневич

Профессор кафедры инженерной психологии и эргономики БГУИР, доктор технических наук, профессор, кавалер медали Франциска Скорины



А.Г.Кравцов

Заместитель академика-секретаря физико-технического отделения наук Национальной академии наук Беларуси, доктор технических наук, профессор

М.В.Тумилович

Начальник управления подготовки научных кадров высшей квалификации БГУИР, доктор технических наук, доцент.

Л.П.Пилиневич

Профессор кафедры инженерной психологии и эргономики БГУИР, доктор технических наук, профессор, кавалер медали Франциска Скорины.

А.Г.Кравцов

Заместитель академика-секретаря физико-технического отделения наук Национальной академии наук Беларуси, доктор технических наук, профессор.

Аннотация. Проведено моделирование влияния характеристик фильтруемой среды на процессы образования и смыва осадка при тангенциальной фильтрации в пористых градиентных материалах. Показано, что существует принципиальная возможность предотвратить процесс забивки (кольматации) пор фильтроэлементов частицами загрязнителя путем соответствующего выбора скорости тангенциального потока и давления фильтрации. При этом следует соблюдать главное требование к структуре фильтроэлементов – она должна обеспечивать режим поверхностной фильтрации.

Ключевые слова: пористые градиентные материалы, тангенциальная фильтрация, поверхностная фильтрация, образование и смыв осадка, кольматация, фильтрующие элементы.

Введение.

В работах [1] и [2] проведено моделирование процессов тангенциальной фильтрации, при которой образующийся осадок постоянно смывается с поверхности фильтрующего элемента (ФЭ) либо собственно потоком суспензии вдоль этой поверхности, либо под воздействием при определенных условиях других факторов (например, центробежных сил). Показано, что свойствами образующегося осадка (толщиной, гранулометрическим составом и гидравлическим сопротивлением) можно управлять посредством изменения параметров фильтрации (скорость суспензии, скорость фильтрата, перепад давления) и конструктивных параметров ФЭ.

При моделировании процесса тангенциальной фильтрации определяющее влияние имеют два фактора: сдвиговые напряжения вдоль поверхности ФЭ t_w и падение (перепад) давления на этом элементе P . Соотношение этих величин используется в качестве критерия блокировки или кольматации пор и названо характеристическим числом Q_0 [3].

$$Q_0 = \frac{P}{t_w} = f \frac{d_c}{\xi_n}, \quad (1)$$

где: d_c – размер частиц;
 ξ_n – размер пор ФЭ.

Число Q_0 и входящие в него величины P и t_w , однозначно определяют некоторый критический размер частицы d' , находящейся в стабильном пристеночном положении: частицы размером больше d' не проникают в ФЭ и уносятся тангенциальным потоком. Путем несложных преобразований можно получить модифицированную формулу для расчета Q_0 по значению перепада давления на ФЭ и скорости тангенциального потока. Кроме того, в формулу входят величины плотности жидкости и коэффициента трения [3].

Для оценки влияния тангенциального потока на сопротивление ФЭ были изготовлены образцы ФЭ из градиентного композиционного пористого материала со слоистой структурой: селективный слой выполнен из порошка меди марки ПМС-Н с размером частиц 0,063–0,1 мм, основа – из волокна меди диаметром 50 мкм. Данные фильтроэлементы имели средний размер пор селективного слоя 28–35 мкм. В дальнейшем они обозначены, как ФЭ–1. Данные ФЭ получали методами порошковой металлургии путем послойного формования, прессования и спекания в защитной атмосфере. Фотография поперечного сечения такого градиентного композиционного пористого материала приведена на рисунке 1.

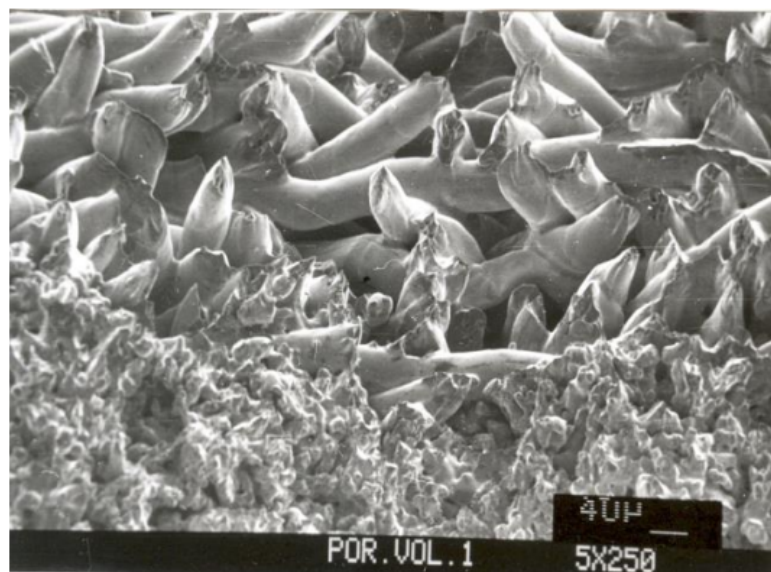
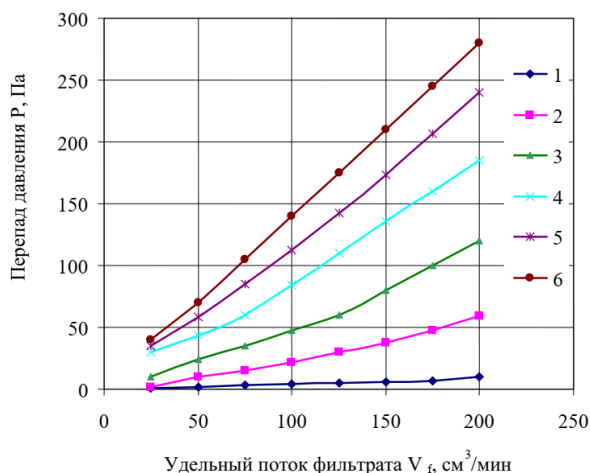


Рисунок 1. Фрактограмма градиентного композиционного пористого материала, полученного из медного порошка и волокна, увеличение $\times 250$

Экспериментальные исследования проводились на дистиллированной воде в диапазоне скоростей тангенциального потока $V = 0,5–10,5$ м/с на установке, описанной в [4]. Результаты исследований представлены на рисунке 1 в виде зависимости перепада давления в центре ФЭ–1 (на расстоянии 150 мм от входа) от удельного потока фильтрата, который искусственно дросселировался при постоянном давлении 400 Па.



1 – $V=0,5$ м/с; 2 – $V=1,39$ м/с;
 3 – $V=3,78$ м/с; 4 – $V=6,03$ м/с; 5 – $V=8,33$ м/с; 6 – $V=10,5$ м/с

Рисунок 2. Зависимость перепада давления на ФЭ – 1 от удельного потока фильтрата V_f при различных скоростях тангенциального потока V

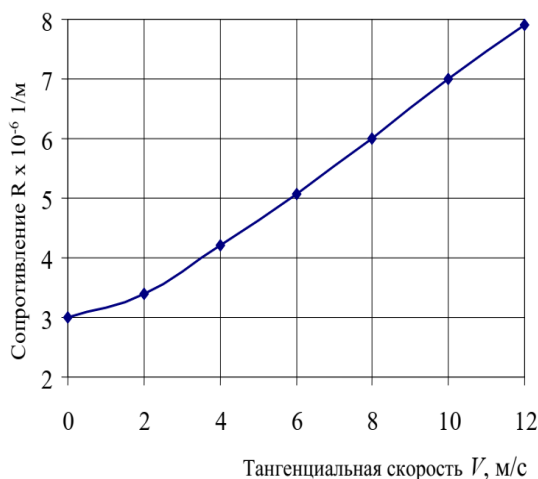


Рисунок 3. Зависимость сопротивления ФЭ – 1 от скорости тангенциального потока

Из зависимостей, приведенных на рисунке 2 видно, что при постоянных скоростях тангенциального потока перепад давления на ФЭ – 1 прямо пропорционален потоку фильтрата. Это, в свою очередь, позволило рассчитать сопротивление потоку R по формуле:

$$R = \frac{P}{V_f} \cdot \eta \quad (2)$$

где η – динамическая вязкость жидкости.

Зависимость сопротивления ФЭ – 1 от тангенциальной скорости потока дистиллированной воды представлена на рисунке 3. Данная зависимость имеет принципиально важное значение. Физический смысл зависимости сопротивления ФЭ от скорости

тангенциального потока пока ясен не до конца. Можно предположить, что причины данного феномена следует изучать в точном рассмотрении соотношения давлений в приближении к поверхности ФЭ.

В работе [3] получена зависимость безразмерной скорости $V^*=V/V_1$ от безразмерной координаты $G=x/L_f$ для модели канала тангенциального фильтра, изображенной на рисунке 4, дающая представление о характере изменения скорости потока по длине канала, часть стенки которого выполнена проницаемой с длиной L_f .

По аналогии с вышеуказанной зависимостью проведено моделирование изменения скорости потока по длине канала для граничных условий ячейки модуля тангенциального фильтра, описанного в работе [4]. Результаты расчетов представлены на рисунках 5 и 6.

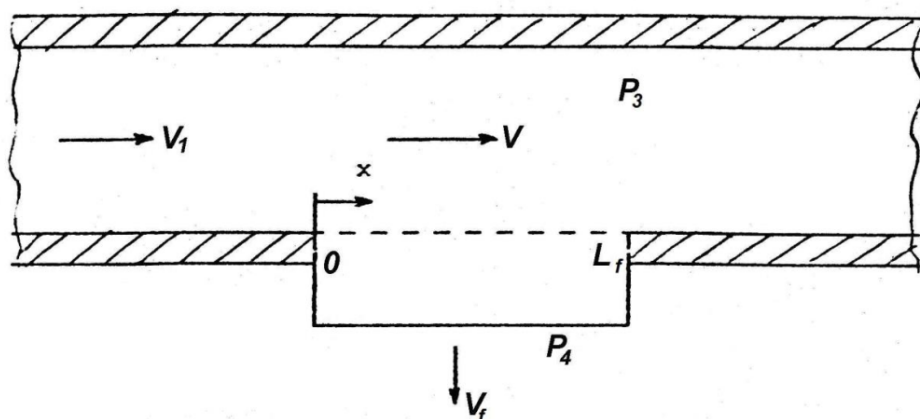
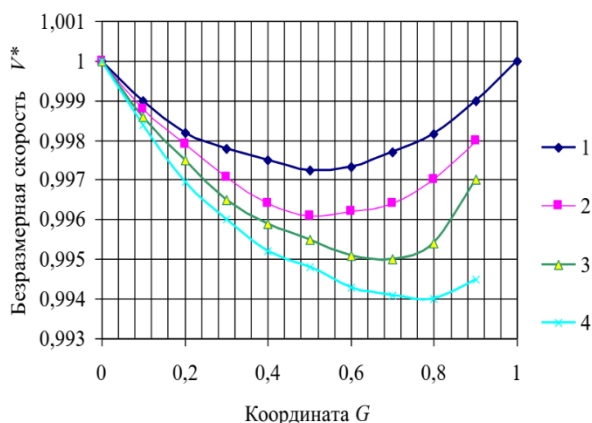


Рисунок 4. Схема модели канала тангенциального фильтра

На рисунке 5 показана зависимость $V^*=f(G)$ при разных скоростях тангенциального потока и нулевой скорости отвода фильтра V_f (при закрытом вентиле), а на рисунке 6 – при фиксированной тангенциальной скорости 8,33 м/с и различных скоростях отвода фильтра.

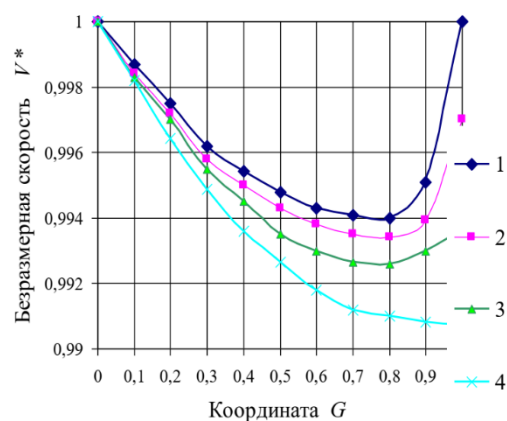
Как видно из рисунка 5, в начале ФЭ безразмерная скорость V^* уменьшается, а затем вновь возрастает до начального значения к концу ФЭ. Хотя это изменение скорости не превышает 1 %, оно увеличивается с ростом скорости тангенциального потока и ростом скорости отвода фильтра (рисунок 6).

На рисунках 7 и 8 показано изменение перепада давления P_3-P_4 по длине фильтроэлемента ФЭ–1 от безразмерной координаты при нулевой скорости отвода фильтра V_f (при закрытом вентиле) и при фиксированной тангенциальной скорости $V_1=6,03$ м/с и различных скоростях отвода фильтра. При этом на рисунке 7 показано изменение перепада давления по длине фильтроэлемента тогда, когда фильтрат течет через фильтроэлемент назад в канал фильтра. Анализ характера изменения зависимостей, приведенных на рисунках 7 и 8, показывает, что в первой рабочей части ФЭ жидкость проникает через пористый материал в пространство для фильтра, а во второй периферийной части вытекает назад в канал. Соответственно этому перепад давления P_3-P_4 в этой области становится отрицательным или положительным (рисунок 7).



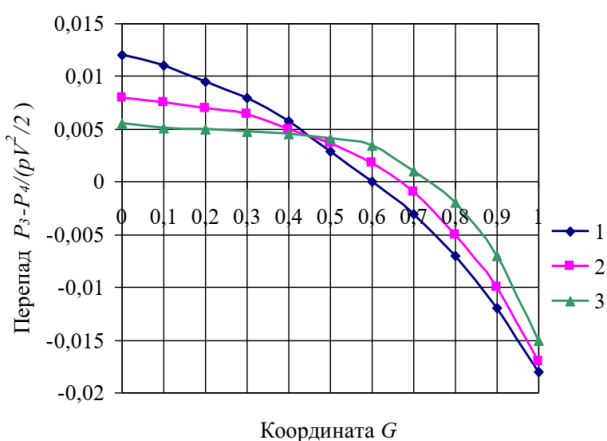
1 – $V_t = 1,39$ м/с; 2 – $V_t = 3,78$ м/с;
3 – $V_t = 6,03$ м/с; 4 – $V_t = 10,5$ м/с

Рисунок 5. Зависимость безразмерной скорости потока V^* от безразмерной координаты G



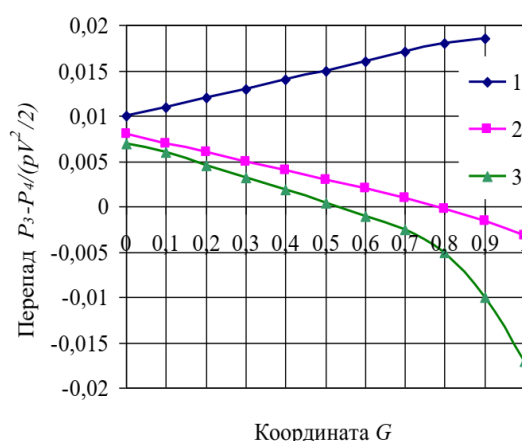
1 – $V_f = 0,008$ м/с; 2 – $V_f = 0,04$ м/с;
3 – $V_f = 0,08$ м/с; 4 – $V_f = 0,12$ м/с

Рисунок 6. Зависимость безразмерной скорости потока V^* от безразмерной координаты G при фиксированной тангенциальной скорости $V_t = 8,33$ м/с



1 – $V_t = 1,39$ м/с; 2 – $V_t = 6,03$ м/с;
3 – $V_t = 10,5$ м/с

Рисунок 7. Изменение перепада давления по длине фильтроэлемента ФЭ – 1 при $V_f = 0$



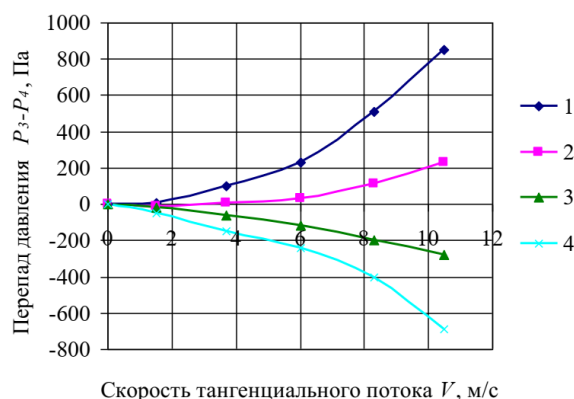
1 – $V_f = 0,16$ м/с; 2 – $V_f = 0,08$ м/с;
3 – $V_f = 0,0016$ м/с

Рисунок 8. Изменение перепада давления по длине фильтроэлемента ФЭ – 1 при $V_t = 6,03$ м/с

Попробуем объяснить этот феномен из предположения, что при $V_f = 0 - V > 0$, а координата точки перехода от «плюса» к «минусу» $G > 0,5$. Для очень высоких значений потоков фильтрата вышеуказанный перепад давлений во всей области положительный и повышается вдоль поверхности фильтра по экспоненте. При небольших же скоростях фильтрации всегда существует еще и область, в которой фильтрат течет через ФЭ в обратном направлении.

Данный эффект подтверждает и эксперимент по составлению зависимостей изменения перепада давления $P_3 - P_4$ от скоростей тангенциального потока при перемещении точки измерения давления P_3 вдоль поверхности ФЭ – 1 (рисунок 9). «Перемещение» точки измерения давления P_3 на практике осуществляли путем его контроля манометрами, установленными последовательно в разных зонах канала ячейки тангенциального микрофильтра. Результаты изменений еще раз показали, что в области значений $G < 0,5$ при скорости отвода фильтрата, равной нулю, перепад давления на фильтре – положительный, а при $G > 0,5$ отрицательный,

причем с ростом тангенциальной скорости потока эта разность растет. Данный эффект имеет важное практическое значение – с его помощью можно легко и просто регенерировать ФЭ обратным потоком фильтрата, увеличивая скорость циркуляции.



1 – $G=0,1$; 2 – $G=0,3$; 3 – $G=0,7$; 4 – $G=0,9$

Рисунок 9. Зависимость перепада давления $P_3 - P_4$ от скорости тангенциального потока V при различных значениях безразмерной координаты G

Таким образом, проведенное моделирование процессов тангенциальной фильтрации в пористых материалах, а также проведенные экспериментальные исследования процессов фильтрации позволили установить, что при определенных условиях можно предотвратить блокировку пор пористого материала твердыми частицами и обеспечить регенерацию ФЭ потоком фильтрата.

Заключение.

На градиентном композиционном пористом материале проведено моделирование влияния характеристик фильтруемой среды на процессы образования и смыва осадка при тангенциальной фильтрации. Показано, что существует принципиальная возможность предотвратить или, по крайней мере, замедлить процесс забивки (кольматации) пор фильтроэлементов частицами загрязнителя путем соответствующего выбора скорости тангенциального потока и давления фильтрации. При этом следует соблюдать главное требование к структуре фильтроэлементов – она должна обеспечивать режим поверхностной фильтрации.

Список литературы

- [1] Тумилович, М. В. Моделирование процессов тангенциальной фильтрации в пористых материалах/ М. В.Тумилович, Л. П.Пилиневич, А. Г.Кравцов // Сб. материалов VII Междунар. науч.-практ. конф. “Big Data and Advanced Analytics”(Минск, Республика Беларусь, 19–20 мая 2021 г.): редкол. :В. А. Богуш [и др.]– Минск: Бестпринт, 2021. – С.305–310.
- [2] Тумилович, М. В. Моделирование процесса образования и смыва осадка при тангенциальной фильтрации в пористых материалах / М. В. Тумилович, Л. П. Пилиневич, А. Г. Кравцов // Сб. материалов VIII Междунар. науч.-практ. конф. “Big Data and Advanced Analytics”(Минск, Республика Беларусь, 11–12 мая 2022 г.): редкол. : В. А. Богуш [и др.]– Минск: Бестпринт, 2022. – С.110–117.
- [3] Fischer; E. Untersuchungen zum Trennprozess bei der Querstromfiltration/ E. Fischer. – Duesseldorf: VDI-Verlag GmbH. – 1987. – 134 S.
- [4] Тумилович, М.В. Очистка жидкостей тангенциальной фильтрацией в пористых порошковых материалах /М.В. Тумилович, В.В. Мазюк, Л.П. Пилиневич, В.В. Савич// Порошковая металлургия. – Минск: Белорусская наука. –2008. – Вып.31. – С. 239–245.

MODELING OF THE EFFECT OF THE CHARACTERISTICS OF A FILTERED MEDIUM ON THE REGULARITIES OF TANGENTIAL FILTRATION IN POROUS GRADIENT MATERIALS

M.V. Tumilovich

Head of the Department for the Training of Scientific Personnel of Higher Qualification of BSUIR, Doctor of Technical Sciences, Associate Professor

L.P. Pilinevich

Professor of Engineering Psychology and Ergonomics BSUIR, Doctor of Technical Sciences, Professor, holder of the Francis Skaryna Medal

A.G. Kravtsov

Deputy Academician-Secretary of the Physical and Technical Department of Sciences of the Presidium of the National Academy of Sciences of Belarus, Doctor of Technical Sciences, Professor

*Belarusian State University of Informatics and Radioelectronics, Minsk, Belarus
E-mail: tumilovich@bsuir.by*

Abstract. The influence of the characteristics of the filtered medium on the processes of sediment formation and washout during tangential filtration in porous gradient materials was modeled. It is shown that there is a fundamental possibility to prevent the process of clogging (clotting) of the pores of the filter elements with pollutant particles by appropriately choosing the tangential flow rate and filtration pressure. At the same time, the main requirement for the structure of filter elements should be observed - it should provide a surface filtration mode.

Keywords: porous gradient materials, tangential filtration, surface filtration, sediment formation and washout, clogging, filter elements.

УДК [659.123]

НАПРАВЛЕНИЯ ТРАНСФОРМАЦИИ WEB-КОПИРАЙТИНГА В УСЛОВИЯХ РАЗВИТИЯ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА



И.В. Марахина
Доцент кафедры экономики БГУИР,
кандидат экономических наук, доцент
marahina@bsuir.by

И.В. Марахина

Окончила Белорусский государственный университет информатики и радиоэлектроники. Область научных интересов связана с инновационным развитием, теорией организации, маркетингом. Доцент кафедры экономики БГУИР, кандидат экономических наук, доцент

Аннотация. Одним из направлений, связанным с развитием сайтов и их продвижением, является написание статей в соответствие с поставленными заданиями – web-копирайтинг. Автор выделяет и обосновывает требования, предъявляемые к таким статьям. Также проводится анализ возможностей повышения производительности труда копирайтера за счет автоматизации. В статье обращается внимание на активизацию возможностей для этого в условиях совершенствования технологий искусственного интеллекта, а также приводится авторское видение дальнейшего развития копирайтинга.

Ключевые слова: копирайтинг, web-копирайтинг, сайт, искусственный интеллект, копирайтер, поисковая выдача.

Введение.

С развитием цифровой экономики все большее число организаций создает и развивает свой сайт. При этом одной из важных составляющих, связанных с такой деятельностью, наряду с программной частью и дизайном, является его информационное наполнение.

Основная часть. Написание статей для сайта в соответствии с поставленными заданиями получило название web-копирайтинг (далее – копирайтинг). Такие статьи могут иметь рекламный, пиар или информационный характер и направлены на достижение маркетинговых задач, в первую очередь продвижение продукта, компании, сайта [1; 2]. Они должны отвечать ряду требований, содержащихся в заданиях. Такие требования к статьям могут быть объединены в следующие группы.

1. Соответствие содержания статьи ожиданиям читателей – посетителей сайта. Статья должна понравиться и удовлетворить их потребности – в т.ч. содержать искомую посетителем информацию. Если статья удерживает внимание и ее дочитают до конца, это с одной стороны, увеличит время нахождения на сайте, что благоприятно отразится на выдаче сайта в поисковых системах. С другой стороны, при выполнении этого требования автору статьи и сайту предоставляются максимальные по объему и времени возможности осуществления воздействия на читателя, например, рекламного. Еще одним плюсом является формирование благоприятного имиджа сайта, например, как экспертного ресурса и лояльности читателей.

2. Соответствовать цели выпуска статьи, включая донесение рекламного и пиар посыла до целевой аудитории, формирование у нее лояльности, продвижение сайта. Очень важно в копирайтинге установить акценты, позволяющие запомнить ключевые послы.

3. Привлечь внимание к статье, не потеряться в информационном шуме. В настоящее время для посетителей характерна «информационная слепота», когда значительная часть

информации, особенно рекламного характера, игнорируется посетителем сайта. Правильный заголовок, грамотно составленный текст формирует «информационную воронку» и побуждает интернет-пользователя кликнуть на заголовок, дочитать до конца и проникнуться идеей статьи.

4. Соответствовать требованиям поисковой системы. Как правило, критически важным является попадание на первые позиции в выдаче поисковых систем. Так, в настоящее время для многих сайтов значительная часть трафика генерируется поисковыми системами, и чем выше их положение в поисковой выдаче, тем больше посетителей откроют статью. Этот пункт связан и с вышеназванными требованиями, так как поисковые системы анализируют длительность просмотра и количество отказов. Важным является качество и уникальность текста, а также графических и видеоматериалов, использованных для иллюстрации. Но кроме указанных характеристик статьи необходим анализ поисковых запросов, трендов, и формирования на их основе требуемой структуры, плана статьи, включения в нее ключевых слов, при сохранении естественного вида.

5. Соответствовать ожиданиям целевой аудитории сайта – быть ей интересной, грамотной и стилистически верной. Соответствовать концепции сайта, зачастую доносить экспертное мнение.

Вышеназванные требования наиболее часто встречаются, но в зависимости о цели могут выдвигаться в разных комбинациях и дополняться более специфическими.

Таким образом копирайтинг требует от специалиста широкого перечня компетенций. Но когда нужно высокое качество материала, появляется необходимость привлечения к работе дополнительных специалистов, например, редактора, экспертов, фотографа, технических специалистов.

С развитием цифровой экономики все чаще сайты отказываются от концепции статичных страниц в пользу динамичного, обновляемого и дополняемого содержания. Актуальный контент определяет их успех в поисковой выдаче и росте посещений сайта. В тоже время такой подход требует и значительного увеличения объема работы копирайтера – а значит роста затрат на оплату его труда.

Избежание значительного удорожания работы копирайтера возможно за счет повышения производительности его труда путем автоматизации части рутинных операций. Как отмечает Квят А. Г., «процедура копирайтинга, как правило, складывается из двух стандартных шагов – постановки маркетинговых задач и непосредственно написания текста» [3, с. 82]. Так, копирайтер уже сейчас активно использует программы по анализу поисковых запросов, выбору ключевых слов, исследованию конкурентов, тем самым ускоряя и совершенствуя часть операций первого шага копирайтинга. В рамках второго этапа копирайтинга – написания текста – традиционно использовались программы по проверке готовой статьи на ошибки, антиплагиат, анализ заспамленности, подбора синонимов.

Дальнейшая автоматизации части рутинных операций копирайтера возможна за счет использования искусственного интеллекта. Согласно некоторым мнениям он может заменить работу копирайтера. Уже сейчас декларируемые возможности искусственного интеллекта в рамках копирайтинга значительны и охватывают оба шага копирайтинга: формирование структуры, выбор ключевых слов, подбор оптимального заголовка и написание контекстной части, рерайтинг, выбор различных целей и задач. В тоже время при практической реализации и использовании существующих программ появляется ряд вопросов, связанных с нарушением авторских прав, истинностью информации, содержанием и качеством текста. Кроме того, конечное решение о выборе концепции и подходящего варианта, стиля задает копирайтер. Тем не менее, следует отметить что скорее всего возможности искусственного интеллекта к написанию статей будут развиваться, а вместе с этим и совершенствоваться получаемый результат.

В будущем при значительном улучшении качества текста, генерируемого искусственным интеллектом, следует ожидать что основные задачи копирайтинга сместятся от написания текста

в сферу управления двумя шагами копирайтинга – планирования, анализа, постановки задач для программ, корректировки и доработки полученных результатов. Также ожидается, что работа копирайтера будет требовать все больше технических и маркетинговых навыков. Еще одним важным направлением копирайтинга будет создание новых и уникальных текстов, экспертных публикаций.

Заключение.

Подводя итоги, можно отметить, что появление искусственного интеллекта в будущем позволит сократить повысить производительность написания статей, удешевить цену условной статьи, при росте их количества и объема и потребует от копирайтера дальнейшего совершенствования и развития технических, экспертных и управленческих компетенций.

Список литературы

- [1] Болдина, К. А. Копирайтинговые и рекламный тексты: к вопросу о тождественности понятий / К. А. Болдина // Проблемы современной науки и образования. – 2016. – № 5 (47). – С. 164–167.
- [2] Жилина, А. В. К вопросу о понятии копирайтинг и его видах / А. В. Жилина // Медиаскоп. – 2020. – Вып. 2. [Электронный ресурс]. URL: <http://www.mediascope.ru/2627> (Дата обращения: 1.03.2023).
- [3] Квят, А. Г. Лингвокогнитивная методика копирайтинга / А. Г. Квят // Альманах теоретических и прикладных исследований рекламы. – 2010. – №1 (1). – С. 79–97.

DIRECTIONS OF TRANSFORMATION OF WEB COPYWRITING UNDER THE DEVELOPMENT OF ARTIFICIAL INTELLIGENCE

I.V. Marahina

*Associate professor, Department of Economics of BSUIR,
Candidate of Economic Sciences, Associate Professor*

Department of Economics

Faculty of Engineering and Economics

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

E-mail: marahina@bsuir.by

Abstract. One of the areas related to the development of sites and their promotion is the writing of articles in accordance with the assigned tasks - web-copywriting. The author highlights and substantiates the requirements for such articles. It also analyzes the possibilities of increasing the productivity of a copywriter through automation. The article draws attention to the activation of opportunities for this in the conditions of improving artificial intelligence technologies, and also provides the author's vision for the further development of copywriting.

Keywords: copywriting, web copywriting, site, artificial intelligence, copywriter, search results.

УДК 159.9.072+159.9.078+331.442+331.446.4

МАТЕМАТИЧЕСКАЯ МОДЕЛЬ ПРОГНОЗИРОВАНИЯ ПОКАЗАТЕЛЕЙ РЕЛАКСАЦИИ РАБОТНИКОВ ЛОКОМОТИВНЫХ БРИГАД



Н. В. Щербина

*Старший преподаватель кафедры инженерной психологии и эргономики
УО «Белорусский государственный университет информатики и
радиоэлектроники», магистр технических наук
shcherbina@bsuir.by*

Н. В. Щербина

Старший преподаватель кафедры инженерной психологии и эргономики УО «Белорусский государственный университет информатики и радиоэлектроники», магистр технических наук, «исследователь» в области технических наук. Проводит научные исследования в областях промышленной безопасности, эргономики, психологии и безопасности труда.

Аннотация. Показано влияние психофизиологических, личностных показателей работников локомотивных бригад на способность к выработке навыка релаксации. Получена факторная структура психофизиологических, личностных показателей работников локомотивных бригад и степень их влияния на успешную выработку навыка релаксации. Получены линейные дискриминантные функции для предсказания принадлежности работника локомотивной бригады к одной из трех групп по успешной выработке навыка релаксации. Получены регрессионные модели зависимостей показателей выработки навыка релаксации от психофизиологических, личностных показателей машинистов, помощников машинистов локомотивных бригад. В среде Tableau разработан дашборд «Профиль бригады машинистов локомотивного депо».

Ключевые слова: биологическая обратная связь, саморегуляция, выработка навыка релаксации, функциональное состояние, электродермальная активность кожи, корреляционный анализ, кластерный анализ, факторный анализ, дискриминантный анализ, множественный регрессионный анализ, локомотивная бригада, психофизиологические и личностные показатели.

Введение.

Актуальность исследования продиктована ростом повышения нагрузки на интеллектуальную и эмоциональную деятельности машинистов, помощников машинистов локомотивных бригад. Профессия машиниста локомотива является одной из наиболее ответственных работ на железнодорожном транспорте. Отличительной чертой данной профессии является постоянный мониторинг ситуации в пути следования, мониторинг показаний приборов, узлов, механизмов всего локомотива. А это значит, что машинист, помощник машиниста локомотива должен обладать набором навыков и умений, таких как ответственность, пунктуальность, эмоциональная устойчивость, монотонноустойчивость, стрессоустойчивость, повышенное внимание, бдительность, скорость и точность двигательных реакций, чувство времени, готовность к экстренному действию в условиях монотонно действующих факторов, максимально концентрироваться на выполнении поставленной задачи и прочее. Повышенная тревожность, утомляемость, снижение внимания, увеличение количества ошибок при выполнении трудовой деятельности и принятие решений способствуют развитию профессиональных заболеваний и влияют на трудоспособность работников локомотивной бригады, что способствует повышению риска возникновения аварий [1].

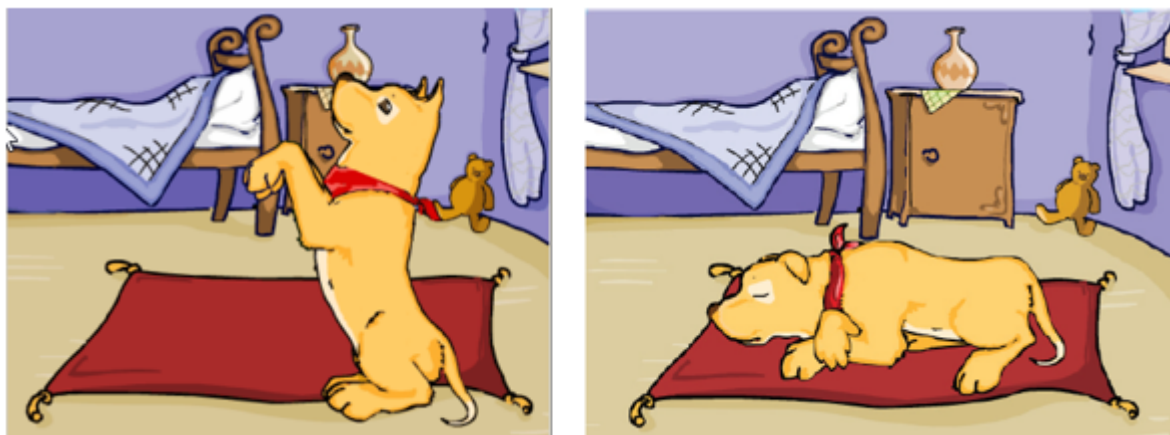
Выработка навыка релаксации позволит работнику локомотивной бригады преодолевать неблагоприятные условия стрессовых (экстремальных) ситуаций в трудовой деятельности,

регулировать свое функциональное состояние на оптимальном уровне во время рабочего рейса и межрейсового отдыха, что благоприятно отразится на надежности выполнения алгоритмов трудовой деятельности и поспособствует сохранению здоровья и продлению трудового долголетия. Прогнозирование способности к выработке навыка релаксации поможет сотрудникам психологической службы осуществлять индивидуальный подход к разработке программ тренинга релаксации и проведению мероприятий по профилактике психологического здоровья.

Материалы и методы.

В исследовании приняли участие 106 машинистов локомотивных бригад Моторвагонного депо г. Минска в возрасте от 18 до 65 лет, мужчины, средний возраст по выборке составил $32 \pm 11,4$ года. Стаж работы участников эксперимента в диапазоне от 1 до 36 полных лет [1, 4].

В качестве модуля обучения релаксации использована система БОС, позволяющая работнику самостоятельно оценивать успешность своего обучения навыкам релаксации, АПК NeuroDog экспериментальный образец 25.11.2013, разработанный АО «Нейроком» г. Москва, Россия. Это биоадаптивная игрушка (БАИ), которая предназначена для обучения пользователя навыкам релаксации путем визуализации его текущего уровня бодрствования на экране монитора. БАИ оценивает текущий уровень бодрствования пользователя путем измерения параметров ЭДА с помощью специального датчика (браслет, перстень). В соответствие текущему уровню бодрствования пользователя один из анимационных фрагментов БАИ выводится на экран монитора. В качестве анимационной картинке на экране монитора пользователь видит виртуального щенка, положение которого соответствует его текущему состоянию (рисунок 1) [1-3].



а)

б)

а – состояние активного бодрствования,

б – состояние полной релаксации (состояние спокойного бодрствования)

Рисунок 1. Визуальное отображение уровня релаксации испытуемого для организации биологической обратной связи

Процесс обучения навыкам релаксации включал в себя сеансы с визуальной биологической обратной связью по параметрам ЭДА с использованием АПК NeuroDog АО «Нейроком». Длительность эксперимента – 10 сеансов по 15 минут. Испытуемому предлагалось «уложить» щенка спать путем достижения собственного состояния спокойного бодрствования. Сеансы проводились индивидуально в присутствии психолога психологической службы Моторвагонного депо г. Минска.

Для оценки профессионально важных психофизиологических, личностных показателей машинистов, помощников машинистов локомотивных бригад использованы основные и дополнительные методики АПК УПДК-МК (АО «Нейроком»).

Результаты.

По результатам БОС-тренинга испытуемые распределены на три группы («успешная» (47 чел.); «менее успешная» (39 чел.); «неуспешная (20 чел.)» по выделенным показателям релаксации.

Введены понятия показатели выработки навыка релаксации: время до КГР-150, сумма КГР-150 за сеанс, время до КГРср-60, сумма КГРср-60 за сеанс, общее время сеанса [1, 4].

В результате проведенного корреляционного анализа по Спирмену установлены достоверные корреляционные положительные и отрицательные взаимосвязи между показателями выработки навыка релаксации и психофизиологическими, личностными показателями машинистов, помощников машинистов локомотивных (рисунок 2). Все связи статистически значимы на уровне $p < 0,05$.

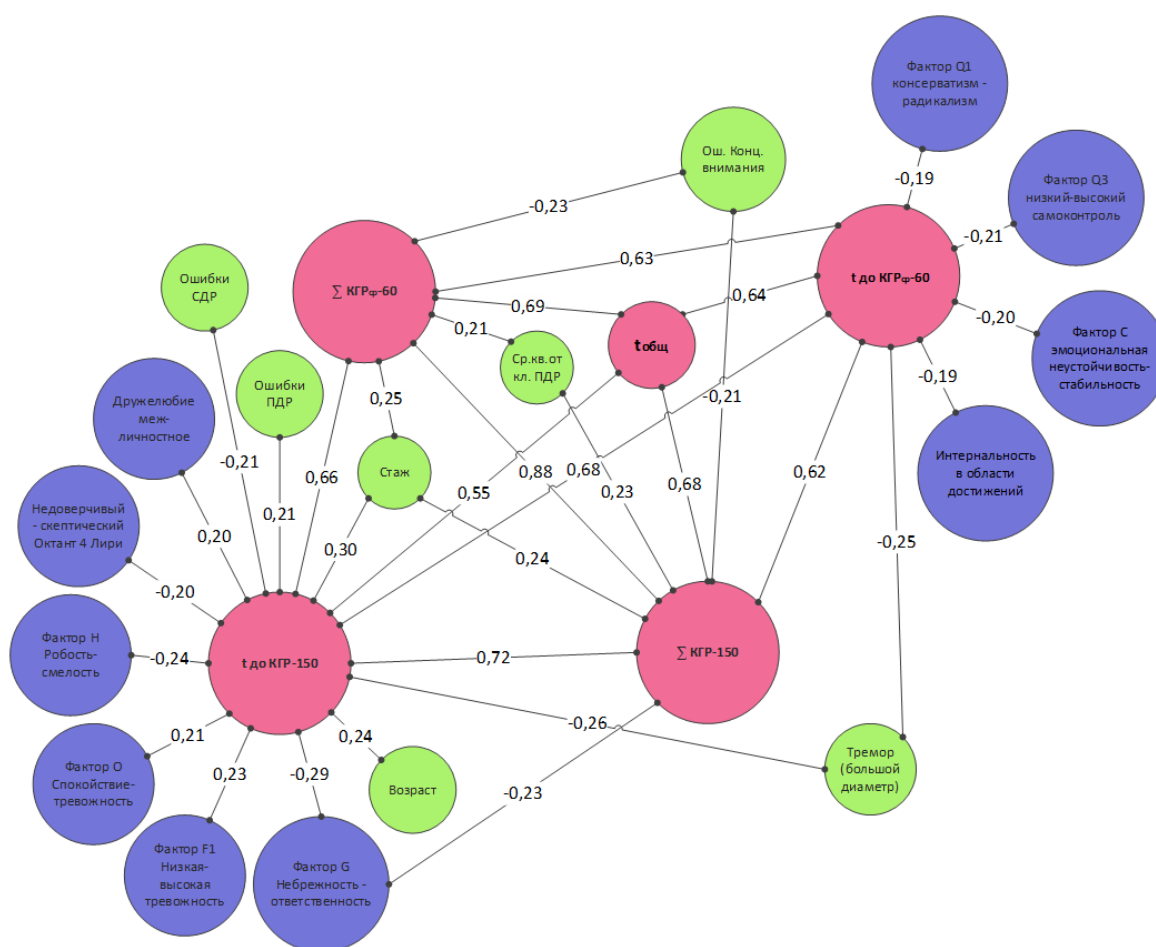


Рисунок 2. Корреляционные плеяды показателей релаксации с психофизиологическими, личностными показателями работников локомотивных бригад

Кластерный анализ с использованием метода k-средних по показателям релаксации машинистов, помощников машинистов локомотивных бригад выявил три кластера испытуемых, различающихся по уровню сформированности навыка выработки релаксации в режиме биоуправления ($\chi^2=65,25$, $p=0,00000$) (рисунок 3).

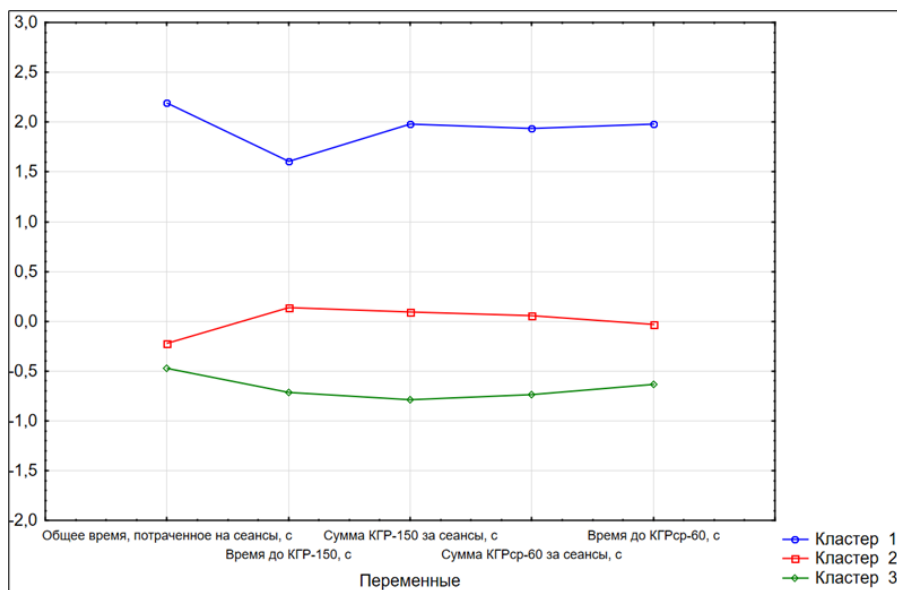


Рисунок 3. График средних значений параметров релаксации в каждом кластере

Кластер 1 на 100 % образован представителями «успешной» группы (14 чел.). Кластер 2 образован представителями «успешной» группы на 60,78 % (31 чел.) и на 39,22 % представителями «менее успешной» группы (20 чел.). В кластере 3 оказалось примерно одинаковое количество представителей «менее успешной» и «неуспешной» групп. Они составили 46,34% (19 чел.) и 48,78% (20 чел.) соответственно от всего числа испытуемых в кластере. Доля в кластере 3 машинистов локомотивных бригад «успешной» группы незначительна – всего 4,88% (2 чел.). Три кластера значительно различаются в показателях: время выполнения задания на эмоциональную устойчивость; количество ошибок, допущенных при выполнении задания на концентрацию внимания; внутриличностный октант 6 (Лири); фактор G (низкая-высокая нормативность поведения) (Кеттелла). Группа «менее успешная» статистически значительно отличается от группы «успешная» по времени выполнения задания на эмоциональную устойчивость, теппинг, внутриличностный октант 6 (Лири). Группа «неуспешная» статистически значительно отличается от групп «успешная» и «менее успешная» по количеству ошибок, допущенных при оценке концентрации внимания. Группа «неуспешная» статистически значительно отличается от группы «менее успешная» по внутриличностному октанту 8 (Лири) и статистически значительно отличается от группы «успешная» по внутриличностному доминированию (Лири) и фактору G [1, 4-8].

Факторному анализу подвергнуты данные по отобранным переменным (психофизиологическим, личностным показателям работников локомотивных бригад) в количестве 64 признаков. Для выделения факторов применен метод главных компонент. Для вращения факторных нагрузок переменных использован метод Варимакс исходных. В результате анализа все переменные объединились в 13 (тринадцать) факторов, объясняющие 60,23 % общей дисперсии переменных. Факторная структура психофизиологических, личностных показателей личности машинистов, помощников машинистов локомотивных бригад, определяющая способность работника к успешной выработке навыка релаксации включает факторы: «Переключение внимания» (собственное значение 6,37) «Стабильность» (собственное значение 4,62), «Эмоциональность» (собственное значение 4,13), «Психоэмоциональное состояние» (собственное значение 3,11), «Скорость реакции в ситуациях принятия решения» (собственное значение 3,09), «Чувство времени» (собственное значение 2,73), «Лидерские качества» (собственное значение 2,56), «Сенситивность и эмоциональная лабильность» (собственное значение 2,35), «Самоконтроль» (собственное значение 2,18),

«Внимание» (собственное значение 2,07), «Интеллекта/мышление» (собственное значение 1,89), «Дружелюбие» (собственное значение 1,82), «Эмоциональная устойчивость» (собственное значение 1,63) [9-11].

Получены линейные дискриминантные функции для групп «успешная», «менее успешная», «неуспешная», которые позволяют предсказать принадлежность машиниста, помощника машиниста локомотивной бригады к одной из трех групп. Полученные стандартизированные коэффициенты позволяют определить частный вклад каждой переменной в предсказание исхода для испытуемого. Полученные данные факторной нагрузки канонических линейных дискриминантных функций могут использоваться для оценки коэффициентов «весомости» переменных при решении диагностических задач [12].

Множественная линейная регрессия позволила определить совокупность психофизиологических, личностных показателей, влияющих на выработку навыка релаксации у машинистов, помощников машинистов локомотивных бригад. Построены пять регрессионных моделей в соответствии с числом показателей релаксации. Модели оказались адекватными (хорошего и достаточного качества на 71 % – 87 % объясняют вариации зависимых переменных [13]), что дает возможность на их основе судить о том, какие именно свойства и каким образом определяют изменения у машинистов, помощников машинистов локомотивных бригад показателей релаксации.

Множественный регрессионный анализ позволяет прогнозировать показатели, характеризующие успешность выработки навыка релаксации, в зависимости от выраженности психофизиологических и личностных показателей машинистов локомотивных бригад. Регрессионное уравнение имеет следующий вид (формула 1):

$$y = \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_n x_n + b + \varepsilon \quad (1)$$

где y – результирующий признак (зависимая, результирующая переменная); n – число факторов, включенных в модель; x_1, x_2, \dots, x_n – признаки-факторы (регрессоры, объясняющие предикторные, предопределенные переменные); $\beta_1, \beta_2, \dots, \beta_n$ – коэффициенты регрессии; b – свободный член регрессии; ε – компонента, отражающая в модели влияние случайных факторов, из-за которых реальное значение показателя может отклониться от теоретического (регрессионный остаток).

Регрессионная модель зависимости «суммы КГР_{ср}-60 за сеансы» от психофизиологических, личностных показателей машинистов, помощников машинистов локомотивных бригад имеет показатели качества: коэффициент детерминации $R^2 = 0,75$ (при $F(49, 56) = 3,51, p = 0,0000043$), статистика Дарбина-Уотсона $d = 2,09$. Качество модели – достаточное [13].

Регрессионная модель зависимости «суммы КГР-150 за сеансы» от психофизиологических, личностных показателей машинистов, помощников машинистов локомотивных бригад имеет показатели качества: коэффициент детерминации $R^2 = 0,71$ (при $F(40, 65) = 4,06, p = 0,00000027$), статистика Дарбина-Уотсона $d = 1,81$. Качество модели – хорошее [13].

Регрессионная модель зависимости «времени до КГР-150» от психофизиологических, личностных показателей машинистов, помощников машинистов локомотивных бригад имеет показатели качества: коэффициент детерминации $R^2 = 0,87$ (при $F(54, 51) = 6,34, p = 0,000000000$), статистика Дарбина-Уотсона $d = 1,96$. Качество модели – хорошее.

Регрессионная модель для зависимой переменной «Время до КГР-150» включает в себя следующие предикторы: объем внимания (количество знаков) ($\beta = -0,65$), ММРІ К (шкала коррекции) ($\beta = 0,47$), недоверчивый тип внутриличностных отношений (октант 4) ($\beta = -0,43$), психологическое утомление (Люшер) ($\beta = 0,43$), независимый тип внутриличностных отношений (октант 2) ($\beta = 0,41$), интернальность общая шкала ($\beta = -0,37$), фактор Н: робость-смелость ($\beta = -$

0,37), количество ошибок в сложной двигательной реакции ($\beta=-0,37$), психологическое напряжение ($\beta=-0,36$), ММРІ F (шкала валидности – достоверность) ($\beta=0,34$), теппинг-3 ($\beta=0,32$), бдительность ($\beta=-0,32$), независимый тип межличностных отношений (октант 2) ($\beta=-0,31$), теппинг-5 ($\beta=0,30$), фактор А: замкнутость-общительность ($\beta=0,30$), ответственный тип межличностных отношений (октант 8) ($\beta=0,29$), среднее время простой двигательной реакции ($\beta=-0,27$), теппинг-1 ($\beta=-0,27$), тремор (малое отверстие) ($\beta=-0,26$), ММРІ 0 – социальная интроверсия ($\beta=-0,25$), фактор М: практичность-мечтательность ($\beta=0,24$), число неправильных нажатий (бдительность) ($\beta=0,23$), ММРІ 3 – эмоциональная лабильность ($\beta=-0,23$), время выполнения задания на эмоциональную устойчивость ($\beta=-0,22$), среднее время реакции на движущийся объект ($\beta=0,22$), количество ошибок (эмоциональная устойчивость) ($\beta=-0,20$), фактор G: низкая-высокая нормативность поведения ($\beta=-0,19$), интернальность в области межличностных отношений ($\beta=-0,19$), суммарное отклонение от аутогенной нормы ($\beta=0,18$), скорость переключения внимания ($\beta=-0,17$). Множественный коэффициент корреляции равен 0,93.

Полученная модель позволяет сделать следующие выводы. Можно ожидать увеличения времени, которое пройдет от начала сеанса до фиксирования первого КГР-150, у машинистов с выраженным психологическим утомлением, демонстрируют более высокий темп рук, характеризуются нарастанием неправильных нажатий при оценке бдительности, увеличением времени реакции на движущийся объект. Они более ответственные, общительны, независимы, практичны, имеют большее суммарное отклонение от аутогенной нормы (т.е. большая неустойчивость).

С высокой вероятностью время до первого КГР-150 сократится при обнаружении у машинистов следующих тенденций в психофизиологических особенностях: нарастание количества правильно воспроизведенных знаков при изучении у них объема внимания; увеличение среднего времени простой двигательной реакции и времени выполнения задания на эмоциональную устойчивость и увеличение скорости переключения внимания; большее количество ошибок при сложной двигательной реакции и оценке эмоциональной устойчивости; увеличение степени психологического напряжения. Сокращение времени до КГР-150 также зависит от того, насколько сильно у машинистов проявляются такие личностные свойства, как бдительности, скептичность, робость, независимость, интровертированность, но при этом достаточно сильное психологическое напряжение; высокая нормативность поведения.

Регрессионное уравнение для прогнозирования временного интервала от начала сеанса до первого импульса КГР длительностью 150 с выглядит следующим образом:

$$\begin{aligned} \text{Время до КГР}_{150} = & 4921,009 - 1,906x_1 + 23,327x_2 - 93,949x_3 + 105,683x_4 + \\ & +122,669x_5 - 134,431x_6 - 130,138x_7 - 114,708x_8 - 101,917x_9 + 15,128x_{10} + 17,097x_{11} - \\ & -5,545x_{12} - 81,665x_{13} + 20,258x_{14} + 99,137x_{15} + 76,962x_{16} - 3,587x_{17} - 14,489x_{18} - \\ & -271,896x_{19} - 15,193x_{20} + 73,535x_{21} + 216,051x_{22} - 16,908x_{23} - 2,878x_{24} + 6,047x_{25} - \\ & -167,630x_{26} - 53,944x_{27} - 70,789x_{28} + 18,656x_{29} - 2,973x_{30} - 11,451x_{31} + 27,278x_{32} - \\ & -32,862x_{33} - 51,478x_{34} + 41,302x_{35} + 71,432x_{36} - 20,303x_{37} + 9,710x_{38} - 9,042x_{39} - \\ & -7,535x_{40} - 1,789x_{41} - 64,186x_{42} + 42,965x_{43} - 34,498x_{44} - 3,794x_{45} - 11,220x_{46} + \\ & +62,833x_{47} + 84,275x_{48} + 71,734x_{49} - 1,006x_{50} + 21,264x_{51} + 49,808x_{52} - 40,813x_{53} - \\ & -35,426x_{54}, \end{aligned} \quad (2)$$

где x_1 – объем внимания (количество знаков), x_2 – ММРІ К (шкала коррекции), x_3 – недоверчивый тип внутриличностных отношений (октант 4), x_4 – психологическое утомление (Люшер), x_5 – независимый тип внутриличностных отношений (октант 2), x_6 – интернальность общая шкала, x_7 – фактор Н: робость-смелость, x_8 – количество ошибок в сложной двигательной реакции, x_9 – психологическое напряжение, x_{10} – ММРІ F (шкала валидности – достоверность), x_{11} – теппинг/3, x_{12} – бдительность, x_{13} – независимый тип межличностных отношений (октант

2), x_{14} – теппинг-5, x_{15} – фактор А: замкнутость-общительность, x_{16} – ответственный тип межличностных отношений (октант 8), x_{17} – среднее время простой двигательной реакции, x_{18} – теппинг/1, x_{19} – тремор (малое отверстие), x_{20} – ММРІ 0 – социальная интроверсия, x_{21} – фактор М: практичность-мечтательность, x_{22} – число неправильных нажатий (бдительность), x_{23} – ММРІ 3 – эмоциональная лабильность, x_{24} – время выполнения задания на эмоциональную устойчивость, x_{25} – среднее время реакции на движущийся объект, x_{26} – количество ошибок (эмоциональная устойчивость), x_{27} – фактор G: низкая-высокая нормативность поведения, x_{28} – интернальность в области межличностных отношений, x_{29} – суммарное отклонение от аутогенной нормы, x_{30} – скорость переключения внимания, x_{31} – ММРІ 9: оптимистичность, x_{32} – F1 (Кеттелла), x_{33} – тремор (большое отверстие), x_{34} – объем внимания (количество символов), x_{35} – концентрация (количество ошибок), x_{36} – интернальность в области неудач, x_{37} – F3 (Кеттелла), x_{38} – ММРІ 4: импульсивность, x_{39} – ММРІ 8: индивидуалистичность, x_{40} – интернальность в отношении здоровья и болезни, x_{41} – среднее квадратическое отклонение времени сложной двигательной реакции, x_{42} – количество ошибок (оценка эмоциональной устойчивости), x_{43} – межличн. октант 6 (Лири), x_{44} – эмоциональный стресс (Люшер), x_{45} – ММРІ L: шкала лжи, x_{46} – ММРІ 1 – невротический сверхконтроль, x_{47} – вегетативный коэффициент (Люшер), x_{48} – интернальность в области достижений, x_{49} – количество ошибок в простой двигательной реакции, x_{50} – среднее квадратическое отклонение времени реакции на движущийся объект, x_{51} – интернальность в семейных отношениях, x_{52} – фактор О: спокойная самоуверенность – склонность к чувству вины, x_{53} – фактор L: доверчивость – подозрительность, x_{54} – внутриличн. октант 6 (Лири).

Регрессионная модель зависимости «времени до КГР_{ср}-60» от психофизиологических, личностных показателей машинистов, помощников машинистов локомотивных бригад имеет показатели качества: коэффициент детерминации $R^2 = 0,84$ (при $F(55, 50) = 4,67$, $p = 0,00000077$), статистика Дарбина-Уотсона $d = 1,76$. Качество модели – хорошее.

Предикторами в данной модели являются: межличностное дружелюбие ($\beta = 0,64$), ММРІ F (шкала валидности – достоверность) ($\beta = 0,53$), фактор А: замкнутость-общительность ($\beta = 0,52$), ММРІ 0 – социальная интроверсия ($\beta = -0,52$), ММРІ К (шкала коррекции) ($\beta = 0,51$), время выполнения третьего задания на переключение внимания ($\beta = -0,48$), ММРІ 3 – эмоциональная лабильность ($\beta = -0,46$), фактор М: практичность-мечтательность ($\beta = 0,43$), психологическое утомление ($\beta = 0,41$), независимый тип межличностных отношений (октант 2) ($\beta = -0,38$), тремор (большое отверстие) ($\beta = -0,36$), внутриличностное дружелюбие ($\beta = -0,35$), фактор С: эмоциональная неустойчивость-стабильность ($\beta = -0,34$), теппинг-1 ($\beta = -0,34$), суммарное отклонение от аутогенной нормы ($\beta = 0,33$), покорный тип межличностных отношений (октант 5) ($\beta = -0,32$), ММРІ L (шкала лжи) ($\beta = -0,31$), ММРІ 9 – оптимистичность ($\beta = -0,31$), интернальность в области неудач ($\beta = 0,30$), теппинг-3 ($\beta = 0,30$), общая интернальность ($\beta = -0,29$), время выполнения второго задания на оценку переключения внимания ($\beta = 0,29$), ответственный тип межличностных отношений (октант 8) ($\beta = 0,26$), фактор L: доверчивость – подозрительность ($\beta = -0,26$), ММРІ 4 – импульсивность ($\beta = 0,24$), объем внимания, количество знаков ($\beta = -0,22$) и символов ($\beta = -0,20$), число пропущенных сигналов с предупреждением ($\beta = 0,19$), среднее арифметическое чувства времени ($\beta = 0,17$). Множественный коэффициент корреляции равен 0,91.

Большее время до первого КГР_{ср}-60 определяется такими психофизиологическими особенностями, как психологическое утомление, увеличение времени выполнения второго задания на переключение внимания, числа пропущенных сигналов с предупреждением, среднего арифметического чувства времени и более высокий темп рук. Больше времени до первого КГР_{ср}-60 проходит у тех машинистов и помощников машинистов, которым свойственна импульсивность, ответственность, высокая интернальность в области неудач.

Время до первого КГР_{ср}-60 сокращается для тех испытуемых, у которых повышен тремор правой руки в большом диаметре и увеличен объем внимания (знаков и символов), уменьшение времени на выполнение третьего задания на переключение внимания. Их отличают такие личностные особенности, как интровертированность, эмоциональная лабильность, оптимистичность, скромность, у которых менее выражено внутриличностное дружелюбие.

Регрессионное уравнение для прогнозирования времени от начала сеанса до первого усредненного импульса КГР_{ср}-60 выглядит следующим образом:

$$\begin{aligned} \text{Время до КГР}_{\text{ср}} 60 = & 4586,410 + 69,489x_1 + 27,992x_2 + 213,850x_3 - 37,974x_4 + \\ & +31,138x_5 - 8,688x_6 - 41,080x_7 + 162,440x_8 + 122,401x_9 - 123,420x_{10} - 107,021x_{11} - \\ & -39,486x_{12} - 128,457x_{13} - 22,344x_{14} + 40,732x_{15} - 82,735x_{16} - 20,215x_{17} - 26,598x_{18} + \\ & +123,192x_{19} + 18,926x_{20} - 126,492x_{21} + 15,688x_{22} + 84,921x_{23} - 93,704x_{24} + 22,218x_{25} - \\ & -0,774x_{26} - 100,406x_{27} + 126,737x_{28} + 0,208x_{29} - 37,140x_{30} - 19,199x_{31} + 6,295x_{32} - \\ & -111,91x_{33} - 45,772x_{34} - 8,180x_{35} - 1,393x_{36} + 1,344x_{37} - 6,853x_{38} - 52,346x_{39} - \\ & -132,175x_{40} + 1,677x_{41} - 22,888x_{42} + 12,757x_{43} + 13,691x_{44} + 126,888x_{45} - 44,044x_{46} - \\ & -6,381x_{47} + 42,966x_{48} - 208,356x_{49} + 67,514x_{50} - 64,786x_{51} + 84,665x_{52} - 58,979x_{53} + \\ & +54,365x_{54} + 41,296x_{55}, \end{aligned} \quad (3)$$

где x_1 – межличностное дружелюбие, x_2 – ММРІ F (шкала валидности – достоверность), x_3 – фактор А: замкнутость-общительность, x_4 – ММРІ 0 – социальная интроверсия, x_5 – ММРІ К (шкала коррекции), x_6 – время выполнения третьего задания на переключение внимания, x_7 – ММРІ 3 – эмоциональная лабильность, x_8 – фактор М: практичность-мечтательность, x_9 – психологическое утомление, x_{10} – независимый тип межличностных отношений (октант 2), x_{11} – тремор (большое отверстие), x_{12} – внутриличностное дружелюбие, x_{13} – фактор С: эмоциональная неустойчивость-стабильность, x_{14} – теппинг-1, x_{15} – суммарное отклонение от аутогенной нормы, x_{16} – покорный тип межличностных отношений (октант 5), x_{17} – ММРІ L (шкала лжи), x_{18} – ММРІ 9 – оптимистичность, x_{19} – интернальность в области неудач, x_{20} – теппинг-3, x_{21} – общая интернальность, x_{22} – время выполнения второго задания на оценку переключения внимания; x_{23} – ответственный тип межличностных отношений (октант 8), x_{24} – фактор L: доверчивость – подозрительность, x_{25} – ММРІ 4 – импульсивность, x_{26} – объем внимания (количество знаков), x_{27} – объем внимания (количество символов), x_{28} – число пропущенных сигналов с предупреждением, x_{29} – среднее арифметическое чувства времени, x_{30} – фактор Q1: консерватизм – радикализм, x_{31} – ММРІ 8 – индивидуалистичность, x_{32} – количество ошибок (оценка эмоциональной устойчивости), x_{33} – число пропущенных сигналов без предупреждения (оценка бдительности), x_{34} – эмоциональный стресс (Люшер), x_{35} – ММРІ 5 – мужественность – женственность, x_{36} – время выполнения первого задания (переключение внимания), x_{37} – ситуативная тревожность, x_{38} – фактор Q3: низкий – высокий самоконтроль, x_{39} – психологическое напряжение (Люшер), x_{40} – количество ошибок (скорость переключения внимания), x_{41} – доминирование внутриличностное (Лири), x_{42} – количество ошибок в сложной двигательной реакции, x_{43} – теппинг/5, x_{44} – ММРІ 7: тревожность, x_{45} – F1 (Кеттелла), x_{46} – тревожность (Люшер), x_{47} – теппинг/ср, x_{48} – межличн. октант 3 (Лири), x_{49} – количество ошибок в простой двигательной реакции, x_{50} – фактор О: спокойная самоуверенность – склонность к чувству вины, x_{51} – межличн. октант 6 (Лири), x_{52} – интернальность в семейных отношениях, x_{53} – фактор I: реализм-чувствительность, x_{54} – фактор Н: робость-смелость, x_{55} – внутриличн. октант 1 (Лири).

В среде Tableau (ver. 2021.4.4.) разработан интерактивный дашборд «Профиль бригады машинистов локомотивного депо» (рисунок 4).

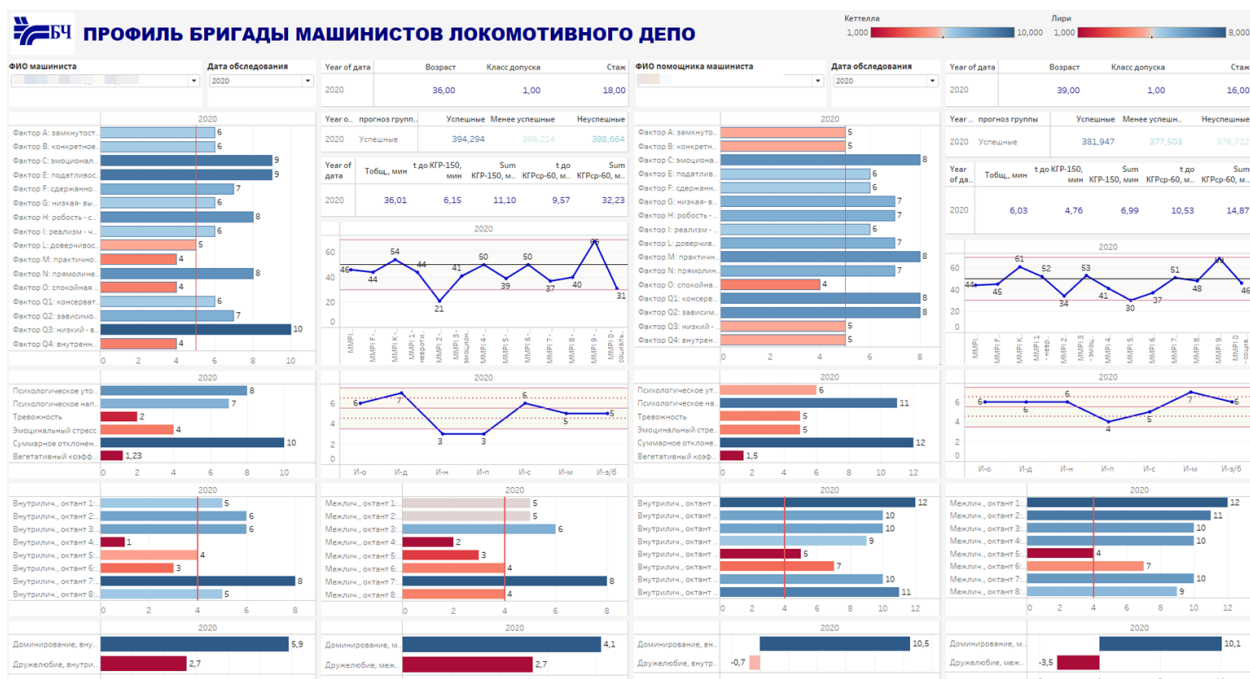


Рисунок 4. Рабочее окно дашборда «Профиль бригады машинистов локомотивного депо»

Профиль представлен следующей информацией о локомотивной бригаде: дату (год) проведения периодического обследования, возраст, класс допуска, стаж, прогнозируемую категорию группы по выработке навыка релаксации (успешная, менее успешная, неуспешная) и прогнозируемые показатели релаксации, т.е. как быстро и как продолжительно машинист (конкретный машинист) будет погружаться в состояние релаксации за весь период обучения и сколько времени он может потратиться для достижения этих результатов. Кроме этих данных на панели дашборда выводятся графические данные опросника Кеттелла, личностного опросника ММРІ, уровень субъективного контроля, данные по психоэмоциональному состоянию и межличностным отношениям. Выбрав в графе «ФИО машиниста/помощника машиниста» фамилию машиниста и его помощника, выбрав дату проведения обследования психолог получает информацию по личностному профилю локомотивной бригады, что позволяет оценить психоэмоциональное состояние машиниста и бригады в целом, а также оценить и подобрать продолжительность обучения навыку релаксации.

Заключение.

Результаты научных исследований подтверждают влияние психофизиологических, личностных показателей работников локомотивных бригад на способность к выработке навыка релаксации.

Понятие «выработка навыка релаксации» представлено следующими показателями: время до КГР-150, сумма КГР-150 за сеанс, время до КГРсп-60, сумма КГРсп-60 за сеанс, общее время сеанса [1, 4].

Получена факторная структура психофизиологических, личностных показателей работников локомотивных бригад и степень их влияния на успешную выработку навыка релаксации [9-11].

Получены линейные дискриминантные функции для групп «успешная», «менее успешная», «неуспешная», которые позволяют предсказать принадлежность работника локомотивной бригады к одной из трех групп. Полученные стандартизированные коэффициенты позволяют определить частный вклад каждой переменной в предсказание исхода для испытуемого, а полученные данные факторной нагрузки канонических линейных

дискриминантных функций могут использоваться для оценки коэффициентов «весомости» переменных при решении диагностических задач [12].

Получены регрессионные модели зависимостей показателей выработки навыка релаксации от психофизиологических, личностных показателей машинистов, помощников машинистов локомотивных бригад. Полученные модели хорошего и достаточного качества на 71 % – 87 % объясняют вариации зависимых переменных [13].

Полученные результаты исследования могут быть использованы для обеспечения психофизиологического сопровождения профессиональной деятельности работников локомотивных бригад; разработки рекомендаций по организации и контролю за режимом труда и отдыха локомотивных бригад; разработки психопрофилактических и психокоррекционных мероприятий, направленных на снижение утомления, переутомления, психоэмоционального напряжения и других нарушений функционального состояния локомотивных бригад.

Список литературы

[1] Щербина, Н. В. Исследование метода выработки навыка на релаксацию с биологической обратной связью по параметрам электродермальной активности / Н.В. Щербина, В.В. Савченко, К.Д. Яшин // Новости медико-биологических наук. – 2019. – № 1/2019. – Том 19. – С. 65-73.

[2] Гедранович Ю. А. Обзор и сравнительный анализ методов и систем для развития навыков релаксации. / Ю. А. Гедранович, В. В. Савченко, К. Д. Яшин, Н. В. Щербина // Журнал «Человеческий фактор: проблемы психологии и эргономики», 2016, № 1 (77), С. 62 – 69. URL: <http://elibrary.ru> (дата обращения 22.09.2016)

[3] Гедранович Ю. А. Обзор и сравнительный анализ методов и систем для развития навыков релаксации. / Ю.А. Гедранович, В.В. Савченко, К. Д. Яшин, Н.В. Щербина // Журнал «Человеческий фактор: проблемы психологии и эргономики», 2016, № 2 (78), С. 44 – 50. URL: <http://elibrary.ru> (дата обращения 22.09.2016).

[4] Щербина, Н.В. Первичная обработка и анализ данных для оценки состояния релаксации машинистов железнодорожного транспорта / Н. В. Щербина, В. В. Савченко, К. Д. Яшин // BIG DATA and Advanced Analytics = BIG DATA и анализ высокого уровня : сборник материалов V Международной научно-практической конференции, Минск, 13–14 марта 2019 г. В 2 ч. Ч. 2 / Белорусский государственный университет информатики и радиоэлектроники; редкол. : В. А. Богуш [и др.]. – Минск, 2019. – С. 121 – 125.

[5] Щербина, Н. В. Анализ различий между группами машинистов локомотивных бригад по профессионально важным психофизиологическим и личностным показателям / Н. В. Щербина // BIG DATA and Advanced Analytics = BIG DATA и анализ высокого уровня : сборник научных статей VIII Международной научно-практической конференции, Минск, 11-12 мая 2022 года / Белорусский государственный университет информатики и радиоэлектроники ; редкол.: В. А. Богуш [и др.]. – Минск, 2022. – С. 265–272.

[6] Щербина Н. В. Выявление индивидуальных особенностей к саморегуляции помощников машинистов и машинистов железнодорожного транспорта / Н. В. Щербина // Компьютерное проектирование и технология производства электронных систем: сборник тезисов 53 научной конференции аспирантов, магистрантов и студентов (Минск, 2–6 мая 2017 года) / отв. ред. Раденюк А. Л. – Минск : БГУИР, 2017. – С. 246.

[7] Щербина, Н. В. Оценка состояния релаксации машинистов железнодорожного транспорта / Н. В. Щербина // Электронные системы и технологии: сборник тезисов докладов 55 юбилейной научной конференции аспирантов, магистрантов и студентов (Минск, 22–26 апреля 2019 г.) / отв. ред. Раденюк А. Л. – Минск, БГУИР, 2019. – С. 607-608.

[8] Щербина Н.В. Анализ различий между группами машинистов локомотивных бригад с разной степенью способности к произвольной регуляции их функционального состояния. / Н.В. Щербина // Доклады БГУИР. 2022;20(2):21-29. <https://dx.doi.org/10.35596/1729-7648-2022-20-2-21-29>.

[9] Щербина Н.В. Регуляция функционального состояния машинистов локомотивных бригад с применением БОС-тренинга: факторный анализ экспериментальных данных. / Н.В. Щербина // Доклады БГУИР. 2021;19(4):28-36. <https://dx.doi.org/10.35596/1729-7648-2021-19-4-28-36>.

[10] Щербина, Н. В. Факторный анализ показателей индивидуальных психофизиологических и личностных характеристик машинистов локомотивных бригад / Щербина Н. В. // Актуальные проблемы практической психологии : материалы Всероссийской научно-практической конференции с международным участием, Тверь, 17 декабря 2021 г. / под ред. Т. А. Попковой, А. В. Антоновского. – Тверь : СФК-офис, 2021. – С. 262–267.

[11] Щербина, Н.В. Визуализация психофизиологических и личностных данных машинистов локомотивных бригад в Tableau / Н.В. Щербина // BIG DATA and Advanced Analytics = BIG DATA и анализ высокого уровня : сборник материалов VIII Международной научно-практической конференции, Минск, 11-12 мая 2022 г. В 2 ч. Ч. 2 / Белорусский государственный университет информатики и радиоэлектроники; редкол. : В. А. Богуш [и др.]. – Минск, 2022. – С. 215 – 218.

[12] Щербина Н.В. Дифференциальная диагностика способности к выработке навыка релаксации у машинистов локомотивных бригад. / Н.В. Щербина // Доклады БГУИР. 2022;20(4):96-103.

<https://dx.doi.org/10.35596/1729-7648-2022-20-4-96-103>.

[13] Щербина Н.В. Прогнозирование способности к выработке навыка релаксации у машинистов локомотивных бригад. / Н.В. Щербина // Доклады БГУИР. 2022;20(7):95-101. <https://dx.doi.org/10.35596/1729-7648-2022-20-7-95-101>.

MATHEMATICAL MODEL FOR PREDICTION OF RELAXATION INDICATORS OF WORKERS OF LOCOMOTIVE BRIGADES

N.V. Shcherbina

Master of Technical Science

Senior Lecturer, Department of Engineering Psychology and Ergonomics, Belarusian State University of Informatics and Radioelectronics

Department of Engineering Psychology and Ergonomics

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

E-mail: shcherbina@bsuir.by

Abstract. The influence of psychophysiological, personal indicators of workers of locomotive crews on the ability to develop the relaxation skill is shown. The factorial structure of psychophysiological, personal indicators of employees of locomotive crews and the degree of their influence on the successful development of the relaxation skill have been obtained. Linear discriminant functions have been obtained to predict whether an employee of a locomotive crew belongs to one of the three groups according to the successful development of a relaxation skill. Regression models of the dependence of the indicators of the development of the skill of relaxation on the psychophysiological, personal indicators of drivers, assistant drivers of locomotive crews have been obtained. In the Tableau environment, the dashboard “Profile of the crew of locomotive depot drivers” was developed.

Keywords: biofeedback, self-regulation, relaxation skill development, functional state, skin electrodermal activity, correlation analysis, cluster analysis, factor analysis, discriminant analysis, multiple regression analysis, locomotive crew, psychophysiological and personal indicators.

УДК 331.101.1:004.774

ИСПОЛЬЗОВАНИЕ БАЗОВЫХ ПРИНЦИПОВ ЭРГОНОМИКИ В WEB-РАЗРАБОТКЕ



И. Ф. Киринович

доцент кафедры инженерной психологии и эргономики БГУИР, к.ф.-м.н, доцент,
Kirinovich.irina@yandex.ru



К. Дадаев

магистрант кафедры инженерной психологии и эргономики БГУИР
frkeodal.rty@gmail.com

И. Ф. Киринович

С 2015 года – доцент кафедры инженерной психологии и эргономики БГУИР.

К. Дадаев

С 2022 года – магистрант кафедры инженерной психологии и эргономики БГУИР.

Аннотация. Рассматриваются принципы эргономики, используемые при разработке веб-приложений, позволяющие создавать удобные и легкие в использовании интерфейсы, доступные для любого пользователя.

Ключевые слова: эргономика, принципы эргономики, web-приложение, элементы интерфейса

Введение.

Актуальный программный продукт должен не только хорошо выполнять свою функцию, но быть понятным и комфортным пользователю. Будь то обычный монитор или компьютерное приложение, оно должно быть сделано с использованием разработок последних десятилетий, направленных на улучшение удобства, скорости взаимодействия, а главное понятности. Это сказывается на эффективности работы персонала, продуктивности, а также качестве услуг. Ответом для решения данной проблемы стало развитие эргономики.

Анализ проблем.

Эргономика – научная и проектировочная дисциплина, сформировавшаяся на стыке психологии, физиологии, гигиены труда, биомеханики, антропологии и ряда технических дисциплин [1].

Эргономика приобретает все большее значение в связи с ростом популярности технологий и количеством часов, которые люди проводят за компьютерами, смартфонами и другими цифровыми устройствами. Важным принципом эргономики при разработке веб-приложений (приложений) является проектирование, ориентированное на пользователя. Дизайн приложений должен быть основан на потребностях и предпочтениях пользователей, а не на предпочтениях разработчика. Это означает, что разработчику необходимо иметь данные о потребностях, предпочтениях и привычках пользователей. Это можно сделать, например, с помощью опросов, интервью и юзабилити-тестирования.

Следующий принцип эргономики заключается в обеспечении четкой и последовательной навигации. Пользователи должны иметь возможность легко перемещаться по приложению и находить нужную информацию. Этого можно достичь с помощью четкой и лаконичной маркировки, организации контента в логические категории и обеспечения последовательных элементов навигации во всем приложении.

Принципом эргономики является минимализм в дизайне. Это означает, что разработчик должен стремиться к уменьшению ненужных элементов в дизайне приложения. Пользовательский интерфейс должен быть простым и интуитивно понятным, с отображением на экране только основных элементов. Это поможет пользователям сосредоточиться на задачах,

которые они пытаются решить, не отвлекаясь на несущественную информацию или визуальный беспорядок. Принцип понятности заключается в том, чтобы пользователи имели возможность легко читать текст в приложении. Этого можно добиться путем имплементирования необходимой информации о функционале в небольшой объёме и доступном виде.

Обеспечение принципа доступности при разработке состоит в том, чтобы приложение было доступно для пользователей, независимо от их способностей и психофизических особенностей развития. Это означает, что приложение должно быть совместимо со вспомогательными технологиями, такими как программы для чтения с экрана и распознавания голоса. Кроме этого, приложение должно соответствовать требованиям стандартов доступности, таких как руководство по доступности веб-контента (WCAG) [2].

Обеспечение обратной связи с пользователями является важным принципом эргономики. Его суть заключается в информировании пользователей о действиях, которые они совершают в работе с приложением, и об успешности этих действий с помощью визуальных подсказок, таких как анимация, шкала прогресса или предоставления обратной связи в виде текста [3].

Реализация принципа контекста использования означает, что приложение должно учитывать место, время и цель использования, чтобы обеспечить наилучший пользовательский опыт. Если пользователь использует приложение в общественном месте, то нужно ограничить использование громких звуков или яркого цвета. Если пользователь использует приложение на мобильном устройстве, то интерфейс должен быть адаптирован к небольшому размеру экрана и использованию пальцев рук для навигации.

Кроме этого, при разработке приложения, необходимо учитывать принцип цифровой безопасности пользователя. Пользователь должен чувствовать себя защищенным от возможных угроз. Для этого необходимо использовать защищенное соединение, чтобы передача конфиденциальной информации (например, паролей) была зашифрована и недоступна для злоумышленников. Также необходимо ограничить доступ к конфиденциальной информации только для авторизованных пользователей, чтобы избежать утечки данных. Важно предупреждать пользователя о возможных угрозах безопасности (например, о том, что пароль должен содержать специальные символы), чтобы пользователь мог принять меры для защиты своих данных. В целом, принцип безопасности помогает создать доверие у пользователей и защитить их данные от возможных угроз.

Каким бы многофункциональным не было приложение, без должной скорости работы пользователю будет некомфортно с ним взаимодействовать. Поэтому важно при разработке приложения руководствоваться принципом быстрого действия. Возможны ошибки в обработке данных и работе отдельных элементов программы. Приложение должно быть производительным и отзывчивым, чтобы пользователи не теряли много времени на ожидание загрузки страниц и ответа на свои запросы. Для этого при разработке приложения необходимо уделить внимание оптимизации кода и базы данных. Это позволит ускорить загрузку страниц и выполнение запросов к базе данных.

При разработке приложения необходимо соблюдение принципа консистентности. Это означает, что все элементы интерфейса должны быть единообразными по стилю и поведению. То есть пользователь должен понимать, как взаимодействовать с различными элементами интерфейса. Например, если кнопки в приложении имеют различный цвет, размер и расположение на разных страницах, то пользователю будет трудно понять, как они работают и как их использовать. Несоблюдение принципа консистентности может привести к тому, что пользователи будут испытывать трудности при использовании приложения. Каждый из вышеперечисленных принципов эргономики оказывает весомое значение на качество приложения, и являются базовыми принципами при разработке любого веб-приложения.

Существует ряд дополнительных принципов эргономики, отвечающих за тот или аспект приложений. Если разработчик не придерживается соблюдения принципов эргономики, то его

приложение считается не эргономичным. Это значит, что пользователь может столкнуться с трудностями и неудобствами при использовании приложения.

Поэтому важным является соблюдение основных принципов юзабилити, что оказывает значительное влияние на результативность (удовлетворенность) работы пользователя с приложением.

Заключение.

Следует отметить, что эргономика является одним из важнейших аспектов при разработке веб-приложений. Применяя подход к проектированию, ориентированный на пользователя, разработчики могут создавать веб-приложения, удобные и эффективные для использования человеком. Эти принципы должны быть интегрированы в процесс проектирования с самого начала, чтобы гарантировать, что созданное веб-приложение будет простым в использовании и удовлетворит потребности целевых пользователей [5].

Список литературы

- [1] Вайнштейн, Л. А. Эргономика: учеб. пособие. В 2 ч. Ч. 1 / Л. А. Вайнштейн. – Минск: БГУИР, 2018. – 208 с.: ил.
- [2] Обеспечение доступности веб-контента: стандарты, критерии, пример реализации – [Электронный ресурс] – <https://habr.com/ru/company/ispring/blog/564446/>;
- [3] "Web Design for Developers" Брайана Хогана -2010 год Издательство Programmatic Bookshelf. 300 страниц;
- [4] Консистентность в дизайне — [Электронный ресурс] – <https://infogra.ru/design/konsistentnost-v-dizajne>;
- [5] Юзабилити сайта: что это такое и как анализировать. – [Электронный ресурс] – <https://mediaaid.ru/blog/instructions/yuzabiliti-sayta/>.

USING BASIC ERGONOMIC PRINCIPLES IN WEB DEVELOPMENT

I.F Kirinovich

*Ph.D., associate professor of
engineering psychology and
ergonomics department BSUIR*

K. Dadaev

*Master student of engineering psychology
and ergonomics department BSUIR*

*Belarussian State University of Informatics and Radioelectronics
E-mail: frkeodal.rty@gmail.com*

Abstract. The principles of ergonomics used in the development of web applications are considered, allowing you to create convenient and easy-to-use interfaces that are accessible to any user.

Keywords: ergonomics, principles of ergonomics, web application, interface elements

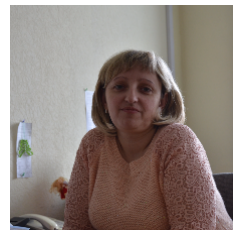
УДК 331.108:004.42

ПОСТРОЕНИЕ ИНДИВИДУАЛЬНОГО ПУТИ ОБУЧЕНИЯ ИТ-СПЕЦИАЛИСТА И СЕТЕВОЕ ВЗАИМОДЕЙСТВИЕ ВУЗОВ В ПРОЦЕССЕ ОБУЧЕНИЯ



В.В. Верняховская

Заместитель декана инженерно-экономического факультета БГУИР,
магистр экономических наук
verniahovskaya@bsuir.by,



О.М. Раптунович

Ассистент кафедры
экономической информатики
БГУИР
oraptunovich@gmail.com

В.В. Верняховская

Окончила Белорусский государственный университет информатики и радиоэлектроники. Магистр экономики. Работает заместителем декана инженерно-экономического факультета БГУИР. Направления исследований: трансфер технологий, инновационная деятельность, информационные технологии в маркетинге.

О.М. Раптунович

Окончила Частный Институт Управления и Предпринимательства. Работает ассистентом кафедры экономической информатики.

Аннотация. На сегодняшний день в экономике провозглашено новое направление развития – цифровая экономика. Основа новой экономической ориентации складывается в функциональном использовании электронных технологий, услуг, большом объеме данных, представленном в электронном варианте в информационной сети. Отсюда, ради эффективной реализации нужны специалисты, владеющие высоким уровнем знаний, умений и навыков в предоставленной области, умеющие ориентироваться в стремительно нарастающем потоке информации, способной к непрерывному самообразованию. Подобные условия подразумевают модернизацию высшего профессионального образования, в том числе и большую индивидуализацию процесса обучения каждого студента. Следовательно, одной из важнейших направленностей развития прогрессивного образования является узкая связь с реальной наукой, производством и экономикой. В ходе данной работы был проведен анализ компетенций со стороны работодателей и выпускников вуза, рассмотрены необходимые навыки, необходимые ИТ-специалисту.

Ключевые слова: профессиональные компетенции, цифровая экономика, профессиональные навыки, личностные качества, методы обучения, интернациональное партнерство, программа обмена.

Введение.

Одна из наиболее важных задач всех высших учебных заведений Беларуси – обеспечение качества обучения, получение студентами таких знаний и умений, которые помогут им находить решения для новых проблем в их будущей профессиональной деятельности [1].

В настоящее время высшее профессиональное образование республики определяется как практико-ориентированное, а компетентностный подход провозглашен в качестве базового. Преимуществами такого подхода является то, что в качестве результата берутся в расчёт не все знания, полученные студентом в процессе обучения, а способность адаптироваться к различным ситуациям и применять полученные навыки на практике.

Однако, какими бы преимуществами не обладал этот подход, в данный момент приходится констатировать несоответствие между компетенциями выпускников и требованиям со стороны работодателей. В белорусском педагогическом обществе противники компетентностного подхода отмечают, что при таком обучении молодых специалистов полностью отсутствует воспитательный эффект, другими словами, высшие образовательные заведения дают отличную базу, однако совсем не развивают душу человека и так называемые «гибкие навыки».

Основным фактором роста ИТ-сферы являются трудовые ресурсы [2].

Белорусский рынок труда в ИТ-сфере характеризуется высоким качеством технического образования, близостью к европейской бизнес-культуре на фоне общего сокращения и старения населения. Сформированная во время СССР система технического образования остается весьма качественной, хотя и требует ряда реформ.

В сентябре 2017 г. - январе 2018 г. в Республике Беларусь в рамках проекта *FOSTERC* был проведен опрос выпускников, преподавателей белорусских вузов и работодателей республики. В ходе данного опроса выпускниками оценивались методы преподавания, значимость компетентностного подхода, их реальный уровень по их собственным оценкам. Так же стоит отметить, что данный опрос проходили именно молодые специалисты, которые недавно устроились на работу, а также работодатели, предоставившие им «первое» рабочее место.

Очевидно, что выпускники вузов и их работодатели по-разному оценивают фактический и требуемый уровень компетенций. Вопрос в том, насколько самооценка компетенций выпускников вузов совпадает с оценкой данных компетенций их работодателями и насколько расходится с их востребованностью на рынке труда. Для оценки уровня сформированности и востребованности компетенций использовалась шкала от 1 до 5 (таблица 1).

Таблица 1. Уровень сформированности профессиональных компетенций у выпускников белорусских вузов, шкала оценки от 1 до 5.

Название компетенции	Сформированность компетенций		
	Мнение работодателей	Мнение выпускников	Разница
Использовать информационно-коммуникационные технологии	4,26	3,64	0,62
Открытость новым возможностям	4,08	3,49	0,59
Продуктивно работать в команде	4,01	3,59	0,42
Приобретать новые знания	4,22	3,86	0,36
Понимать сложные проблемы в целом	3,71	3,62	0,09
Диагностировать новые проблемы	3,48	3,45	0,03
Презентовать продукт, идеи или отчёты на публике	3,59	3,61	-0,02
Мобилизовать других	3,59	3,22	0,37
Эффективно вести переговоры	3,58	3,31	0,27
Общаться на иностранном языке	3,20	3,05	0,15

В ходе анализа представленных данных можно сделать вывод, что выпускники по всем показателям оценивают свою компетентность в различных областях ниже, чем их работодатели. Выше всего работодатели оценили использование информационно-коммуникационных технологий, кроме того, именно здесь разница между оценками составила 0,62 балла. Это может объясняться тем, что большинство работодателей как минимум половину своей жизни не были тесно связаны с использованием информационных технологий. Молодое поколение, в свою очередь, практически с самого детства получало возможность пользоваться продуктами технического процесса. Такая разбежка говорит о том, что выпускникам намного легче усваивать новые методы и технологии в сфере информационных технологий, что очень ценится работодателями, которым это сделать на порядок сложнее. Ниже всего оценили навыки презентации продуктов, мобилизации других, эффективное ведение переговоров. Определенно, любой компании необходимо иметь не только работников с профессиональными знаниями, а также эмоционально-устойчивых, легко адаптируемых к изменениям сотрудников. Эффективность работы зависит от умения вести себя в команде, грамотно излагать идеи [3]. По данным таблицы выпускникам белорусских вузов не хватает знаний в этих областях. Далее проанализируем пять самых востребованных компетенций на рынке труда.

Таблица 2. ТОП-5 самых востребованных компетенций среди работодателей и выпускников белорусских вузов

Название компетенции	Компетенции, востребованные на рынке труда		
	Мнение работодателей	Мнение выпускников	Разница
Приобретать новые знания	4,60	4,22	0,38
Использовать информационно-коммуникационные технологии	4,58	4,12	0,42
Продуктивно работать в команде	4,57	4,01	0,56
Аналитическое мышление	4,51	4,11	0,40
Презентовать продукт, идеи или отчёты	4,37	3,56	0,81

Согласно анализу, можно сделать вывод, что работодатели высоко оценивают не только профессиональные навыки, но и личностные качества. Ярким примером этого является востребованность умения презентовать продукт, который работодатели оценили в 4,37 балла, а выпускники на 0,81 балл ниже. В целом, можно сделать вывод, что так как у выпускников не так много опыта, то они не могут оценить востребованность своих компетенций.

Стоит отметить, что работодатели высоко ценят профессиональные знания, которые выпускники получают в процессе обучения в белорусских вузах, однако отмечают, что тренировка личностных качеств также необходима, чего явно не хватает в процессе обучения.

Исходя из анализа компетенций выпускников белорусских ВУЗов, мы получили, что по мнению работодателей студентам не хватает таких навыков как умение презентовать продукт и идеи, работать в команде продуктивно, проявлять лидерские качества. Несомненно, в некоторой степени групповые занятия помогают развивать эти навыки, но порой их оказывается недостаточно, ведь в основном студенты учатся самостоятельно. Разберем одни из самых необходимых личностных качеств для IT-специалиста и возможные варианты их развития в рамках университета.

1. *Творческое мышление.* Везде высоко ценятся программисты, которые не только могут написать код, обладая общими знаниями, но и подойти к нему творчески: изящно упростить строки и посмотреть на задачу под другим углом. Несомненно, дать хорошие базовые знания студентам – залог качественного обучения, однако обучение всех «под копирку» не даст желаемых результатов. Хороший способ развить творческое мышление – перенести задачу на пример из жизни, примером такого подхода могут быть задачи по мотивам игр, фильмов и т.д. Проводя аналогию между жизненными ситуациями и непосредственно написанием кода, поможет студентам взглянуть на задачи с разных сторон.

2. *Сила убеждения.* Данный навык поможет студентам научиться не бояться осуждения своих новых идей при высказывании их на публике, презентовать их, а также воспринимать критику в свою сторону. Для развития данного навыка в белорусских вузах предусмотрены научно-исследовательские конференции, однако формат их проведения недостаточно эффективный. Намного эффективнее было бы, если бы конференции проходили в виде диалога между студентами, обсуждения идеи и способов ее доработки, а не сухого чтения с листочка. Кроме этого, результатами своих работ студенты могут делиться в соцсетях, если им сложно контактировать с живой публикой, записывать видеоролики и оставлять комментарии.

3. *Сотрудничество.* Во время работы в команде необходимо верно распределить нагрузку между её членами, чтобы все выполняли приблизительно одинаковое количество работы. Необходимо следить за успеваемостью выполнения модулей, поэтому сотрудничество в команде – ключевой навык её члена. Научиться сотрудничеству помогут не только групповые занятия, но и занятия, где обучение проходит по принципу «говорю-слушаю», где студенты научатся выслушивать чужие точки зрения и высказывать свои, не нарушая границ членов группы [4].

4. *Адаптивность.* Сфера IT настолько непостоянная и изменчивая, с каждым годом направлений и технологий становится все больше, поэтому IT-специалисту придется постоянно адаптироваться к новым условиям работы. Данный навык помогут развить разнообразные формы занятий, вместо монотонных. Например, мозговые штурмы, турниры. Также этот навык можно развивать, ставя различные условия при выполнении работ, например, уложиться в короткий временной интервал или выполнить работу другим способом.

5. *Эмоциональный интеллект.* Многие карьеристы не обращают внимания на свои эмоции, а потом проваливаются «в дыру» и не могут работать. Учиться понимать свои эмоции, уметь их контролировать и выражать необходимо не только для специалиста в IT-сфере, но и для любого человека [5]. Периодические занятия с психологами помогли бы студентам научиться правильно отдыхать, распределять время, рассказывать о своих эмоциях другим и морально отдыхать от высоких темпов обучения.

Заключение.

Включение компетенций в структуру индивидуального интеллекта, позволяет в полной мере уделить внимание современному человеку и специалисту как когнитивному агенту, объединив его разумное (интеллектуальное) и деятельностное начало: мало обладать высоким уровнем интеллекта, необходимо научиться им пользоваться для достижения успеха. В ходе данной работы, как необходимая часть построения индивидуального пути обучения, было предложено обратить внимание на развитие личностных качеств студента, что было отмечено необходимым при анализе реальных компетенций работодателями и студентами. Данный подход поможет выпускникам не только стать профессионалами в своей области, но и быть востребованными специалистами.

Список литературы

- [1] Лабынцева, И. В. Адаптированность студентов к обучению в вузе: моногр. / И. В. Лабынцева. - М.: LAP Lambert Academic Publishing, 2012. - 112 с.
- [2] Клименко, А. В. Инновационное проектирование оценочных средств в системе контроля качества обучения в вузе / А.В. Клименко. - М.: Прометей, 2015. - 153 с.
- [3] Вахтина Е.А. Гуманизация обучения в вузе: технология дидактического проектирования: моногр. / Е. А. Вахтина. - М.: LAP Lambert Academic Publishing, 2014. - 232 с.
- [4] Акопян, М. А. Использование коммуникативных технологий в обучении студентов вуза / М. А. Акопян. - М.: LAP Lambert Academic Publishing, 2012. -164 с.
- [5] Головин, Ю. А.; Коханая О. Е. Инновационные Методы Обучения Студентов Университета В Институте Масс Медиа: Материалы Учебно-Методической Конференции - Москва, 18 Февраля 2008 Г / Евгеньевна Ю. А. Головин; О. Е. Коханая. - Москва: СИНТЕГ, 2008.- 789 с.

BUILDING AN INDIVIDUAL LEARNING PATH FOR IT SPECIALIST AND NETWORK INTERACTION OF UNIVERSITIES IN THE PROCESS OF EDUCATION

V. Vernyakhovskaya

Master of economics, Deputy Dean of the faculty of engineering and Economics at BSUIR

O. Raptunovich

Assistant of the Department of Economic Informatics at BSUIR

*Belarusian State University of Informatics and Radioelectronics, Republic of Belarus
E-mail: verniahovskaya@bsuir.by*

Abstract. A new direction of development has been announced in the economy today – the digital economy. The basis of the new economic orientation is formed in the functional use of electronic technologies, services, a large amount of data presented in electronic form in the information network. Hence, for the sake of effective implementation we need specialists who have a high level of knowledge, skills and abilities in the given area, who are able to navigate the rapidly growing flow of information capable of continuous self-emergence. Such conditions imply the modernization of higher professional education, including greater individualization of the learning process of each student. Consequently, one of the most important directions of the development of progressive education is a narrow connection with real science, production and economy. In the course of this work, an analysis of competencies on the part of employers and university graduates was carried out, the necessary skills needed by an IT specialist were considered.

Keywords: professional competencies, digital economy, professional skills, personal qualities, teaching methods, international partnership, exchange program

УДК 336.774.3

ОСОБЕННОСТИ УЛУЧШЕНИЯ ПОЛЬЗОВАТЕЛЬСКОГО ИНТЕРФЕЙСА МОБИЛЬНОГО ПРИЛОЖЕНИЯ



И. Ф. Киринович

доцент кафедры инженерной психологии и
эргономики БГУИР, к.ф.-м.н, доцент
Kirinovich.irina@yandex.ru



К. Дадаев

магистрант кафедры инженерной
психологии и эргономики БГУИР
frkeodal.rty@gmail.com

И. Ф. Киринович

С 2015 года – доцент кафедры инженерной психологии и эргономики БГУИР.

К. Дадаев

С 2022 года – магистрант кафедры инженерной психологии и эргономики БГУИР.

Аннотация. Ограниченное экранное пространство является одной из главных проблем проектирования пользовательского интерфейса для мобильных устройств. С ростом сложности приложений и функциональности приложений возрастают требования к удобству интерфейса. Изложены основные аспекты проектирования эргономичного пользовательского интерфейса для мобильных устройств.

Ключевые слова: пользовательский интерфейс, эргономика, дизайн интерфейса, эргономические показатели

Введение.

Проектирование пользовательского интерфейса включается в процессы разработки и тестирования программного продукта как часть системы качества. Разработка пользовательского интерфейса ведется параллельно с разработкой дизайна программного продукта в целом и в основном предшествует его реализации и является важным подготовительным этапом проектирования. Именно на этом этапе требуется учесть параметры эргономики при проектировании интерфейса программного продукта [1].

Исследования показывают, что незначительное улучшение эргономических свойств приводит к значительному повышению эффективности работы пользователей.

Следовательно, необходимо совершенствовать пользовательский интерфейс систем управления цифровым контентом в соответствии с критериями качества пользования приложения [2].

Анализ проблем.

Проектирование интерфейсов для мобильных устройств имеет ряд важных особенностей:

- различные типы устройств и различные размеры экранов: планшеты все чаще могут конкурировать по эргономическим показателям с настольными компьютерами и ноутбуками, а телефоны с сенсорными экранами, хоть и имеют небольшие по размеру дисплеи, позволяют реализовать полноценное приложение или мобильное приложение для мониторинга, если речь идет о больших системах и сами являются частью системной архитектуры;

- ориентация устройства в пространстве позволяет строить интерфейсы с индивидуальным расположением элементов, а альбомная ориентация присуща исключительно для мобильных устройств.

Эргономичные показатели качества – показатели степени соответствия параметров продукта психологическим и психофизическим характеристикам пользователей, подразделяемые на следующие группы [1]:

антропометрические, свойственные не только для мобильных устройств [2]:

- по содержанию – элементы отображения информации должны адекватно отображать объекты управления, внешнюю среду и состояние самого продукта;
- по количеству информации – элементы отображения информации обязаны обеспечить информационный баланс, т. е. представлять исчерпывающую информацию;
- сохранение текущего рабочего состояния приложения при повторном открытии приложения;
- по форме и композиции – следует обеспечить соответствие элементов отображения информации задачам, поставленным перед разработчиками интерфейса.

физиологические:

- приложение должно быстро реагировать на жесты пользователя, тем самым давая обратную связь;
- использование основных мобильных платформ есть четко выработанный набор рекомендаций, правил и принципов от создателей платформы, благодаря которым приложения под эти платформы выглядят единообразно;
- предопределенные жесты операционной системы, а также жесты, которые будут интуитивно понятны пользователю.

психофизические:

- скорость работы приложения при выполнении сложных операций, а также отзывчивость пользовательского интерфейса во время загрузки приложения;
- использование заряда батарейки устройства, а также его температура на пике нагрузок;
- автономность работы приложения.

С учетом того, что программное средство со временем увеличивает количество своих функций, проектирование пользовательского интерфейса не заканчивается на начальном этапе разработки программного средства, но предполагает дальнейшую работу по улучшению пользовательского интерфейса.

Улучшение пользовательского интерфейса приложения предполагает сбор и аналитику качественной информации о продукте, существует две стратегии получения пользовательских отзывов: пассивная и активная, что подразумевает под собой возможность использование широкого инструментария.

Пассивная:

- использование альфа-бета тестирования на ограниченных выборках пользователей;
- сбор обратной связи от пользователей во время фазы сопровождения приложения;

Активная:

- аналитика и анализ использования функций приложения [2];
- построение тепловых карт приложения, программа помогает пользователю спроектировать более эргономичный интерфейс, исходя из полученных данных на температурных картах. Это ускорит работу пользователя, уменьшит количество его ошибок, что полезно как в обучающей, так и в производственной сфере [3, 4];
- опрос пользователей, позволяющий получить качественную оценку эргономики приложения.

Выводы.

Таким образом, при проектировании интерфейса важным является постоянное взаимодействие дизайнера и программиста для достижения соответствия разрабатываемого интерфейса всем требованиям к качеству, в том числе и эргономическим.

Другим вариантом решения данной проблемы может быть наличие необходимой знаний в области дизайна и эргономики у разработчика интерфейса. Таким образом, показатель коммерциализации программного продукта с разработанным интерфейсом на основе элементов эргономики и дизайна будет значительно выше.

В данном тезисе определены основные эргономические показатели интерфейса мобильного приложения и методы, по которым эти показатели будут улучшаться.

Список литературы

- [1] Т.Ю. Забавникова Элементы эргономики проблеме проектирования интерфейса ISSN 1810-0198. Вестник ТГУ, т.14, вып.1, 2009.
- [2] Shneiderman B. Designing the User Interface: Strategies for Effective Human-Computer Interaction, 2010.
- [3] Эргодизайн пользовательского интерфейса: методы юзабилити-исследований В.С. Компаниец, А.Е. Лызь Инженерный вестник Дона, No3 (2017).
- [4] Шульга Т. Э., Данилов Н. А. Программный комплекс для сбора и визуализации данных активности пользователя настольного приложения: No 2014662094: Программа для ЭВМ. Россия, 06 Октябрь. 2014.

FEATURES OF IMPROVING THE USER INTERFACE OF THE MOBILE APP

I.F Kirinovich

*Ph.D., associate professor of engineering
psychology and ergonomics department
BSUIR*

K. Dadaev

*Master student of engineering psychology
and ergonomics department BSUIR*

*Belarussian State University of Informatics and Radioelectronics
E-mail: Kirinovich.irina@yandex.ru*

Abstract. Limited screen real estate is one of the major challenges in designing user interfaces for mobile devices. As the complexity of applications and the functionality of applications increase, the requirements for user-friendliness increase. The main aspects of designing an ergonomic user interface for mobile devices are outlined.

Keywords: user interface, ergonomics, interface design, ergonomic indicators

УДК 339.138.2:004.67

БУДУЩЕЕ МАРКЕТИНГА: АВТОМАТИЗАЦИЯ ВЗАИМОДЕЙСТВИЯ С КЛИЕНТАМИ С ПОМОЩЬЮ ИНТЕРНЕТА



Н.В. Лежнев

Студент инженерно-экономического факультета БГУИР
kolyaleshiov@gmail.com



О.Н. Шкор

Старший преподаватель кафедры экономики БГУИР
shkor@bsuir.by

Н.В. Лежнев

Родился в 2001 году в Орше. В 2019 году закончил ГУО «СШ №18 г. Барань». В этом же году поступил в УО «БГУИР», был зачислен на платную форму обучения по специальности «Электронный маркетинг» инженерно-экономического факультета.

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 рассмотрено влияние Интернета на бизнес и взаимодействие с клиентами, подчеркивается рост цифрового маркетинга и автоматизация обслуживания клиентов. В ней объясняется, как предприятия могут автоматизировать взаимодействие с клиентами через свой веб-сайт, используя такие технологии, как чат-боты, формы обратной связи, возможность заказать товары или услуги через сайт, анализ данных и машинное обучение.

В статье также рассматривается, как автоматизация может использоваться в других маркетинговых каналах, таких как кампании по электронной почте и платформы социальных сетей. Описаны преимущества автоматизации обслуживания клиентов, а именно более быстрое и эффективное обслуживание, более высокий уровень удовлетворенности и лояльности клиентов, а также высвобождение ценных ресурсов для бизнеса.

В заключении статьи прогнозируется, что по мере дальнейшего развития интернет-технологий предприятия будут находить еще более инновационные способы использования автоматизации для улучшения качества обслуживания клиентов.

Ключевые слова: Интернет, клиенты, цифровой маркетинг, будущее маркетинга, персонализация, автоматизация, веб-сайт, чат-боты, анализ данных, машинное обучение, удовлетворенность клиентов, лояльность.

Введение.

Интернет резко изменил мир маркетинга в современную эпоху. Развитие цифровых технологий и социальных сетей предоставило компаниям новые возможности для охвата своей целевой аудитории и взаимодействия с ней способами, которые раньше были невозможны.

Одним из наиболее значительных изменений, вызванных Интернетом, является возможность для бизнеса более эффективно ориентироваться на определенную аудиторию. С помощью социальных сетей предприятия могут создавать целевые рекламные кампании, которые охватывают людей с учетом их интересов, демографических данных и поведения в Интернете. Это позволило более эффективно использовать маркетинговые бюджеты и повысить рентабельность инвестиций для бизнеса.

Интернет также предоставил предприятиям новые способы взаимодействия со своей аудиторией. Через социальные сети и другие цифровые платформы компании могут

взаимодействовать с клиентами в режиме реального времени, отвечать на отзывы и решать проблемы. Это помогло укрепить доверие и лояльность среди клиентов, что может привести к увеличению продаж и повторных сделок [1].

Кроме того, Интернет упростил для предприятий измерение эффективности их маркетинговых кампаний. С помощью инструментов аналитики компании могут отслеживать трафик веб-сайта, взаимодействие с социальными сетями и коэффициенты конверсии, чтобы определить, какие стратегии наиболее эффективны. Это помогло предприятиям усовершенствовать свои маркетинговые стратегии и повысить общую эффективность.

Появление Интернета оказало глубокое влияние на то, как компании взаимодействуют со своими клиентами. С развитием цифрового маркетинга у компаний теперь есть доступ к множеству возможностей для обеспечения персонализированного и беспрепятственного обслуживания клиентов. Однако одним из наиболее значительных воздействий интернет-технологий на взаимодействие с клиентами стала автоматизация самого взаимодействия с клиентами [3].

Один из основных способов, с помощью которых предприятия могут автоматизировать работу с клиентами, - это веб-сайт. Теперь клиенты могут совершать транзакции, заказывать товары или услуги, получать доступ к информации и оставлять обратную связь, даже не разговаривая с представителем-человеком. Эта автоматизация стала возможной благодаря различным технологиям, включая формы обратной связи, чат-боты, анализ данных и машинное обучение.

Чат-боты – это виртуальные помощники, использующие обработку естественного языка (NLP) и искусственный интеллект (ИИ) для оказания мгновенной поддержки и помощи клиентам. Они могут выполнять множество задач, в том числе отвечать на часто задаваемые вопросы, направлять клиентов к соответствующим ресурсам и даже обрабатывать заказы.

Аналитика данных и машинное обучение также являются важными компонентами автоматизации веб-сайтов. Эти технологии позволяют компаниям анализировать данные о клиентах, выявлять закономерности и адаптировать маркетинговые сообщения для отдельных клиентов. Таким образом, предприятия могут предвидеть потребности своих клиентов и предоставлять актуальные и персонализированные рекомендации [2].

Формы обратной связи стали важным инструментом для предприятий, позволяющими автоматизировать свои операции и улучшить общее качество обслуживания клиентов. Одним из основных преимуществ форм обратной связи является то, что они позволяют компаниям быстро и эффективно собирать отзывы клиентов.

С помощью форм обратной связи клиенты могут высказать свое мнение и предложения о продуктах или услугах компании, что позволяет компании совершенствоваться и лучше обслуживать своих клиентов. Формы обратной связи также могут помочь предприятиям определить области, в которых им необходимо улучшить свои продукты или услуги, что приведет к повышению удовлетворенности клиентов и повышению их лояльности.

В дополнение к формам обратной связи функции записи на определённые услуги на веб-сайтах также помогают предприятиям автоматизировать свои операции. Позволяя клиентам записываться на услуги или заказывать товары через веб-сайт, предприятия могут оптимизировать свои процессы и снизить потребность в человеческом вмешательстве. Это не только экономит время и деньги, но и предоставляет клиентам более удобный и эффективный способ доступа к необходимым им услугам.

В целом, формы обратной связи и функции регистрации на веб-сайтах – это мощные инструменты, которые предприятия могут использовать для автоматизации своих операций и повышения качества обслуживания клиентов. Собирая ценные данные о клиентах, оптимизируя процессы и предоставляя более персонализированный опыт, предприятия могут повысить удовлетворенность и лояльность клиентов, что в конечном итоге приведет к увеличению прибыльности и успеха.

Помимо автоматизации деятельности с помощью веб-сайтов, предприятия также могут автоматизировать взаимодействие с клиентами через другие маркетинговые каналы, такие как кампании по электронной почте и платформы социальных сетей. Маркетинговые кампании по электронной почте могут быть адаптированы для отдельных клиентов с использованием анализа данных, а платформы социальных сетей позволяют компаниям взаимодействовать с клиентами в режиме реального времени без вмешательства человека.

Преимущества автоматизации клиентского опыта очевидны.

Устраняя необходимость вмешательства человека, предприятия могут предоставлять своим клиентам более быстрое и эффективное обслуживание. Это приводит к более высокому уровню удовлетворенности и лояльности клиентов, что может оказать положительное влияние на итоговый результат. Кроме того, автоматизация может высвободить ценные ресурсы для предприятий, позволяя им сосредоточиться на других аспектах своей деятельности [3].

Кроме того, автоматизация также может помочь предприятиям оптимизировать свои операции и сократить расходы. Автоматизируя рутинные задачи, такие как обработка заказов, управление запасами и обслуживание клиентов, предприятия могут работать более эффективно и снизить вероятность ошибок. В конечном итоге это может привести к снижению затрат и увеличению прибыльности.

Еще одно преимущество автоматизации заключается в том, что она может помочь предприятиям легче масштабировать свою деятельность. По мере роста числа клиентов предприятиям может быть сложно обрабатывать запросы клиентов и выполнять другие задачи вручную. Автоматизация может помочь смягчить эту проблему, предоставляя масштабируемое и эффективное решение, способное обрабатывать большие объемы запросов без ущерба для качества [4].

Тем не менее, предприятия должны быть осторожны, чтобы не слишком полагаться на автоматизацию. Хотя автоматизация может дать значительные преимущества, важно помнить, что клиенты по-прежнему ценят человеческое взаимодействие и персональное внимание. Таким образом, предприятия должны стремиться найти баланс между автоматизацией и взаимодействием с человеком, чтобы обеспечить наилучшее качество обслуживания клиентов.

Заключение.

Интернет открыл новую эру взаимодействия с клиентами за счет автоматизации обслуживания клиентов. Благодаря автоматизации веб-сайтов и другим возможностям цифрового маркетинга предприятия могут предоставлять своим клиентам более быстрое и эффективное обслуживание, одновременно высвобождая ценные ресурсы.

Поскольку использование интернет-технологий продолжает расти, мы можем ожидать появления еще более инновационных способов, с помощью которых предприятия используют автоматизацию для улучшения качества обслуживания клиентов.

Список литературы

[1] Что такое интернет-маркетинг [Электронный ресурс]. – Режим доступа: <https://pixelplus.ru/samostoyatelno/stati/marketing/chto-takoe-internet-marketing.html>

[2] Итинсон К.С. Цифровые технологии: четвертая промышленная революция // Региональный вестник. – 2020. – № 1. – 68- 69 с.

[3] Автоматизация интернет-продаж [Электронный ресурс]. – Режим доступа: <https://vc.ru/trade/94738-avtomatizaciya-internet-prodazh>

[4] Автоматизация бизнеса в интернет-магазине [Электронный ресурс]. – Режим доступа: <https://winner-bagz.ru/news/avtomatizatsiya-biznesa-v-internet-magazine/>

THE FUTURE OF MARKETING: AUTOMATING CUSTOMER INTERACTION WITH THE HELP OF THE INTERNET

N.V. Leshchev

Student of engineering and economics at the BSUIR

O.N. Shkor

Senior Lecturer at the Department of Economics at the BSUIR

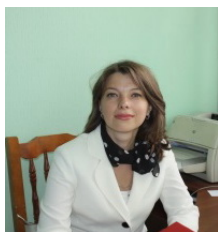
*Belarusian State University of Informatics and Radioelectronics
Minsk, Republic of Belarus
E-mail: kolyaleshiov@gmail.com*

Abstract. The article examines the impact of the Internet on business and customer interaction, highlighting the growth of digital marketing and customer service automation. It explains how businesses can automate interactions with customers through their website using technologies such as chatbots, feedback forms, the ability to order goods or services through the site, data analysis and machine learning. The article also looks at how automation can be used in other marketing channels such as email campaigns and social media platforms. The benefits of customer service automation are described, namely, faster and more efficient service, higher levels of customer satisfaction and loyalty, as well as the release of valuable resources for the business. The paper concludes by predicting that as Internet technologies continue to evolve, businesses will find even more innovative ways to use automation to improve the customer experience.

Keywords: Internet, customers, digital marketing, the future of marketing, personalization, automation, website, chatbots, data analysis, machine learning, customer satisfaction, loyalty.

УДК [304.2]

НОВЫЕ ПОДХОДЫ В ПОДГОТОВКЕ HR-СПЕЦИАЛИСТОВ В УСЛОВИЯХ ЦИФРОВОЙ ТРАНСФОРМАЦИИ



В.В. Шаталова
директор БГУИР филиал МРК,
канд. техн. наук, доцент
shatalova@bsuir.by



Т.В. Казак
заведующий кафедрой инженерной
психологии и эргономики, член-корреспондент
Международной академии психологических
наук, доктор психологических наук,
профессор
kazak@bsuir.by

В.В. Шаталова

Окончила Белорусский государственный университет информатики и радиоэлектроники, факультет компьютерного проектирования. Работает в должности директора БГУИР филиал «Минский радиотехнический колледж»

Т.В. Казак

Заведующий кафедрой инженерной психологии и эргономики Белорусского государственного университета информатики и радиоэлектроники, доктор психологических наук, профессор, член-корреспондент Международной академии психологических наук

Аннотация. В статье рассмотрены особенности подготовки специалистов в области управления человеческими ресурсами в условиях глобальной цифровой трансформации.

Ключевые слова: цифровизация, учебный план, профессиональные компетенции, универсальные компетенции, компетентностный подход, профилизация.

Глобальная цифровая трансформация – уже давно не абстрактный тренд, а современная реальность, в полной мере определяющая развитие компаний и глобальных рынков. Производственная среда, здравоохранение и медицина, транспорт и логистика, сельское хозяйство и прочие отрасли претерпевают существенные изменения из-за повсеместного внедрения информационных технологий и систем.

В эпоху таких глобальных и быстрых изменений предприятия и компании не могут работать по старым моделям. Перед современными бизнес-организациями стоит актуальная задача цифровизации абсолютно всех бизнес-процессов, в том числе и в рамках системы управления человеческими ресурсами [2].

Цифровизация процесса управления персоналом – это процесс организации удобной цифровой среды для сотрудников с целью повышения их мотивации и заинтересованности в работе, а также улучшения HR-бренда самой организации [1].

Управление человеческими ресурсами предполагает широкое применение цифровых технологий, которые являются необходимыми атрибутами конкурентоспособного предприятия.

В рамках Государственной программы «Цифровое развитие Беларуси на 2021-2025 годы» предусматривается выполнение мероприятий по созданию (развитию) современной информационно-коммуникационной инфраструктуры, внедрению цифровых инноваций в отраслях экономики [6].

Целью Государственной программы является обеспечение внедрения информационно-коммуникационных и передовых производственных технологий в отрасли национальной

экономики и сферы жизнедеятельности общества. Для достижения данной цели в рамках Государственной программы должна быть решена одна из ключевых задач – обеспечение доступности образования, основанного на применении современных информационных технологий для подготовки граждан к жизни и работе в условиях цифровой экономики.

Новое поколение образовательных стандартов и учебных планов по специальностям (поколение 3+) позволяет решить поставленные задачи: обеспечение фундаментальности, практико-ориентированности и актуальности содержания подготовки, а также обеспечение связи с рынком труда и возможности оперативного реагирования на его запросы.

Новое поколение образовательных стандартов и учебных планов по специальностям высшего образования разрабатываются на основе следующих принципов: компетентностный подход; преемственность содержания образования на различных уровнях (ступенях) основного образования; модульный принцип проектирования содержания образовательных программ.

В соответствии с ОКРБ 011-2022 «Специальности и квалификации» специальность 7-06-0611-01 Информационные системы и технологии относится к профилю образования «Информационно-коммуникационные технологии», группе специальностей «Прикладные информационные и коммуникационные технологии», что позволяет сформировать углубленный уровень знаний, умений и навыков для осуществления научно-исследовательской, научно-педагогической, учебно-методической, проектной, организационно-управленческой, профессиональной, инновационной деятельности в области цифровых технологий [5].

Учебный процесс в магистратуре предполагает подготовку по профилизациям в рамках одной специальности. Каждая профилизация включает набор общепрофессиональных и специальных дисциплин, в том числе по выбору, формирующих различные компетенции: универсальные, углубленные профессиональные и специализированные компетенции.

Углубленные профессиональные компетенции – компетенции, формируемые в соответствии с требованиями к магистру и отражающие его способность решать инновационные задачи профессиональной деятельности в соответствии с полученной специальностью.

Универсальные компетенции – компетенции, формируемые в соответствии с требованиями к магистру и отражающие его способность применять углубленные научно-теоретические, методологические знания и исследовательские умения, а также социально-личностные качества, соответствующие запросам государства и общества.

Формирование специализированных компетенций профессиональной деятельности возможно с учетом профилизации образовательной программы магистратуры по специальности на основе требований рынка труда, обобщения зарубежного опыта, проведения консультаций с организациями, имеющими потребность в подготовке магистров.

Маркетинговые исследования анализа управления человеческими ресурсами на предприятиях Республики Беларусь показали, что организации сталкиваются с существенными изменениями контекста управления человеческими ресурсами, условий и организации работы.

Ряд сложившихся глобальных тенденций неизбежно ведет к значительным изменениям в сфере управления человеческими ресурсами, среди которых можно выделить наиболее значимые направления:

- персонализация подходов к управлению человеческими ресурсами (HR);
- рост цифровизации и автоматизации HR-процессов.

Чтобы осуществить цифровую трансформацию в управлении человеческими ресурсами, специалисту необходимо обладать не только цифровой грамотностью, но и системными знаниями в трех областях: в образовании, компьютерных технологиях и управлении. Это и явилось предпосылкой открытия подготовки с 2023 года профилизации «Цифровизация в психологии управления человеческими ресурсами» по специальности 7-06-0611-01 «Информационные системы и технологии» впервые в Республики Беларусь [4].

Проектирование примерного учебного плана по специальности углубленного высшего образования 7-06-0611-01 «Информационные системы и технологии» выполнялось с учетом

методических рекомендаций по проектированию новых образовательных стандартов и учебных планов (поколение 3+) [3].

Соотнесение учебных дисциплин (модулей) и компетенций (за исключением отдельных универсальных компетенций) в новом типовом учебном плане специальности «Информационные системы и технологии» в целях создания условий для диагностирования компетенций построено следующим способом:

а) каждой компетенции соответствует одна учебная дисциплина либо группа учебных дисциплин, объединенная в модуль и предусматривающая единую форму контроля по модулю;

б) проектирование нескольких уровней компетенций, когда формулируются более общие (интегральные) компетенции, соответствующие одному либо нескольким модулям, которые затем делятся на составляющие компетенции, каждая из которых формируется одной входящей в модуль учебной дисциплиной либо группой учебных дисциплин, объединенной в модуль и предусматривающей единую форму контроля по модулю.

Соотношение часов на изучение дисциплин государственного компонента и компонента учреждения образования составляет 35% и 65% соответственно.

Государственный компонент включает два основных модуля: проектирование информационных систем и технология информационных систем.

В результате освоения дисциплин этих модулей магистранты должны обладать следующими универсальными компетенциями (применять методы научного познания в исследовательской деятельности, генерировать и реализовывать инновационные идеи; обеспечивать коммуникации, проявлять лидерские навыки, быть способным к командообразованию и разработке стратегических целей и задач, развивать инновационную восприимчивость и способность к инновационной деятельности) и углубленные профессиональные компетенции (применять методы оценки и экспертизы проектных решений при моделировании и проектировании сложных эргатических систем; моделировать, разрабатывать и исследовать пользовательские интерфейсы; проектировать и моделировать архитектуру информационных систем; применять технологию блокчейн для децентрализованных баз данных; разрабатывать информационные системы больших данных и владеть методами машинного обучения).

В процессе обучения магистранты будут получать проектные, производственно-технологические, информационно-аналитические и иные навыки, необходимые для создания умного промышленного производства, развития сопутствующих цифровых услуг и перехода промышленного предприятия к цифровым бизнес-моделям.

Областью действия выпускника является стратегический и логический подход к управлению наиболее ценным активом любой организации, работающими людьми, которые коллективно и индивидуально вносят вклад в решение ее задач; проектирование формальных систем организации управления, которые обеспечат эффективное использование человеческих знаний, навыков и талантов для достижения организационных целей; оптимизация и адаптация программных продуктов для различных отраслей экономики; проведение исследований и внедрение результатов научно-исследовательской деятельности в адаптацию и оптимизацию программных продуктов, представляющих практический интерес для социально-экономического развития Республики Беларусь.

Особый интерес вызывает перечень специализированных компетенций с учетом профилизации «Цифровизация в психологии управления человеческими ресурсами». Важным отличием данного направления профилизации, указанной специальности, является фундаментальная подготовка не только в области цифровых технологий на промышленных предприятиях, использование информационных технологий в экономической и управленческой деятельности, но и будущие специалисты будут дополнительно обладать компетенциями в области психологического сопровождения управленческих процессов предприятия, что

особенно важно в эпоху Индустрии 4.0 и цифровой трансформация промышленных предприятий.

В учебный план включены такие уникальные авторские дисциплины, как «Коммуникативные технологии управления человеческими ресурсами», «Нейролингвистическое программирование», «Психологическое сопровождение искусственного интеллекта», «HR-трансформация в цифровой среде» и многие другие.

При этом новое направление специальности ориентировано главным образом на промышленный сектор. По окончании программы студенты смогут проектировать цифровые среды, управлять проектами в области цифровой трансформации, проводить экспертизу образовательных программ, продуктов и сервисов.

Сегодня тренд цифровизации очевиден. Он охватывает все сферы жизни, начиная от изменения быта каждого, разработки более совершенной продукции, моделирования развития рынков, социума и заканчивая прогнозированием поведения людей и формированием экономики и общества нового типа.

Список литературы

[1] Дроздов, И.Н. Цифровизация управления человеческими ресурсами в бизнес-организации: значимость и актуальные направления / И.Н. Дроздов, Лю Вэйсяо // Креативная экономика. – 2022 – Том 16 – № 6 – С. 2433–2444.

[2] Кибербезопасность цифровой индустрии. Теория и практика функциональной устойчивости к кибератакам / Под ред. профессора РАН, доктора технических наук Д.П. Зегжды. – М.: Горячая линия – Телеком, 2020. – 560 с: ил.

[3] Методические рекомендации по проектированию новых образовательных стандартов и планов (поколение 3+), утверждено министром образования Республики Беларусь, 28.05.2018.

[4] Об открытии подготовки по специальности углубленного высшего образования, приказ Министерства образования Республики Беларусь, № 683, 25.11.2022.

[5] Об утверждении, введении в действие и отмене общегосударственного классификатора Республики Беларусь: Постановление Министерства образования Республики Беларусь, 24.03.2022, №54 // Национальный правовой Республики Беларусь [Электронный ресурс]. – 2022. – Режим доступа : <https://pravo.by/>.

[6] О Государственной программе «Цифровое развитие Беларуси» на 2021–2025 годы: постановление Совета Министров Республики Беларусь, 02.02.2021, №66 // Национальный правовой Республики Беларусь [Электронный ресурс]. – 2023. – Режим доступа : <https://pravo.by/>.

NEW APPROACHES OF TRAINING HR SPECIALISTS IN THE CONTEXT OF DIGITAL TRANSFORMATION

V.V. Shatalova

*Deputy Dean of the Faculty of
Computer-Aided Design, PhD,
Associate Professor*

T.V. Kazak

*Head of the Department of Engineering Psychology and
Ergonomics, Corresponding Member of the International
Academy of Psychological Sciences, Doctor of
Psychological Sciences, Professor*

*Belarusian State University Informatics and Radioelectronics, Republic of Belarus
E-mail: shatalova@bsuir.by*

Abstract. The article discusses the features of training specialists in the field of human resource management in the context of global digital transformation.

Keywords: digitalization, curriculum, professional competencies, universal competencies, competence approach, profiling.

УДК 004.021:004.75

СРАВНИТЕЛЬНЫЙ АНАЛИЗ ЭЛЕКТРОЭНЦЕФАЛОГРАММ



П.Ю. Бранцевич

Докторант, кандидат технических наук, доцент
branc@bsuir.edu.by

П.Ю. Бранцевич

С 1985 года работал в области разработки архитектуры, методов, алгоритмов и программного обеспечения компьютерных систем и комплексов для определения метрологических характеристик виброизмерительных преобразователей и виброустановок, решения задач вибрационного контроля, мониторинга, диагностики и автоматизации защиты многоопорных механизмов с вращательным движением. Разработал математическое и программное обеспечение компьютерных систем поддержки принятия решений по оценке технического состояния сложных механизмов и агрегатов по вибрационным параметрам и характеристикам. Автор трех монографий. Являлся научным руководителем и исполнителем четырех заданий Государственной научно-технической программы (ГНТП) «Энергетика», двух заданий ГНТП «Защита от чрезвычайных ситуаций», четырех заданий ГПНИ «Диагностика», более ста хозяйственных договоров.

Аннотация. Медицинская и технической диагностика, как набор правил, методов, алгоритмов, которые позволяют принять решение о состоянии наблюдаемого объекта или субъекта имеют много общего. Создание новых и совершенствование уже существующих методов оценки состояния человека является важным направлением медицинских исследований. Несомненный научный интерес представляет исследование нейронной сети головного мозга человека с использованием электроэнцефалограмм, представляющих собой сложные цифровые сигналы. В настоящее время для лечения резистентных форм психических и поведенческих расстройств применяется метод электросудорожной терапии. Актуальным является определение информативно-значимых параметров ЭЭГ, которые свидетельствовали бы об её изменении по результатам ЭСТ. Представлены результаты некоторых способов цифровой обработки временных реализаций электроэнцефалограмм, примененных для для сравнительного анализа ЭЭГ, полученных до и после проведения сеансов ЭСТ.

Ключевые слова: сигнал, электроэнцефалограмма, параметр, характеристика, цифровая обработка, решение.

Введение.

Медицинская диагностика, как набор правил, методов и решений, которые позволяют прийти к заключению о наличии или вероятности наличия у человека того или иного заболевания, имеет много общего с технической диагностикой [1]. Разработка новых и совершенствование уже существующих методов оценки состояния и лечения человека является важным направлением медицинских исследований. Поэтому применение разнообразных способов обработки параметров и характеристик человеческого организма и формализация систем принятия решений являются весьма актуальными. Вычислительная и информационная мощность современных технических средств, в том числе и мобильных, позволяет существенно расширить подходы к решению разнообразных задач, связанных с обработкой длинных реализаций разнообразных сигналов.

В результате научных исследований выяснено, что основным функциональным элементом мозга является нейрон [2]. Однако вопрос оценки результативности взаимодействия нейронов головного мозга остается проблемным. Можно предположить гипотезу, что устройство головного мозга представляет собой ядро и окружающую его оболочку. Базовое состояние ядра даётся человеку от рождения. Деятельность ядра головного мозга определяет наши способности и осуществляет верховное управление.

Вторая часть мозга – это самообучаемая нейронная сеть, которая имеет возможность реконфигурироваться, получать, накапливать информацию и производить настройку обрабатывающих функций и коэффициентов передачи для сигналов, поступающих через дендриты и синапсы в нейрон, формируя, тем самым, для данного момента времени систему принятия решений, сущность которых зависит от окружающего пространства и состояния организма, информацию о которых передают в нейронную сеть первичные человеческие преобразователи информации. Все действия и ощущения, которые получает и воспроизводит человек есть сущность и отражение решений, принимаемых нейронной сетью головного мозга. Самообучение нейронной сети производится эмпирически, или под целенаправленным воздействием, в том числе и по желанию нейронной сети, т.е. принятых ею решений на проведение определённых действий по получению новой информации.

Если следовать данной модели, то можно аргументировать вывод, что поведение человека обуславливается первоначальной настройкой нейронной сети головного мозга, и последующей реконфигурацией самообучаемой, многоуровневой, объёмной нейронной сети его головного мозга, происходящей по мере получения новой или переработки имеющейся информации.

Работа мозга сопровождается изменением электромагнитного поля, которое можно зафиксировать специальными первичными преобразователями и преобразовать в виде изменяющихся параметров тока или напряжения, что и происходит, когда снимают электроэнцефалограмму [3]. Электроэнцефалограмма головного мозга (ЭЭГ) – процедура, которая проводится для определения электрической активности головного мозга для выявления очагов повышенной судорожной готовности его коры, что характерно для: эпилепсии; опухолей; состояний после перенесенного инсульта; структурных и метаболических энцефалопатий; расстройств сна и других заболеваний. Электроэнцефалограмма показывает симптомы нарушения работы головного мозга, позволяет оценить характер отклонений и степень их распространенности [4-6].

Электроэнцефалограмма, как цифровой сигнал.

Электроэнцефалограмма отражает колебания напряжения в результате ионного тока в нейронах головного мозга и является электрическим сигналом, как результатом спонтанной электрической активности мозга в течение определенного периода времени, записанной с нескольких электродов на мозге или поверхности скальпа. Стандартной системой размещения электродов на поверхности головы, рекомендованной Международной федерацией электроэнцефалографии и клинической нейрофизиологии является система «10–20%» [6].

Результатом ЭЭГ являются цифровые сигналы, к которым исследуют с помощью методов цифровой обработки сигналов (ЦОС) [7-8].

Наиболее распространен спектральный метод исследования сигналов ЭЭГ, в ходе которого анализируется мощность сигнала в частотных полосах, называемых: альфа (8-13 Гц); бета (13-35 Гц); гамма (35-70 Гц); дельта (0.3-4 Гц); тэта (4-8 Гц) [5-6].

При этом, в большинстве случаев вычисление амплитудного спектра производится на 1024-х или меньшем количестве дискретных точек [9-10], что приводит к достаточно грубому частотному разрешению. Однако, в настоящее время для исследования электроэнцефалограмм в режиме реального времени можно применить и другие способы обработки для выявления новых параметров и информативно-значимых признаков. Это цифровая фильтрация, вейвлет обработка, преобразование Гильберта-Хуанга, гистограммы распределения, разложение сигнала на детерминированные и шумоподобные компоненты, построение временных трендов и их обработка [11].

На рис.1 показан один из сигналов ЭЭГ (отведение FP2-АА), а на рис. 2 его амплитудный спектр, вычисленный для частотного разрешения 0.03 Гц.

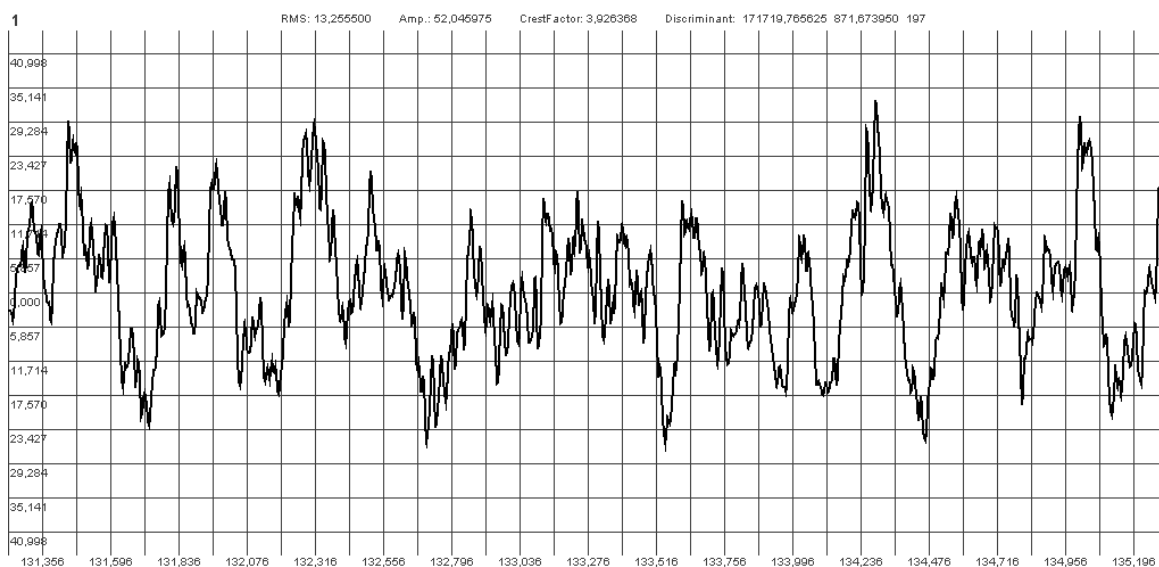


Рисунок 1. Сигнал ЭЭГ, полученный при частоте дискретизации 250 Гц

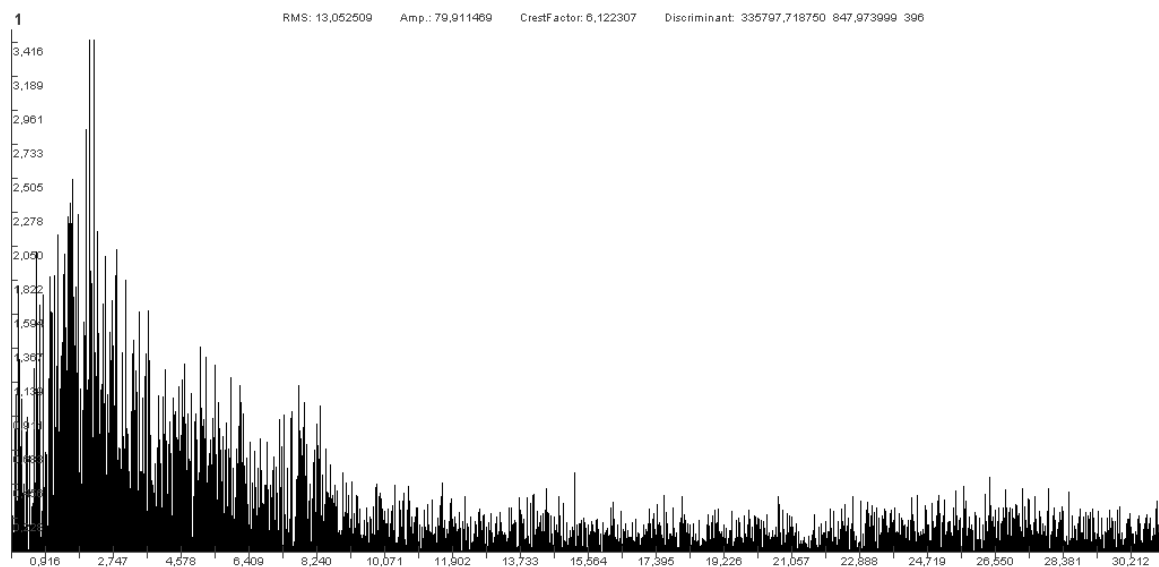


Рисунок 2. Амплитудный спектр сигнала ЭЭГ на временном интервале 32,768 с.

Проведенные исследования сигналов ЭЭГ показали, что изменение частотного разрешения спектрального анализа приводит к изменению структуры их амплитудных спектров, что свидетельствует о случайном характере сигналов ЭЭГ[12]. При этом на определенных временных интервалах анализа гистограмма распределения сигнала ЭЭГ по уровням близка к нормальному закону распределения (рис. 3).

Полосовой спектр данного сигнала для диапазонов альфа (8-13 Гц); бета (13-35 Гц); гамма (35-70 Гц); дельта (0.3-4 Гц); тэта (4-8 Гц) приведен на рис. 4.

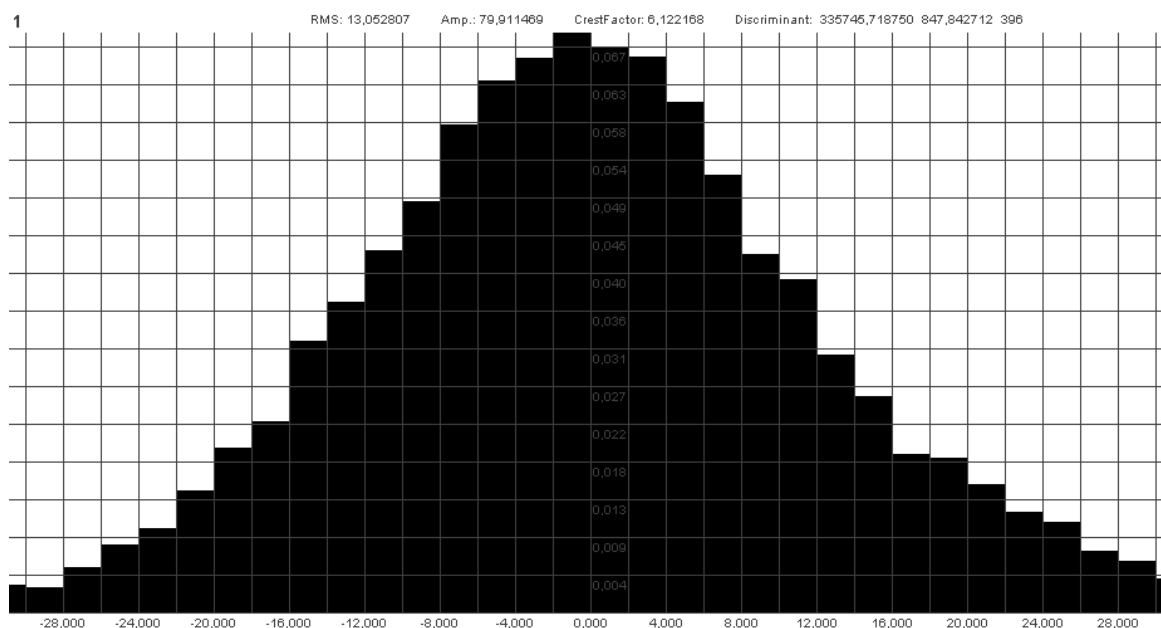


Рисунок 3. Гистограмма распределения по уровням сигнала ЭЭГ на временном интервале 32,768 с.

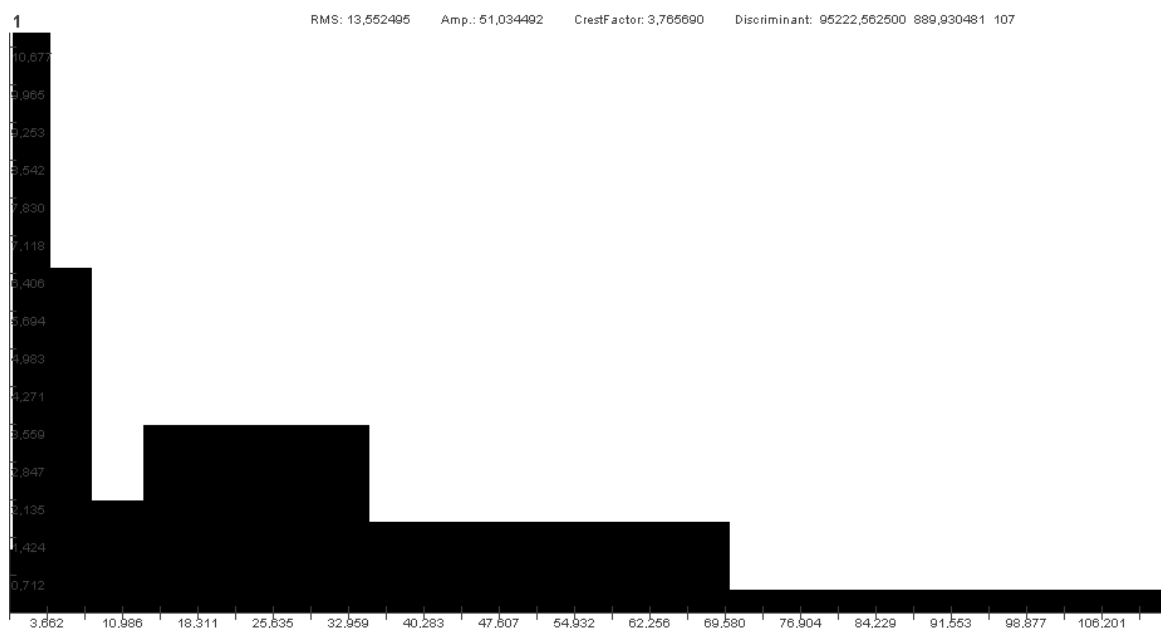


Рисунок 4. Полосовой спектр сигнала ЭЭГ на временном интервале 8,192 с.
Границы частотных полос 0,3, 4, 8, 13, 35, 70, 125 Гц

Число заболеваний, которые приводят к изменениям ЭЭГ человека достаточно велико. Наиболее часто с её помощью диагностируют эпилепсию. Данный недуг является очень опасным и проявляется в приступах и судорогах, во время которых больной теряет сознание.

Электросудорожная терапия.

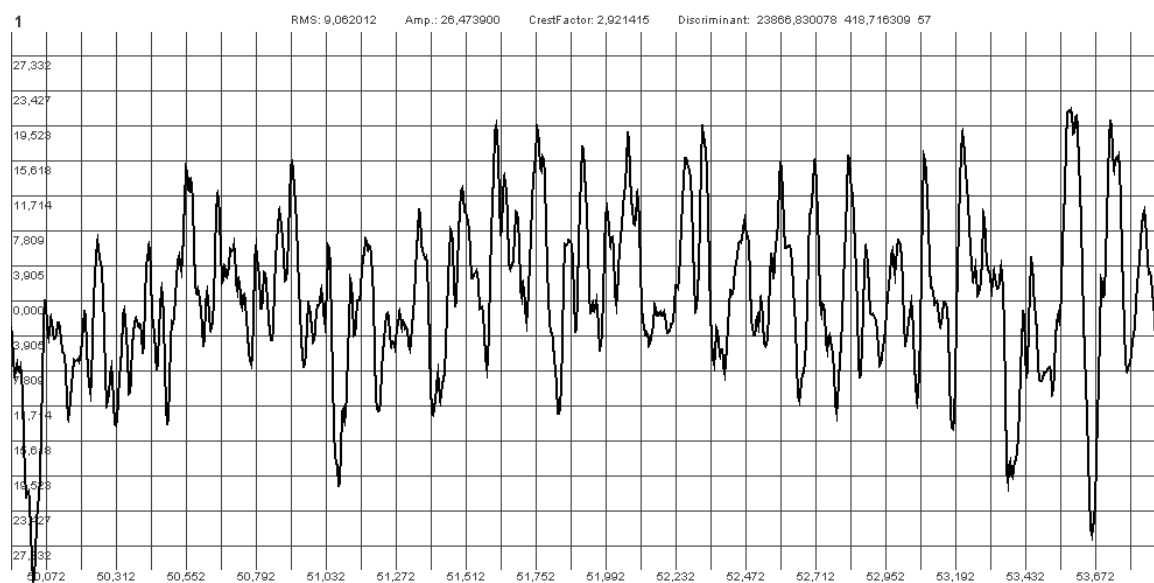
Электросудорожная терапия (ЭСТ), иначе называемая электроконвульсивной терапией (ЭКТ), ранее известная как электрошок (ЭШ) или электрошоковая терапия (ЭШТ), – метод психиатрического и неврологического лечения, при котором эпилептиформный большой судорожный приступ вызывается пропусканием электрического тока через головной мозг

пациента с целью достижения лечебного эффекта [13]. ЭСТ можно использовать, только если другие мероприятия, например, лекарства и психотерапия, не дали результата [14].

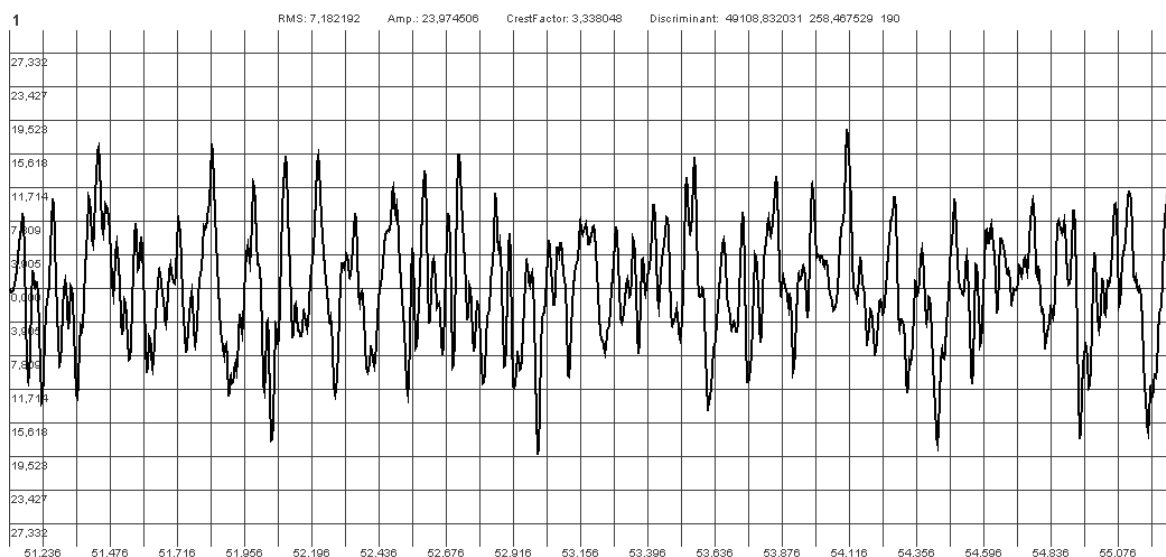
Нейронная сеть головного мозга человека подвергается сильному воздействию, однако пока не существует объективных информативно-значимых признаков на основе анализа ЭЭГ, которые показывали бы эффективность проведенных процедур ЭСТ. В данной работе предпринята попытка проведения сравнительного анализа ЭЭГ, полученных до и после проведения сеансов ЭСТ.

Сравнительный анализ ЭЭГ.

Рассмотрим сигналы ЭЭГ (отведение О1-АА), полученные при обследовании пациента до и после ЭСТ, рис.5.



а) до процедуры ЭСТ



б) после процедуры ЭСТ

Рисунок 5. Отрезки сигналов ЭЭГ, отведение О1-АА

Оценить изменчивость сигнала ЭЭГ можно анализируя временной тренда её параметров [15]. Параметры среднеквадратическое значение (СКЗ), пик-фактор, эксцесс,

асимптота вычислялись на 1024 точках, что при частоте дискретизации 250 Гц соответствует временному интервалу 4,096 с. Шаг по времени – одна секунда. На рис. 6 показаны временные тренды указанных параметров.

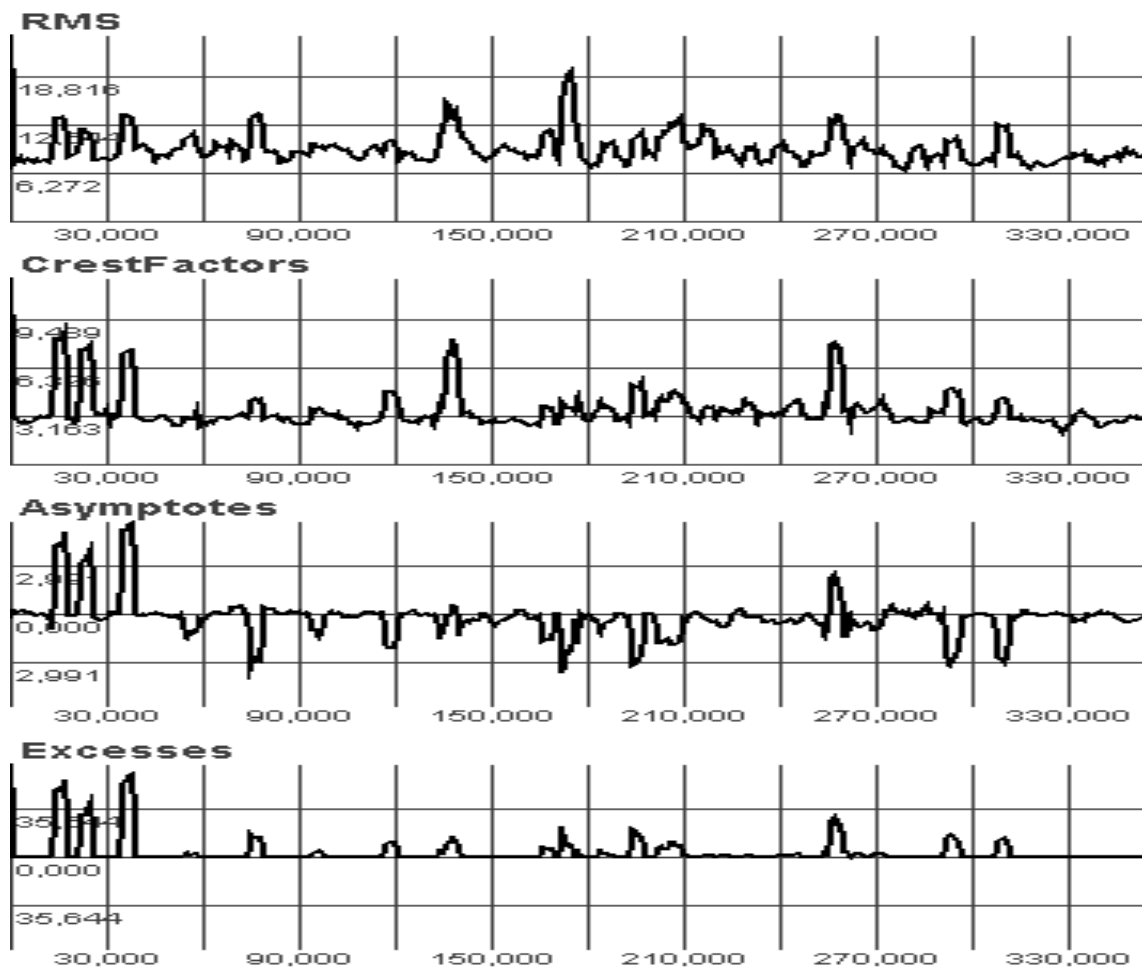


Рисунок 6. Временные тренды параметров сигнала ЭЭГ до ЭСТ.
Ось абсцисс – время, с. Ось ординат – значение параметра

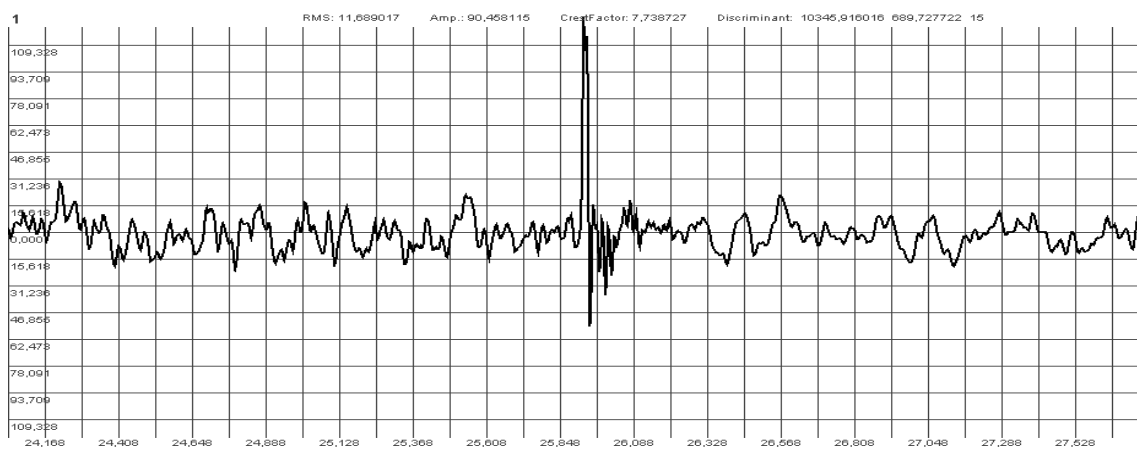


Рисунок 7. Отрезок временной реализации сигнала ЭЭГ с артефактом до ЭСТ

Амплитудные всплески параметров на временном тренде являются отражением возмущений на сигнале ЭЭГ (см. рис. 7). Следовательно, анализируя временные тренды можно локализовать ряд артефактов на сигналах ЭЭГ.

На рис. 8 показаны временные тренды указанных параметров ЭЭГ после проведения сеансов ЭСТ.

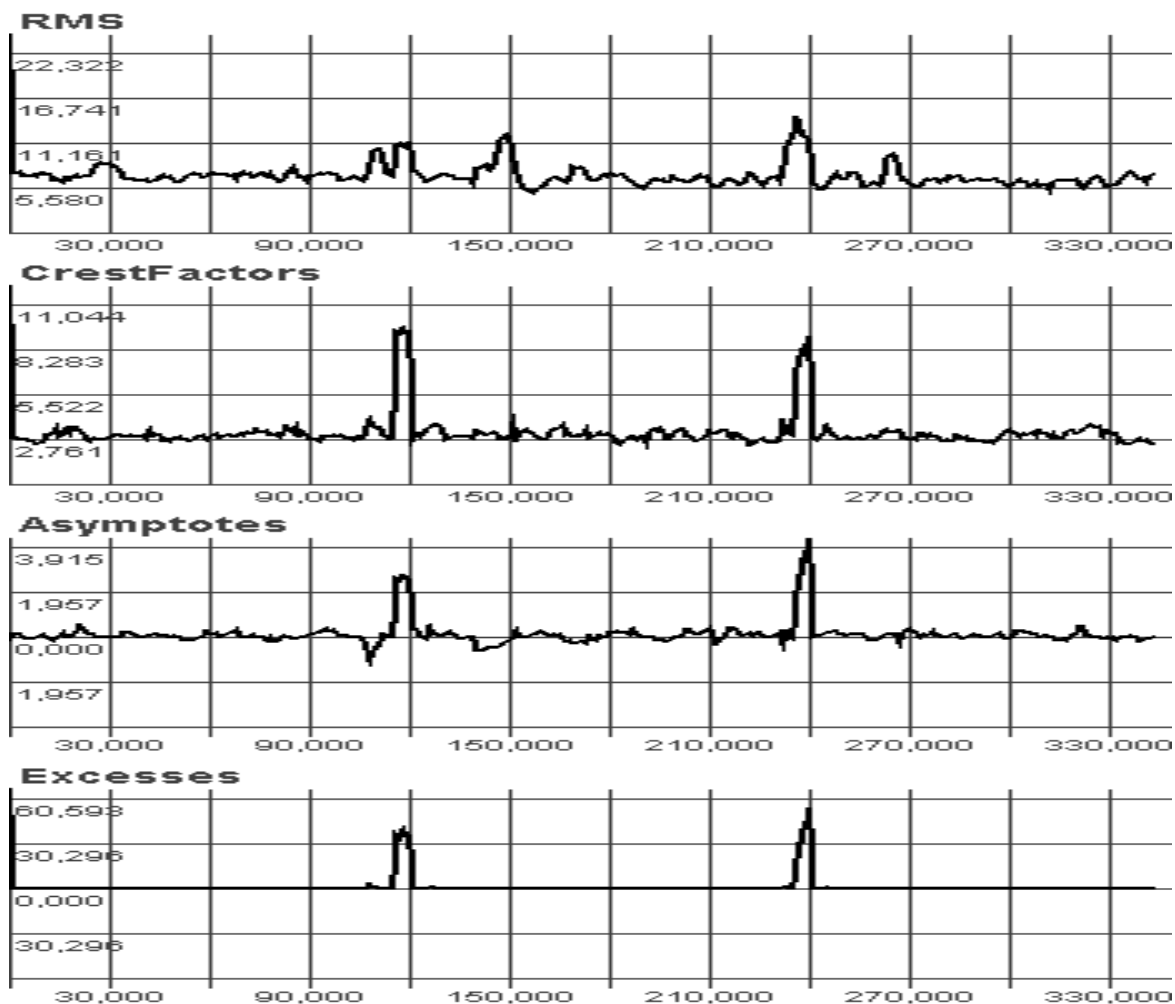


Рисунок 8. Временные тренды параметров сигнала ЭЭГ после ЭСТ.
Ось абсцисс – время, с. Ось ординат – значение параметра

Сравнивая временные тренды параметров одноименного сигнала ЭЭГ до и после процедур ЭСТ можно заметить, что вариативность параметров после ЭСТ значительно меньше, что свидетельствует о большей однородности исследуемого сигнала.

Представляет интерес изменение параметров сигнала в отдельных частотных полосах. Для выяснения данного вопроса можно сравнить изменение СКЗ сигналов ЭЭГ в частотных полосах альфа, бета, гамма, дельта, тэта (рис. 9-12).

На рис. 9-10 показаны полосовые спектры сигналов ЭЭГ до и после процедуры ЭСТ, а на рис. 11-12 временные тренды СКЗ сигналов ЭЭГ в частотных полосах альфа, бета, гамма, дельта, тэта

Заметно, что после ЭСТ произошло увеличение интенсивности сигнала в частотных полосах тэта и альфа относительно полосы дельта.

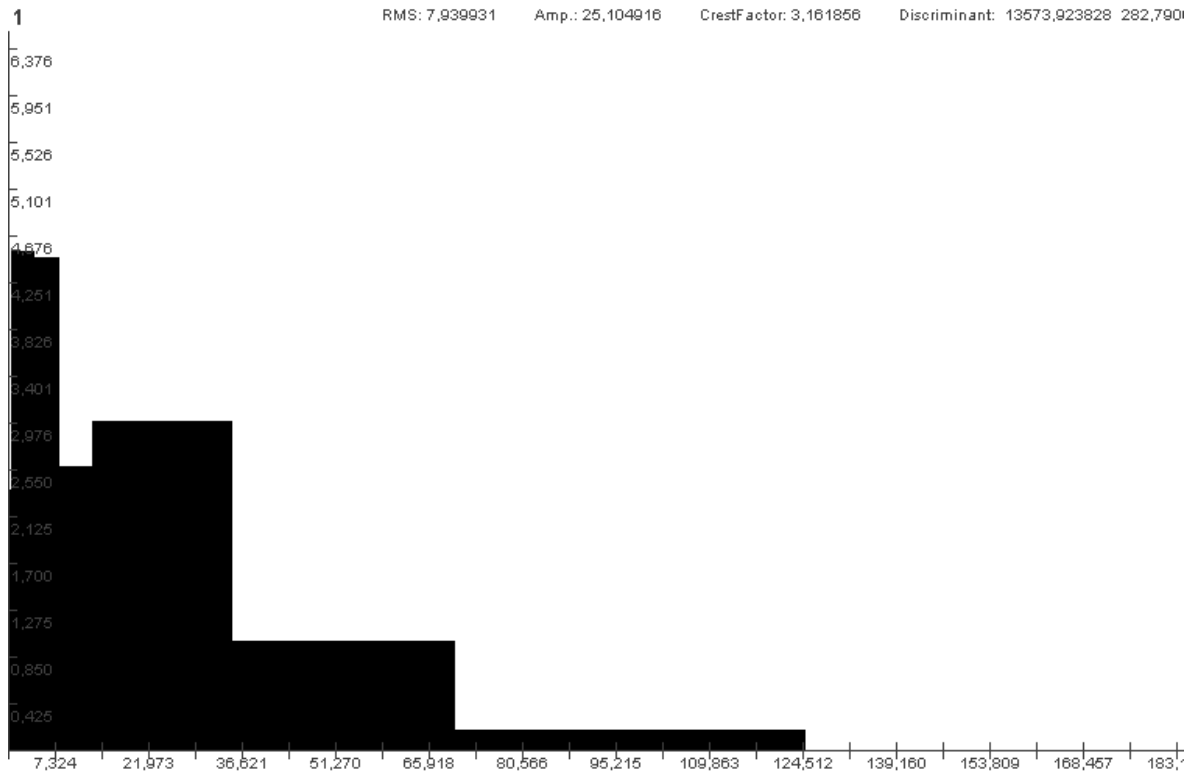


Рисунок 9. Полосовой амплитудный спектр сигнала ЭЭГ до ЭСТ

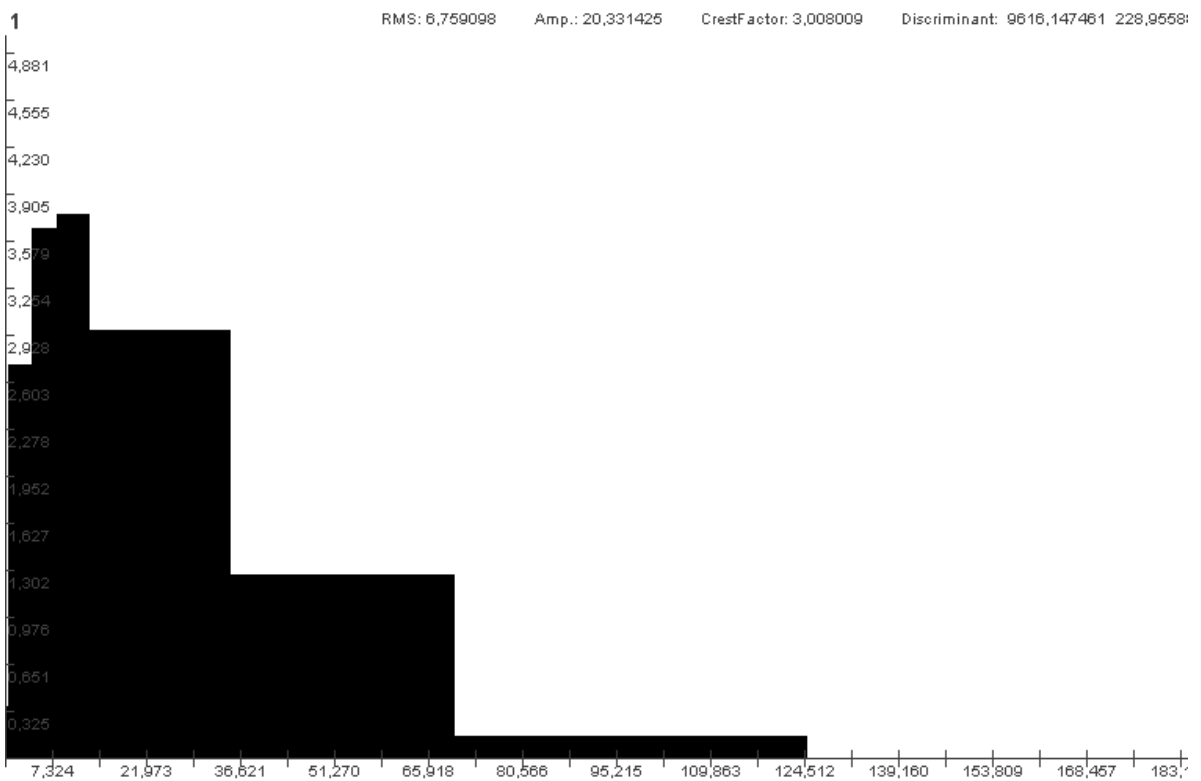


Рисунок 10. Полосовой амплитудный спектр сигнала ЭЭГ после ЭСТ

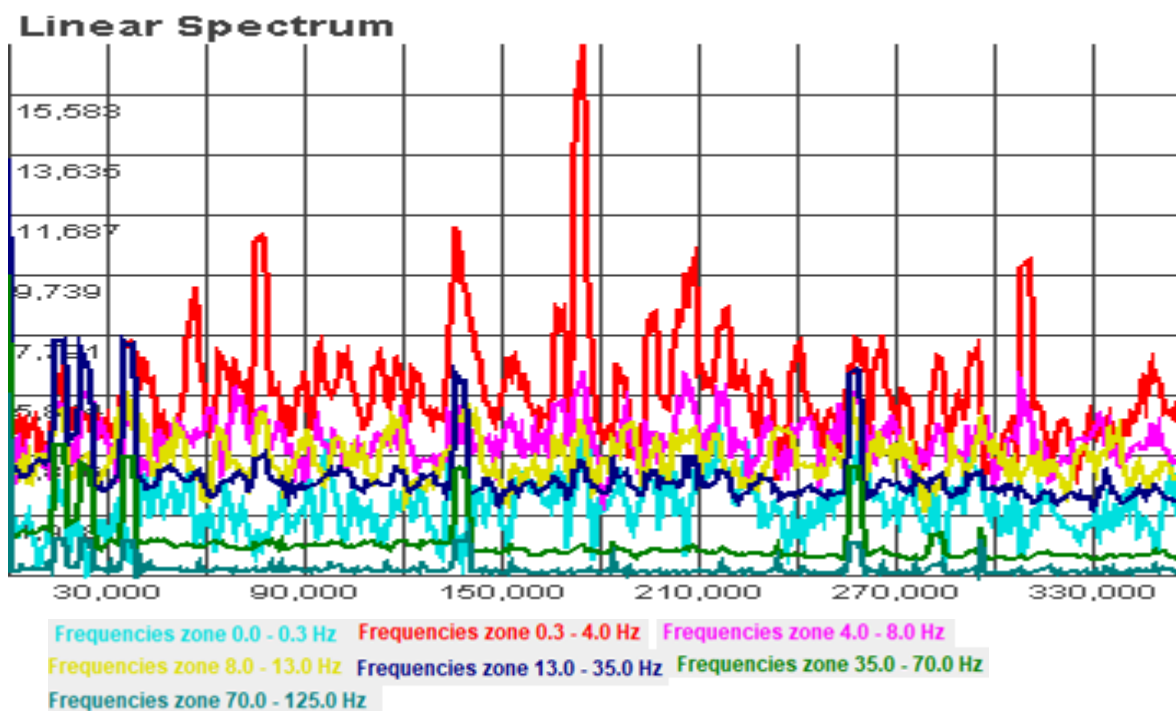


Рисунок 11. Временные тренды СКЗ в частотных полосах 0,3-4, 4-8, 8-13, 13-35, 35-70, 70-125 Гц сигнала ЭЭГ до ЭСТ

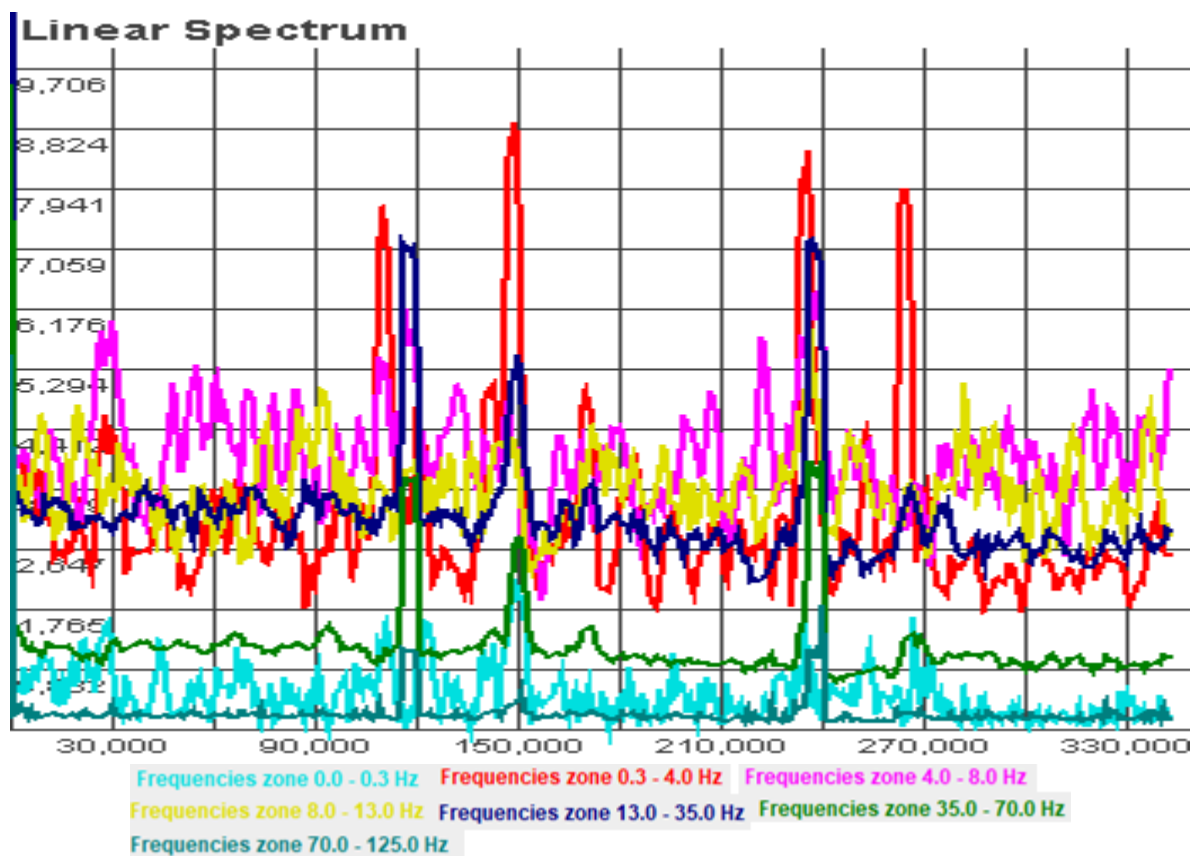


Рисунок 12. Временные тренды СКЗ в частотных полосах 0,3-4, 4-8, 8-13, 13-35, 35-70, 70-125 Гц сигнала ЭЭГ после ЭСТ

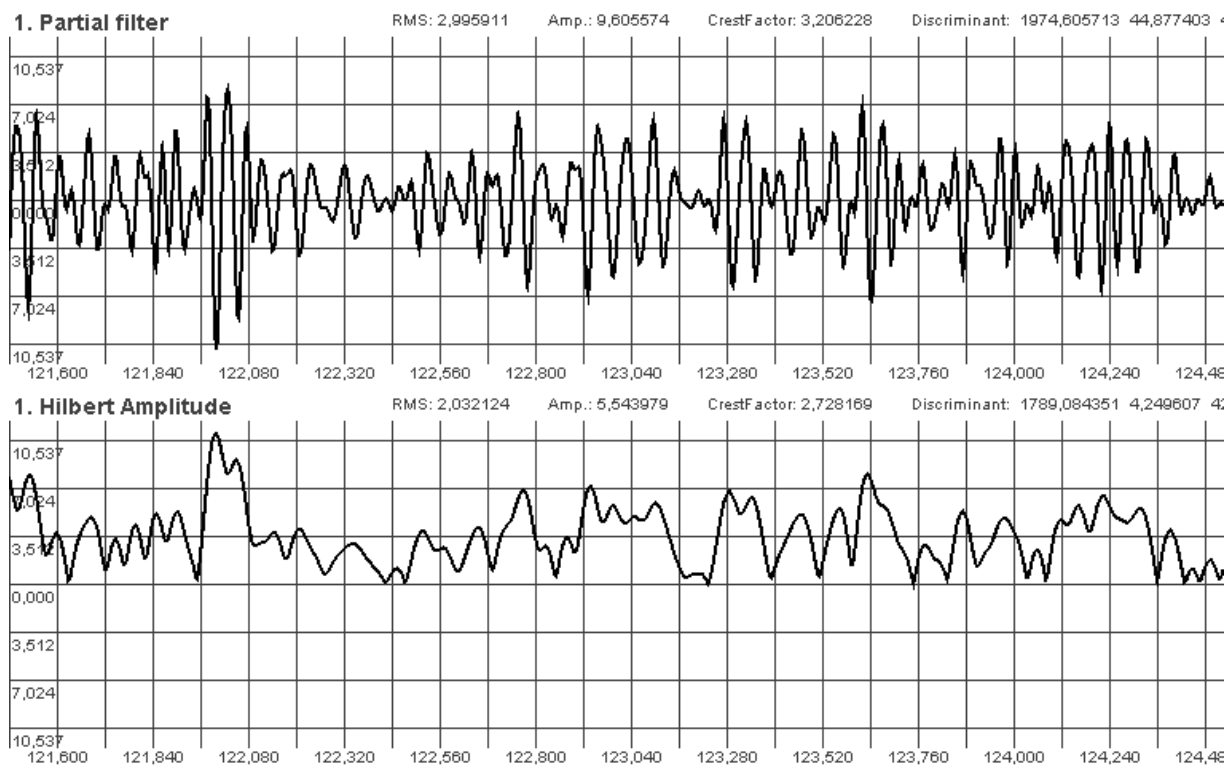


Рисунок 13. Временная реализация сигнала ЭЭГ в частотной полосе 13-35 Гц и её огибающая до ЭСТ

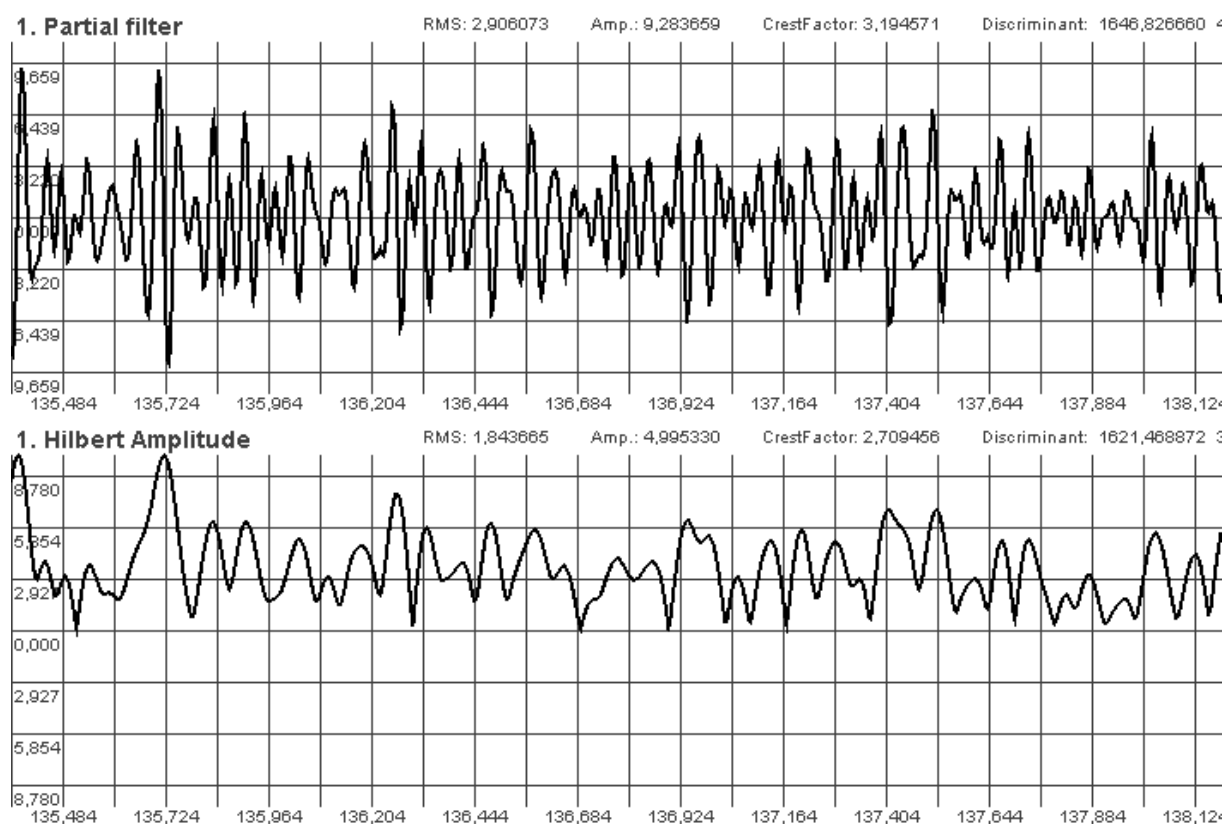
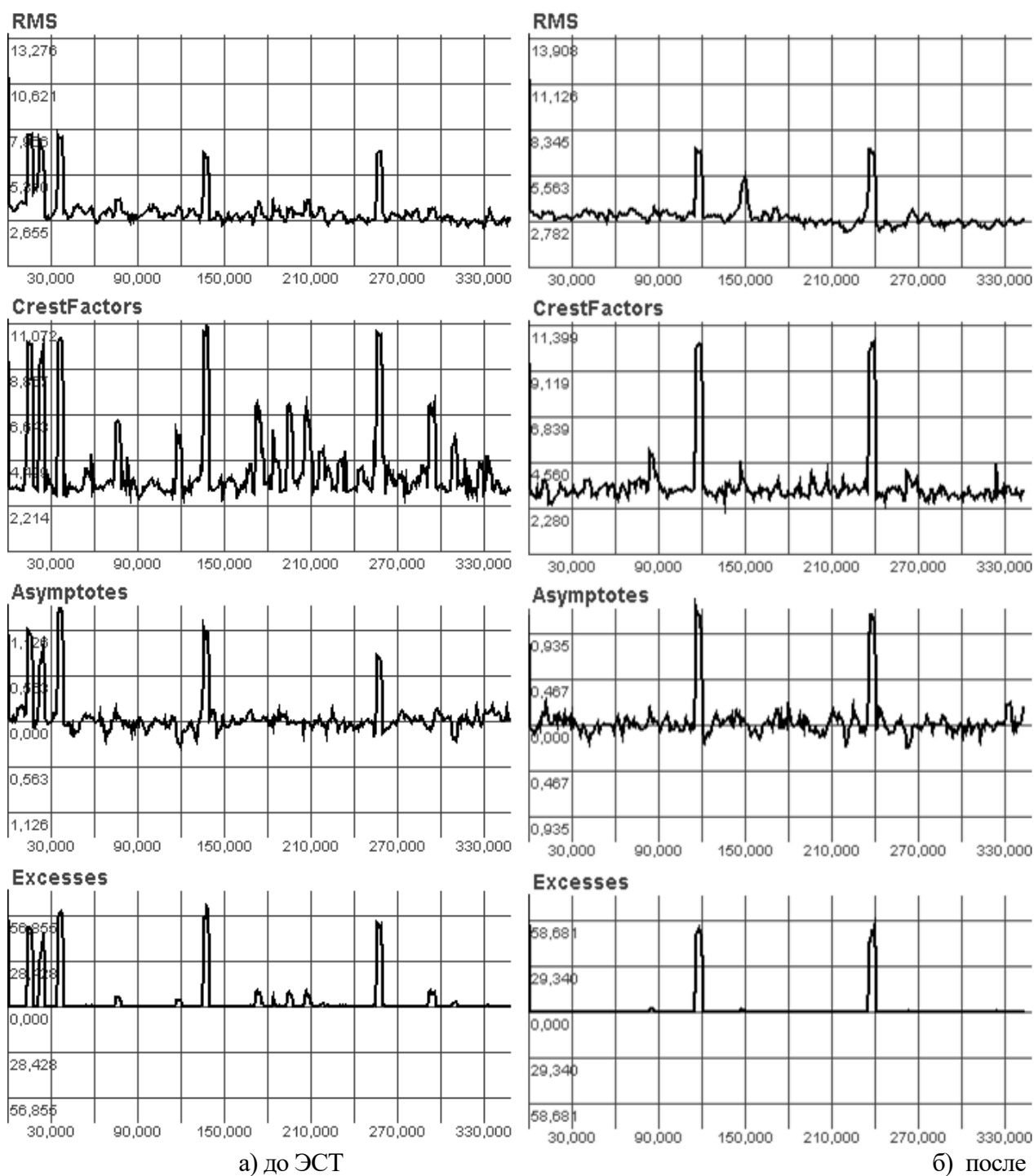


Рисунок 14. Временная реализация сигнала ЭЭГ в частотной полосе 13-35 Гц и её огибающая после ЭСТ



ЭСТ

Рисунок 15. Временные тренды параметров сигнала ЭЭГ в частотной полосе 13-35 Гц до и после ЭСТ. Ось абсцисс – время, с. Ось ординат – значение параметра

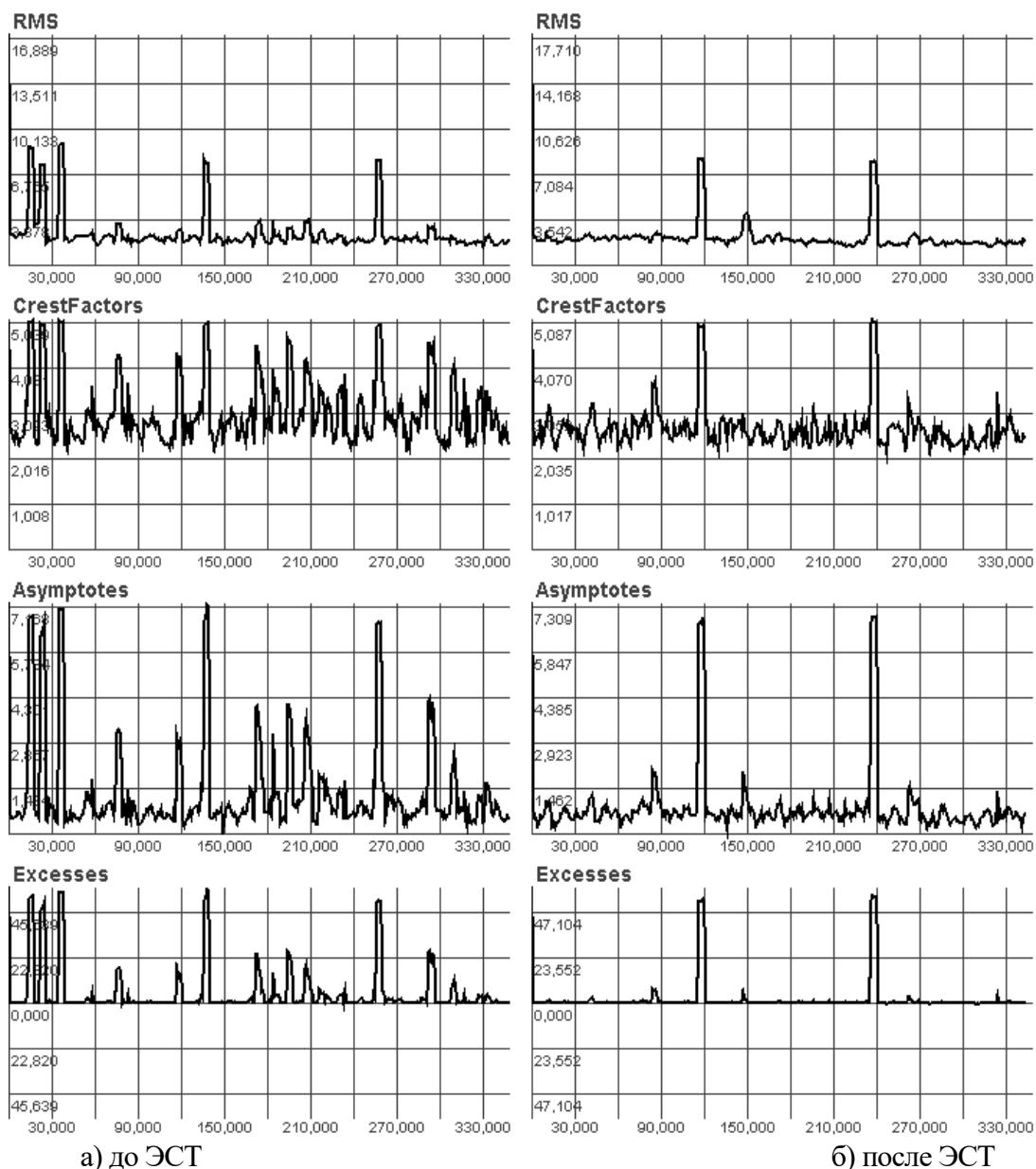


Рисунок 16. Временные тренды параметров огибающей сигнала ЭЭГ в частотной полосе 13-35 Гц до и после ЭСТ. Ось абсцисс – время, с. Ось ординат – значение параметра

Можно более детально исследовать параметры сигналов в отдельных частотных полосах. Предварительный анализ сигналов ЭЭГ показал, что достаточно существенные отличия сигналов ЭЭГ наблюдаются в частотной полосе 13-35 Гц. На рис. 13-14 представлены сигналы ЭЭГ в частотной полосе 13-35 Гц и их огибающие. Для этих сигналов вычислены временные тренды СКЗ, пик-фактора, эксцесса и асимптоты (рис. 15-16). Сравнение временных трендов параметров для сигналов ЭЭГ до процедуры ЭСТ и после показывает их существенное отличие.

Закключение.

По результатам исследования состояния нейронной сети головного мозга человека с использованием электроэнцефалограмм получают данные, которые впоследствии находят применение в медицинской практике. Представленные результаты обработки ЭЭГ позволяют сделать предположение о возможности определить группу информативно-значимых признаков, по которым можно будет делать обоснованное заключение об эффективности процедуры ЭСТ. Однако данная гипотеза требует подтверждения на большом объеме реальных данных. Также

следует учитывать, что рассмотрен только один из вариантов обработки сигналов ЭЭГ. Не менее интересные результаты могут быть получены с использованием вейвлетов или преобразования Гильберта-Хуанга [12].

Список литературы

- [1] Барков, А. В. Мониторинг и диагностика роторных машин по вибрации / А.В. Барков, А.Н. Баркова, А.Ю. Азовцев. – СПб.: Изд. центр СПбГМТУ, 2000. –
- [2] Осовский, С. Нейронные сети для обработки информации. Пер. с польского И. Д. Рудинского. – Москва: Финансы и статистика, 2002. – 344 с.
- [3] Павлова, Л. П. Доминанты деятельного мозга человека. Системный психофизиологический подход к анализу ЭЭГ / Л.П. Павлова. – СПб.: Информ-навигатор, 2017. – 430 с.
- [4] Зенков, Л. Р. Клиническая электроэнцефалография (с элементами эпилептологии) / Л. Р. Зенков. 2-е изд., испр. и доп. – М.: МЕДпресс-информ, 2002. – 368 с.
- [5] Докукина, Т.В. Визуальная и компьютерная ЭЭГ в клинической практике / Т. В. Докукина, Н. Н. Мисюк. – Минск: Книгзбор, 2011. – 187 с.
- [6] Татум, У.О. Клиническая интерпретация электроэнцефалографии / У.О. Татум. А.М. Хусейн, С.Р. Бенбадис, П.В. Каплан. Пер. с англ. – М.: Издательский дом БИНОМ. 2020. – 264 с.
- [7] Лайонс, Р. Цифровая обработка сигналов / Р. Лайонс. – М.: ООО «Бином-Пресс», 2006. – 656 с.
- [8] Айфичер, Э.С. Цифровая обработка сигналов: практический подход / Э.С. Айфичер, Б.У. Джервис. – М.: Издательский дом «Вильямс», 2008. – 992 с.
- [9] Кулаичев, А.П. Компьютерная электрофизиология и функциональная диагностика / А.П. Кулаичев. – Москва: ИНФРА-М, 2018. – 469 с.
- [10] Электроэнцефалография: руководство / М. В. Александров [и др.]. Под ред. М. В. Александрова. 3-е изд., перераб. и доп. – Санкт-Петербург: СпецЛит, 2020. – 224 с.
- [11] Бранцевич, П. Ю. Цифровая обработка вибрационных сигналов / П. Ю. Бранцевич. – Минск: Бестпринт, 2022. – 297 с.
- [12] Бранцевич, П. Ю. Примеры цифровой обработки электроэнцефалограмм / П. Ю. Бранцевич // Медэлектроника–2022. Средства медицинской электроники и новые медицинские технологии: сб. науч. Ст. XIII Междунар. Науч.-техн. конф. (Республика Беларусь, Минск, 8-9 декабря 2022 года). – Минск: БГУИР, 2022. – С. 314-318.
- [13] Метод лечения резистентных форм психических и поведенческих расстройств с использованием электросудорожной терапии. Инструкция по применению / Докукина Т.В. [и др.]. Утв. М-вом здравоохранения Респ. Беларусь 20.01.2015. – Минск: Министерство здравоохранения РБ, 2015. – 11 с.
- [14] Бауэр, М. Клинические рекомендации Всемирной федерации обществ биологической психиатрии по биологической терапии униполярных депрессивных расстройств. Часть 3: Острое и продолженное лечение униполярных депрессивных расстройств по состоянию на 2013 год / Бауэр М., Пфенниг А., Северус Э., Вайбрау П.С., Ж. Анст, Мюллер Х.-Ю. от имени и по поручению Рабочей группы по униполярным депрессивным расстройствам / Современная терапия психических расстройств. – 2016. – № 2. – С. 27-40.
- [15] Вентцель, Е.С. Теория вероятностей и ее инженерные приложения / Е.С. Вентцель, Л.А. Овчаров. – Москва: Наука. Гл. ред. Физ.-мат. лит., 1988. – 480 с.

COMPARATIVE ANALYSIS OF ELECTROENCEPHALOGRAMS

P.J. Brancevich
Grand PhD courses

Educational institution "Belarusian State University of Informatics and Radioelectronics", Republic of Belarus
E-mail: branc@bsuir.edu.by

Abstract. Medical and technical diagnostics, as a set of rules, methods, algorithms that allow you to make a decision about the state of the observed object or subject, have much in common. The creation of new and improvement of existing methods for assessing the human condition is an important area of medical research. Of undoubted scientific interest is the study of the neural network of the human brain using electroencephalograms, which are complex digital signals. Currently, the method of electroconvulsive therapy is used to treat resistant forms of mental and behavioral disorders. Relevant is the determination of informatively significant parameters of the EEG, which would indicate its change according to the results of ECT. The results of some methods of digital processing of temporal realizations of electroencephalograms used for comparative analysis of EEG obtained before and after ECT sessions are presented.

Keywords: signal, electroencephalogram, parameter, characteristic, digital processing, solution

УДК 331.108:004.42

ИННОВАЦИОННЫЕ ТРЕНДЫ В СФЕРЕ ОБРАЗОВАНИЯ И РОЛЬ ПРЕПОДАВАТЕЛЯ



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. Инновационные решения в сфере образования придадут новый импульс развитию современного общества. Доказательное образование, микрообучение, генеративный искусственный интеллект, технологии виртуальной и дополненной реальности станут неотъемлемой частью не только образования, но частью жизни многих, даже очень далеких от ИТ сферы людей. Однако, как бы далеко ни продвинулись бы технологии, процесс обучения всегда будет во многом зависеть от взаимоотношений между учителем и учеником. Например, искусственным интеллект может адаптировать образовательный контент под стиль обучения конкретного ученика, но именно преподаватель сможет это использовать, чтобы лучше понимать, что мотивирует его студента, поддерживает интерес к учебе.

Ключевые слова: EdTech-технологии, доказательное образование, искусственный интеллект, технологии виртуальной и дополненной реальности, микрообучение.

Введение.

Образование является одной из крупнейших отраслей в мире, на которую приходится более 6% мирового ВВП. Объем мирового рынка образования в 2021 году, по данным из открытых источников, достиг отметки 6,5 трлн долл. США, а объем рынка EdTech-отрасли 270 млрд долл. США – это, по максимальным экспертным оценкам, всего лишь чуть более 4% от всего объема рынка образования, по итогам 2020 года было порядка 3,6%.

EdTech-технологии – это инновационные решения в сфере образования. Они охватывают разработку и внедрение оборудования и устройств (планшетов, школьных досок), разработку программных комплексов (авиационных, медицинских и других специализированных 3D-симуляторов).

Огромный процент EdTech – это интернет-образование, в которое комплексно входят не только интерактивные IT-сервисы, но и информационные платформы, которые еще называют инфобизнесом. Ожидается, что к 2025 году объем международного EdTech-рынка вырастет в диапазоне от 404 до 434 млрд долл. США, а по итогам 2030 года доля международного рынка онлайн-образования должна вырасти до 5,6% (585 млрд долл. США) [1].

ТОП-33 мировых «EdTech-единорогов» в 2022 году по направлениям деятельности можно разделить на три основных сегмента: общее и дополнительное школьное образование (36%), высшее образование (24%) и дополнительное профессиональное образование (27%), направленное на дальнейшее карьерное развитие (в большей степени) или смену

профессиональной деятельности (в основном, в сторону развития различных цифровых компетенций) [1].

Даже с учетом максимально ожидаемого 12 % роста объемов мирового EdTech-рынка до 300 млрд долл. США в 2022 году, доля рынка онлайн-образования по итогам 2022 года, скорее всего, не превысит 5,5% от общего объема рынка, но продолжает неуклонно расти. По данным «Центр демографии и глобального человеческого капитала имени Витгенштейна» (IASA, VID/ÖAW, WU), выпускников школ, колледжей, университетов и альтернативных высших учебных заведений в мире к 2050 году станет на 2 млрд больше, чем в 2020 году.

По данным «HolonIQ», за двадцать один год, начиная с февраля 2000 года, количество публичных образовательных компаний из 33 стран мира увеличилось с 41-й в 2000 году, до 224-х в феврале 2020 года, до 268-ми по итогам 2020 года и до 307 по итогам 2021 года [2].

Основными технологическими трендами в этой сфере по-прежнему остаются:

1. Mobile learning и кросс-платформенность становятся базовыми требованиями к любой обучающей платформе

2. Носимые устройства будут играть большую роль в обучении

3. Big Data необходима для улучшения технологий обучения

К тематическим трендам можно отнести:

1. Курсы со знаменитыми людьми (актерами, спортсменами, режиссерами)

2. Подбор коучей и репетиторов

3. Мобильный тьютор в определенной сфере

4. Курсы для творческих профессий

5. Компании учат услугам, косвенно связанным с продуктом - повышают product awareness

6. Узкая специализация – весь сайт про одну тему

К методологическим трендам относятся сближение рекрутмента и обучения, Serious Games – геймификация в B2B-сегменте, обучение в игровой форме, обучение через наблюдение и копирование

Рассмотрим некоторые тренды в образовании.

1. Доказательное образование.

Термин «доказательное образование» происходит от английского evidence-based education, которое дословно переводится как «образование, основанное на доказательствах». Что и кому нужно доказывать? Если кратко – всем и все, но в первую очередь себе. Выражаясь научным языком, доказательное образование – это такой подход ко всем составляющим процесса образования, в котором используются методы, основанные на надежных доказательствах, полученных экспериментальным путем. И это касается всего, начиная от разработки образовательной политики и заканчивая каждой ежедневной учебной практикой.

Такой подход считается противоположным традиционному образованию, где много чего основано на субъективном восприятии педагога: подача учебного материала, соотношение объяснения учителя и самостоятельной работы учеников на занятиях, проверка знаний, выставление отметок и т.д.

В данном подходе роль учителя (преподавателя) заключается в умении оценивать достижения ученика (студента) так, чтобы стимулировать новые достижения (так называемый эффект зеленых чернил). Кроме этого командная работа педагога и ученика создает синергетический эффект. Самооценка оценки – когда ученикам предлагают самим оценить уровень своих познаний.

2. Искусственный интеллект.

Количество запросов на «нейрообразование» в Интернете увеличилось на 405% за последние 10 лет [3].

Системы искусственного интеллекта тестируют обучающихся, отслеживают прогресс обучения и уровень вовлеченности, рекомендуют сопутствующие игры и симуляторы,

обучающие видео и другие полезные ресурсы, выстраивают систему персональных рекомендаций.

Платформы онлайн-оценки на основе AI, например, «iSpringSuite», «Nearpod», «Glider.ai» и «ProProfs», помогают преподавателям и потенциальным работодателям производить оценку уровня знаний и компетенций с помощью анкет, тестирования и онлайн-викторин. Так программа TutorAI позволяет автоматически создавать целые образовательные курсы с помощью ИИ. В поиск вводится тема, а искусственный интеллект самостоятельно собирает и систематизирует информацию, создает модули, тестирует и дает ссылки на дополнительные источники. Это, безусловно, большая помощь как преподавателям, так и студентам, занимающимся научной деятельностью.

В то же время, искусственный интеллект используется учебными заведениями не только для анализа действий пользователя, его предпочтений, успеваемости и компетенций, построения системы рекомендаций и индивидуальной траектории обучения, но и для оптимизации маркетинговой и операционной деятельности, включая набор и зачисление студентов. Необходимо обратить внимание на образовательный процесс лиц с особыми потребностями имеет ряд особенностей, связанных с конкретными особенностями каждого человека.

Так, например, только лица с нарушениями слуха имеют разные потребности в образовательном контенте: люди, потерявшие слух во время жизни, могут воспринимать субтитры видеоматериала учебного контента, в то время как для людей глухих от рождения человеческая речь является непонятной и для восприятия видеоматериала учебного контента для них необходимо предусмотреть сурдоперевод. Поэтому использование искусственного интеллекта для создания оригинального контента для людей с ограниченными возможностями имеет огромные перспективы как в социальном плане, так и в экономическом.

Есть конечно и опасения, что ИИ полностью вытеснит преподавателя из аудитории, заменив его искусственно созданным шаблоном с его мимикой и даже голосом. Но сможет ли он удержать внимание учеников и студентов на протяжении часа и как будет реагировать на их реплики и посторонние шумы?

3. Технологии виртуальной и дополненной реальности.

VR и AR также все чаще используются в образовании. На мировом рынке EdTech уже есть успешные примеры использования технологии VR и AR. Например, Oxford Medical Simulation позволяет будущим врачам отрабатывать свои навыки в виртуальной реальности. Платформа уже получила 2,1 млн. фунтов стерлингов на разработку новых программ и расширение своего присутствия на рынке США. Согласно отчету Market Research Future (MRFR), рынок AR и VR в образовании будет расти в среднем на 18,2% в период с 2022 по 2027 год. А в Lenovo отмечают, что к 2025 году образование станет 4-м по величине сектором использования технологий дополненной реальности и будет иметь объем \$700 млн.[3].

Однако, процессом использования этой технологии должен кто-то руководить. Кто как ни преподаватель-практик должен обучить студентов-медиков всем правилам использования технологии виртуальной и дополненной реальности на практике?

4. Микрообучение.

Это обучение небольшому объему знаний или навыков за отрезок времени в пределах 5 минут. Изначально такой формат был опробован для обучения в условиях ограниченного времени, отведенного на усвоение знаний. В первую очередь это касалось получения профессиональных навыков и адаптации на конкретном рабочем месте. Позже оказалось, что метод вполне приемлем для обучения отстающих студентов и школьников, которым тяжело долго удерживать внимание на предмете изучения. Более того, есть мнение, что растущая популярность формата объясняется все снижающейся способностью к концентрации внимания у детей и молодежи. Все чаще принцип микро- и нанообучения адаптируется создателями всевозможных образовательных курсов, что позволяет их клиентам вписывать обучение даже в самые плотные рабочие графики. Формат микро- и нано-уроков обычно представляет собой

чередование видео-, аудио- и текстового форматов. Изучение одного урока должно занимать не более 5 минут.

Формат микрообучения может быть построен исключительно на EdTech и обычно направлен на достижение одной образовательной цели. Специалисты отмечают, что микро- и микрообучение позволяют более эффективно запоминать материал, а также делают возможным обучение в условиях сниженной концентрации внимания современного человека. Этот формат обучения не сможет полностью заменить традиционные, более длительные формы обучения, но уже сейчас мы можем наблюдать появление все большего количества программ, предлагающих краткие учебные материалы. Например, один из самых известных EdTech-стартапов с белорусскими корнями – EduDo. И здесь в сочетании с искусственным интеллектом можно адаптировать образовательный контент под стиль обучения конкретного ученика, студента, а преподавателю лучше понимать, что его мотивирует, поддерживает интерес к учебе.

Заключение.

Таким образом, можно сделать вывод, что инновационные решения в сфере образования дадут новый импульс развитию современного общества.

Доказательное образование, микрообучение, генеративный искусственный интеллект, технологии виртуальной и дополненной реальности станут неотъемлемой частью не только образования, но частью жизни многих, даже очень далеких от ИТ сферы людей.

В ближайшем будущем будут возникать платформы по трансформации запроса пользователя в код, текст, картинку, видео, музыку. Одним нажатием кнопки мы сможем получить игру или скрипт рекламной кампании, которые будут сгенерированы искусственным интеллектом. Уже сейчас можно преобразовать текст, лекцию в видеоформат, наложить фото и голос преподавателя, но это не исключает его участия в процессе. И это подтверждает мнение Билл Гейтса по поводу использования искусственного интеллекта в образовании. Он считает, что как бы далеко ни продвинулись бы технологии, процесс обучения всегда будет во многом зависеть от взаимоотношений между учителем и учеником.

Список литературы

- [1] Мировой рынок онлайн-образования (аналитический обзор: история, тенденции, перспективы, прогнозы). [Электронный ресурс]
- [2] Искусственный интеллект в образовании. [Электронный ресурс]
- [3] 10 трендов EdTech образования. [Электронный ресурс]

INNOVATIVE TRENDS IN EDUCATION AND THE ROLE OF THE TEACHER

O.N. Shkor

Senior Lecturer

at the Department of Economics BSUIR

*Senior Lecturer at the Department of Economics BSUIR Belarusian State University of Informatics and Radioelectronics G. Minsk, Republic of Belarus, Senior Lecturer at the Department of Economics
E-mail: shkor@bsuir.by*

Abstract. Innovative solutions in the field of education will give a new impetus to the development of modern society. Evidence-based education, microlearning, generative artificial intelligence, virtual and augmented reality technologies will become an integral part of not only education, but also a part of the lives of many people, even those who are very far from the IT sphere. However, no matter how far technology advances, the learning process will always depend largely on the relationship between teacher and student. For example, artificial intelligence can adapt educational content to the learning style of a particular student, but it is the teacher who can use this to better understand what motivates his student and maintains interest in learning.

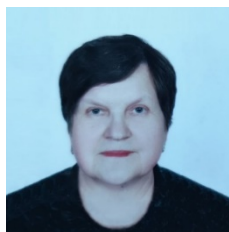
Keywords: EdTech technologies, evidence-based education, artificial intelligence, virtual and augmented reality technologies, microlearning.

УДК 611.08

ИССЛЕДОВАНИЕ ВЛИЯНИЯ ПРИМЕНЕНИЯ ФИЛЬТРОВ НА СИГНАЛ ЭЛЕКТРОКАРДИОГРАММЫ



А.В. Гордиевич
ОАО «АГАТ-системы
управления» - управляющая
компания холдинга
«Геоинформационные
системы управления»,
инженер-программист
alinagordievich78@gmail.com



Л.Ф. Васковская
ОАО «АГАТ-системы
управления» - управляющая
компания холдинга
«Геоинформационные
системы управления»,
ведущий системный
аналитик
lidavask@yandex.by



П.В. Камлач
Заместитель декана ФКП,
доцент кафедры
электронной техники и
технологии БГУИР,
кандидат технических наук,
доцент
kamlachpv@bsuir.by

А.В. Гордиевич

Окончила Белорусский государственный университет информатики и радиоэлектроники. Область научных интересов связана с разработкой методов и алгоритмов автоматизированных систем сбора и обработки информации для контроля за состоянием здоровья.

Л.Ф. Васковская

Окончила Минский радиотехнический институт. Область научных интересов связана с разработкой автоматизированных систем сбора и обработки информации о воздушной, наземной и надводной обстановках.

П.В. Камлач

Окончил Белорусский государственный университет информатики и радиоэлектроники. Область научных интересов связана с разработкой и моделированием работы систем медицинского назначения.

Аннотация. Выполнено сравнение эффективности фильтров ЭКГ в системе MatLab для обработки сигналов ЭКГ. Проведен анализ фильтров Савицкого-Голея, Чебышева, и комбинации фильтров Савицкого-Голея и Чебышева по критериям эффективности сглаживания, удаления шума и сохранения морфологии ЭКГ сигнала.

Результаты сравнения показали, что фильтр, основанный на комбинации фильтров Савицкого-Голея и Чебышева, обеспечил наилучшую эффективность по всем критериям. Он позволяет удалить шумы и артефакты без искажения морфологии сигнала и сохранения высокой точности результатов. Предложенный подход может быть применен в широком спектре задач – от анализа ЭКГ сигналов до мониторинга сердечной деятельности.

Ключевые слова: ЭКГ, обработка сигналов ЭКГ, эффективность фильтров ЭКГ, фильтр Савицкого-Голея, фильтр Чебышева.

Введение.

Одним из важных методов диагностики заболеваний сердечно-сосудистой системы является электрокардиография.

Этот метод позволяет выявить признаки коронарной недостаточности, нарушений ритма сердца и проводимости, гипертрофии различных отделов сердца.

Электрокардиография (ЭКГ) – метод графической регистрации электрических явлений, возникающих в сердечной мышце при ее возбуждении. ЭКГ позволяет изучить автоматизм, проводимость, возбудимость и рефрактерность сердечной мышцы.

Актуальность.

Сигнал ЭКГ очень чувствителен к помехам, которые могут возникать в процессе исследования, такие как движения пациента, электромагнитные помехи, мускульный шум и т.д.

Фильтрация сигнала ЭКГ позволяет удалить эти помехи, что позволит более точно определить наличие или отсутствие патологий сердца.

Фильтрация сигнала ЭКГ позволяет выделить некоторые особенности сигнала, такие как амплитуда, частота и продолжительность, которые могут быть связаны с различными патологиями сердца. Без фильтрации, получение точной информации из электрокардосигнала (ЭКС) может быть затруднительным, что может повлиять на диагностику и лечение пациентов.

Кроме того, фильтрация ЭКГ имеет особую актуальность в условиях современной медицины, где пациентам проводят длительные мониторинги ЭКГ и выявляют некоторые заболевания, которые можно диагностировать только по длительному анализу сигнала.

Таким образом, фильтрация ЭКГ является необходимым элементом в диагностике не только частных пациентов, но и в широких исследованиях сердечно-сосудистой системы в целом, что подчеркивает ее актуальность.

Моделирование идеального ЭКГ сигнала.

Для исследования влияния фильтров на ЭКГ сигнал в системе MatLab был построен идеальный ЭКГ сигнал (рисунок 1).

Для моделирования была использована встроенная функция $ecg(n)$, которая строит один кардиоцикл длиной n точек [2].

Для ЭКГ сигнала задали следующие входные параметры:

1. Частота дискретизации сигналов: $F_d = 1000$ Гц;
2. Амплитуда кардосигнала: $A_{ecg} = 200$ мВ;
3. Частота сердечных сокращений: $F_{ecg} = 60$ уд/мин;
4. Длительность требуемого сигнала: $T_{sign} = 6$ секунд;

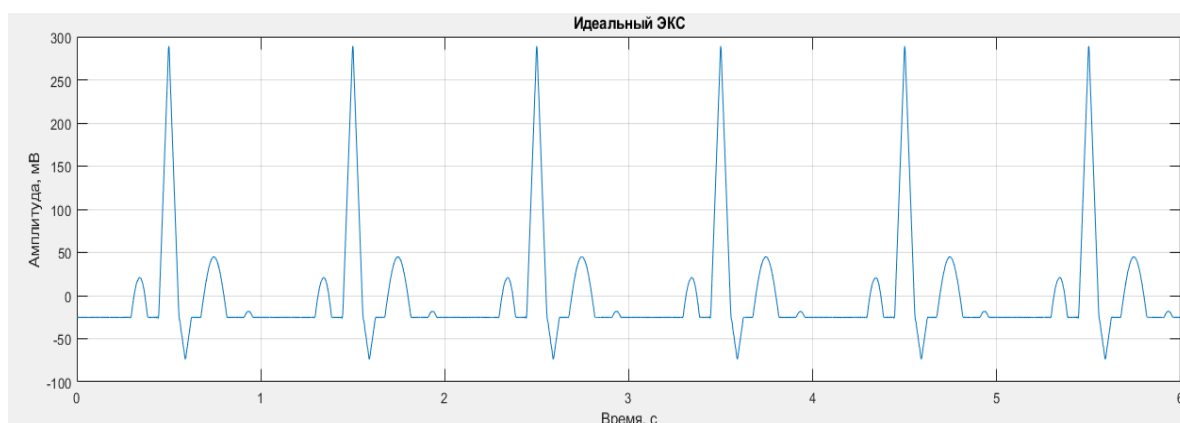


Рисунок 1. Идеальный ЭКС в системе MatLab

Затем к идеальному ЭКС добавили белый и гармонический шумы, со следующими входными параметрами:

1. Амплитуда сигнала белого шума: $A_{Wn} = 50$ мВ;
2. Амплитуда сигнала гармонического шума: $A_{sin} = 20$ мВ;
3. Частота дискретизации сигналов: $F_d = 1000$ Гц;
4. Частота синусоидального сигнала: $F_{sin} = 150$ Гц;

Построенный ЭКС с помехами представлен на рисунке 2.

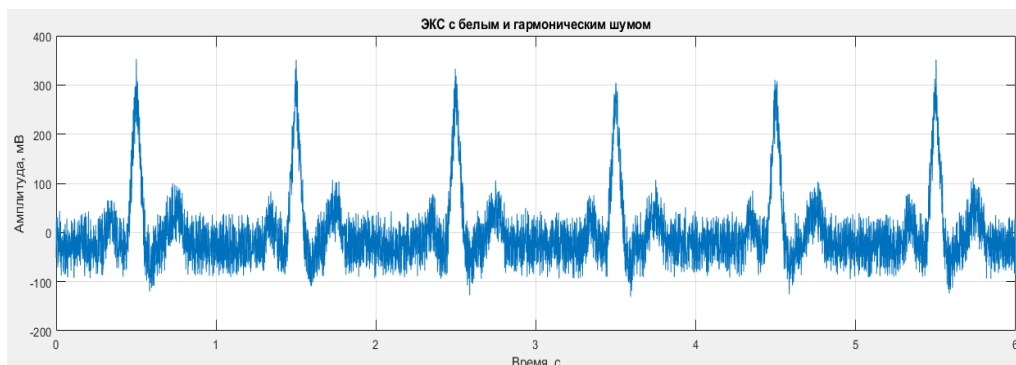


Рисунок 2. ЭКС с белым и гармоническим шумом в системе MatLab

Фильтр Савицкого-Голея.

Фильтр Савицкого-Голея [3, 4] имеет входной параметр – размер окна фильтра.

В первую очередь производились исследования по изменению среднеквадратического отклонения и коэффициента подавления случайной помехи от размера окна фильтра для 10 ЭКС с фиксированными уровнями помех на входе фильтра.

Результаты исследования показали, что наибольшее подавление помех происходит при размере окна фильтра Савицкого-Голея 13 точек.

Так как большинство фильтров не обеспечивают сохранение формы полезного сигнала на локальных участках (зубцах, комплексах и сегментах), исследования проводились в следующих направлениях:

1. Нахождение среднеквадратической ошибки отклонения, полученной по всему исследуемому фрагменту ЭКС, включающему 6 кардиоциклов;
2. Нахождение среднеквадратической ошибки отклонения на критичных для диагностики участках:
 - a. Зубцы P, Q, R, S, T, U;
 - b. Интервалы PQ, QT, ST.
 - c. Сегменты PQ, ST;
 - d. Комплекс QRS;

Все расчеты проводились в системе MatLab. Результаты расчетов представлены в таблице 1.

Таблица 1. Результаты исследования изменения среднеквадратической ошибки отклонения на разных компонентах электрокардиограммы после применения фильтра Савицкого-Голея

Компоненты электрокардиограммы		Среднеквадратическая ошибка отклонения, %
Зубец	P	107,18
	Q	46,12
	R	41,56
	S	59,68
	T	47,45
	U	34,71
Сегмент	PQ	45,36
	ST	43,15
Интервал	PQ	47,14
	QT	43,44
	ST	46,76
Комплекс	QRS	43,25
Весь участок (6 кардиоциклов)		43,08

На рисунке 3 представлен увеличенный (отмасштабированный) ЭКС после применения фильтра Савицкого-Голея.

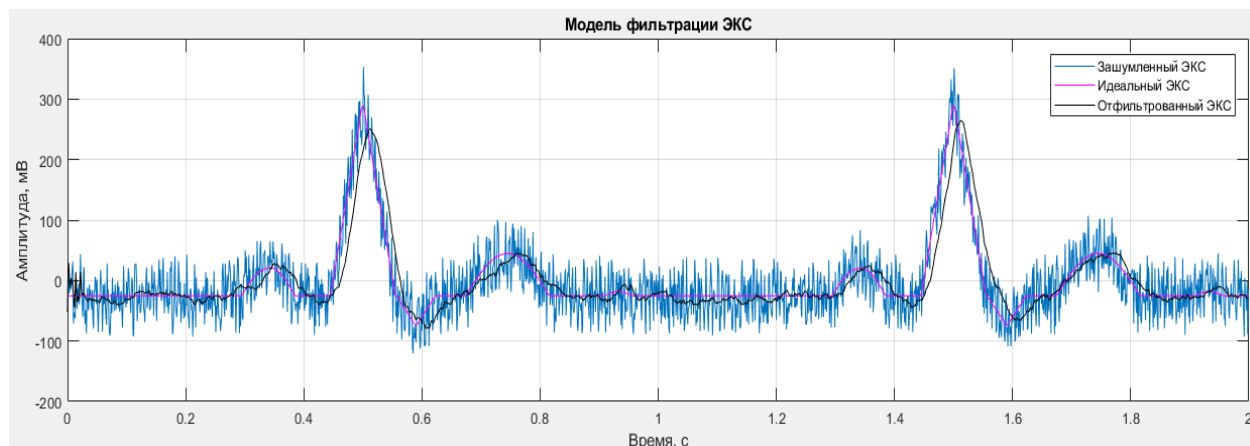


Рисунок 3. Увеличенный ЭКС после применения фильтра Савицкого-Голея

Использование фильтра Савицкого-Голея позволяет снизить долю высокочастотных составляющих ЭКС. Данный фильтр справляется с задачей фильтрации шумов ЭКС на 56,92%, максимальное значение СКО отклонения сигнала наблюдалось в зубце Р – 107,18% и минимальное значение в зубце U – 34,71%. На рисунке 4.5 видно, что фильтр Савицкого-Голея дает смещение по оси времени и полученный отфильтрованный ЭКС труден для восприятия, т.к. подавляет помехи не в полной степени.

Фильтр Чебышева.

Фильтр Чебышева [5, 6] также был исследован в нескольких направлениях: нахождение среднеквадратической ошибки отклонения, полученной по всему исследуемому фрагменту ЭКС, включающему 6 кардиоциклов и нахождение среднеквадратической ошибки отклонения на критичных для диагностики участках (зубцы, интервалы, сегменты, комплекс).

Все расчеты проводились в системе MatLab. Результаты расчетов представлены в таблице 2.

Таблица 2. Результаты исследования изменения среднеквадратической ошибки отклонения на разных компонентах электрокардиограммы после применения фильтра Чебышева

Компоненты электрокардиограммы		Среднеквадратическая ошибка отклонения, %
Зубец	P	31,36
	Q	10,67
	R	9,94
	S	14,64
	T	19,92
	U	27,45
Сегмент	PQ	12,53
	ST	13,49
Интервал	PQ	13,07
	QT	11,03
	ST	18,85
Комплекс	QRS	12,47
Весь участок (6 кардиоциклов)		15,68

На рисунке 4 представлен увеличенный (отмасштабированный) ЭКС после применения фильтра Чебышева.

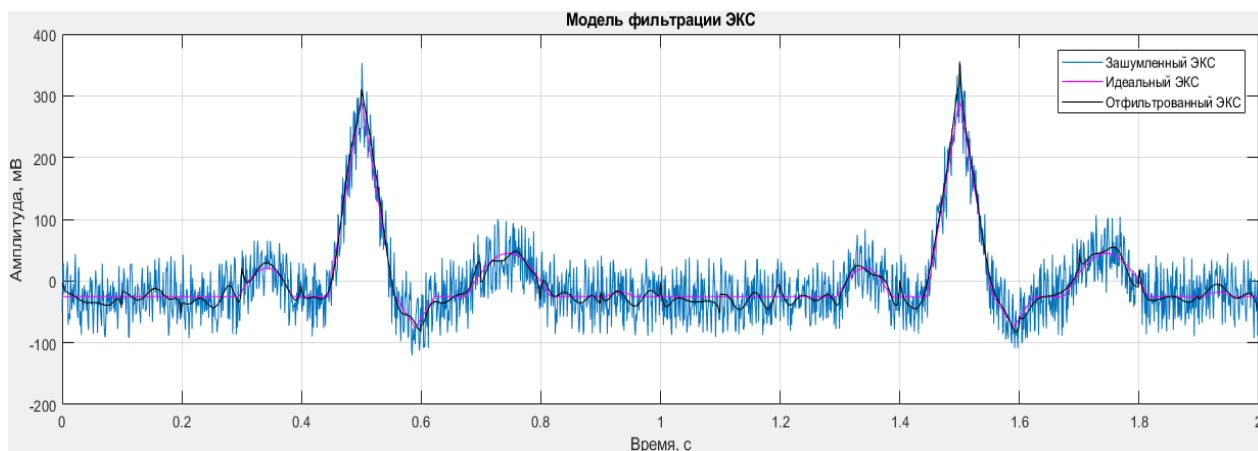


Рисунок 4. Увеличенный ЭКС после применения фильтра Чебышева

Использование фильтра Чебышева позволяет снизить долю высокочастотных составляющих ЭКС, при этом увеличивая амплитуду зубцов, что может повлиять на результат постановки диагноза. Фильтр Чебышева подавил шумы на 84,32%, при этом максимальное значение СКО отклонения сигнала наблюдалось в зубце Р – 31,36% и минимальное значение в зубце R – 9,94%.

Комбинация фильтров Савицкого-Голея и Чебышева.

Разработанный фильтр имеет входной параметр – размер окна фильтра Савицкого-Голея. В первую очередь производились исследования по изменению среднеквадратического отклонения и коэффициента подавления случайной помехи от размера окна фильтра Савицкого-Голея, для 10 ЭКС с фиксированными уровнями помех на входе фильтра. Результаты исследования показали, что наибольшее подавление помех происходит при размере окна фильтра Савицкого-Голея 19 точек.

Так как большинство фильтров не обеспечивают сохранение формы полезного сигнала на локальных участках (зубцах, комплексах и сегментах), исследования проводились в следующих направлениях:

1. Нахождение среднеквадратической ошибки отклонения, полученной по всему исследуемому фрагменту ЭКС, включающему 6 кардиоциклов;

2. Нахождение среднеквадратической ошибки отклонения на критичных для диагностики участках:

- a. Зубцы P, Q, R, S, T, U;
- b. Интервалы PQ, QT, ST.
- c. Сегменты PQ, ST;
- d. Комплекс QRS;

Все расчеты проводились в системе MatLab.

Результаты расчетов представлены в таблице 3.

Таблица 3. Результаты исследования изменения среднеквадратической ошибки отклонения на разных компонентах электрокардиограммы после фильтрации ЭКС методом комбинации фильтров Савицкого-Голея и Чебышева

Компоненты электрокардиограммы		Среднеквадратическая ошибка отклонения, %
Зубец	P	13,71
	Q	10,17
	R	7,27
	S	13,05
	T	12,91
	U	12,55
Сегмент	PQ	11,82
	ST	13,08
Интервал	PQ	12,05
	QT	9,08
	ST	16,76
Комплекс	QRS	8,39
Весь участок (6 кардиоциклов)		9,07

На рисунке 5 представлен увеличенный (отмасштабированный) ЭКС после применения метода комбинации фильтров Савицкого-Голея и Чебышева.

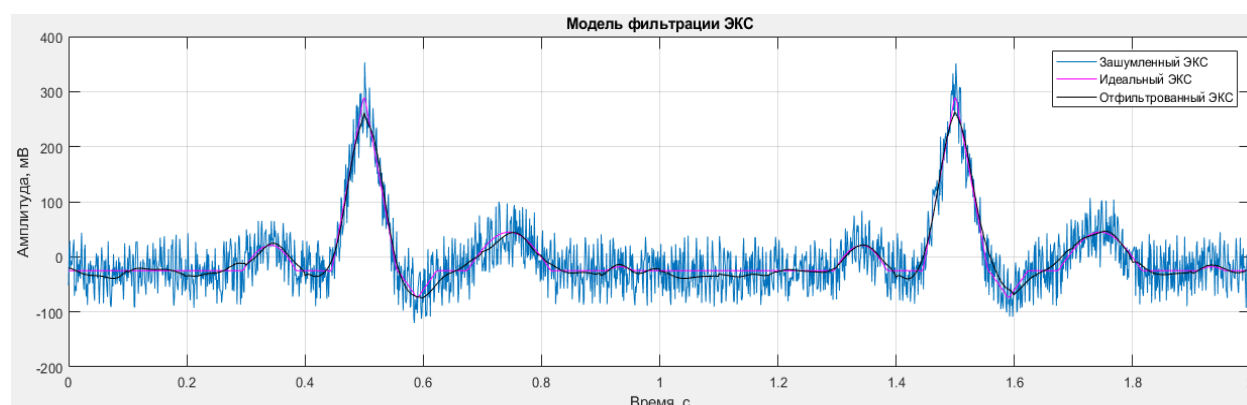


Рисунок 5. Увеличенный отфильтрованный ЭКС после применения метода комбинации фильтров Савицкого-Голея и Чебышева

Использование метода комбинации фильтров Савицкого-Голея и Чебышева снижает долю высокочастотных составляющих ЭКС, подавляет шумы на 90,93%, при этом максимальное значение СКО отклонения сигнала наблюдалось на интервале ST – 16,76% и минимальное значение в зубце R – 7,27%.

Заключение.

По результатам исследования можно сделать вывод, что метод комбинации фильтров Савицкого-Голея и Чебышева устраняет помехи без потерь качества ЭКС как на всем исследуемом фрагменте ЭКС, так и на критичных для диагностики участках. На рисунке 5 видно, что амплитуда зубцов и продолжительность интервалов, по которым ставится диагноз, не изменилась в значительной степени. По обработанному сигналу можно поставить правильный диагноз, следовательно, разработанный фильтр справляется с поставленной задачей: фильтрует шумы и помехи, не искажая исходный ЭКС.

Список литературы

- [1] Шарыкин, А. С. Врожденные и приобретенные заболевания клапанов сердца / А. С. Шарыкин. – Москва; Рязань: ГУП РО «Рязанская обл. тип.», 2019. – 216 с.
- [2] Полковникова, Н. А. Научные и инженерные расчеты в среде MatLab. Учебное пособие / Н.А. Полковникова. – М.: Моркнига, 2019. – 143 с.
- [3] Savitzky A., Smoothing and Differentiation of Data by Simplified Least Squares Procedures / A. Savitzky, J.E. Golay // Analytical Chemistry, 1964. – Vol. 36, № 8. – С. 1627 – 1639.
- [4] Schaferm, R.W. What Is a Savitzky-Golay Filter / R.W. Schaferm // IEEE signal processing magazine, 2011. – № July. – С. 111 – 117.
- [5] Данилов Ю.А. Многочлены Чебышева / Ю.А. Данилов – М.: Высш. школа, 1984. – 158 с.
- [6] Грибкова В.П. Эффективные методы равномерных приближений, основанные на полиномах Чебышева / В.П. Грибкова. – Москва: Спутник, 2017– 194 с.

FILTERS INFLUENCE ON ECG SIGNALS RESEARCH

A.V. Hardziyevich

*Open joint-stock company
«AGAT – Control Systems –
Managing Company of
Geoinformation Control
Systems Holding»,
engineer-programmer*

L.F. Vaskovskaya

*Open joint-stock company
«AGAT – Control Systems –
Managing Company of
Geoinformation Control
Systems Holding»,
leading system analyst*

P.V. Kamlach

*Deputy Dean of the Faculty
of Computer-Aided Design, Associate
professor of the Academic
department of
electronic technique and technology,
PhD, associate professor*

*Department of Electronic technique and technology Faculty of Computer Engineering
Belarusian State University of computer science and Radio Electronics, Republic of Belarus
E-mail: alinagordievich78@gmail.com*

Abstract. The ECG filters effectiveness comparison for ECG signals processing was completed in MatLab (computing platform). Savitzky-Golay filter, Chebyshev filter, combined Savitzky-Golay and Chebyshev filter analysis was carried out in according with the criterions of smoothing effectiveness, noise removal and the ECG signal morphology saving

The comparison results showed, that combined Savitzky-Golay and Chebyshev filters provided the best effectiveness according to all of the criterions. It allows to remove noises and artifacts without any distortion of the signal morphology and maintains high accuracy. Proposed approach can be used in wide spectrum of tasks – from ECG signals analysis to heart activity monitoring.

Keywords: ECG (ElectroCardioGram, ECG signals processing, ECG filters effectiveness, Savitzky-Golay filter, Chebyshev filter.

УДК 004.65+004.8

КАК БОЛЬШИЕ ДАННЫЕ И ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ МОГУТ ПРИНЕСТИ РЕЗУЛЬТАТЫ



Т.В. Казак

заведующий кафедрой инженерной психологии и эргономики, доктор психологических наук Республики Беларусь, доктор психологических наук Российской Федерации, член-корреспондент Международной академии психологических наук, профессор,
kazak@bsuir.by



А.В. Свороб

студент кафедры инженерной психологии и эргономики УО «Белорусский государственный университет информатики и радиоэлектроники»



А.Н. Василькова

ассистент кафедры инженерной психологии и эргономики, магистр УО «Белорусский государственный университет информатики и радиоэлектроники»,
a.vasilkova@bsuir.by

Т.В. Казак

Научный руководитель, заведующий кафедрой инженерной психологии и эргономики, доктор психологических наук Республики Беларусь, доктор психологических наук Российской Федерации, член-корреспондент Международной академии психологических наук, профессор.

А.В. Свороб

Студент кафедры инженерной психологии и эргономики УО «Белорусский государственный университет информатики и радиоэлектроники».

А.Н. Василькова

Ассистент кафедры инженерной психологии и эргономики УО «Белорусский государственный университет информатики и радиоэлектроники», магистр.

Аннотация. Большие данные являются необходимым условием для ИИ и основным фактором, который способствует повышению скорости и точности распознавания ИИ. С развитием и широким применением Интернета вещей количество генерируемых данных увеличилось в геометрической прогрессии с большим ежегодным приростом. Помимо увеличения числа, также была расширена размерность данных. Эти большие объемы многомерных данных делают данные более полными и более достаточными для поддержки.

Ключевые слова: большие данные, искусственный интеллект, бизнес, аналитика, машинное обучение.

Введение.

Термины «большие данные» и «искусственный интеллект (ИИ)» часто используются при построении будущего бизнеса. Потенциал их применения огромен как для оценки данных, так и для мгновенного получения необходимых решений в самых различных секторах бизнеса. Организации создали огромные хранилища данных. Однако простое хранение и управление большими объемами данных не дает большой ценности. В частности, машинное обучение используется для выявления закономерностей и предоставления рациональных рассуждений в обширных наборах данных, что дает возможность применять следующий уровень аналитики, необходимый для извлечения пользы из их данных.

Основная часть. Большие данные состоят из больших объемов информационных активов, которые требуют инновационной и экономической обработки, чтобы предоставить ценную информацию для принятия более эффективных решений. Затем эти данные можно просмотреть,

чтобы обнаружить корреляции или скрытые закономерности, которые в противном случае были бы недоступны.

Просто иметь вычислительную мощь и место для хранения, необходимые для накопления огромных объемов данных, недостаточно. Должен быть способ понять это. Поскольку ни один человек не может сканировать огромные объемы данных в поисках закономерностей или связей, которые можно использовать для стратегического планирования, для этой задачи используется искусственный интеллект. Данные позволяют организациям больше узнать о конкретных демографических показателях и их мотивах. Когда потребители используют технологию пассивно или активно, генерируются данные, описывающие ее. Это включает в себя кредитные карты, смартфоны, камеры и любое электронное устройство, которое расширяет их профиль данных. Когда анализ проведен правильно, учреждения могут многое узнать о поведении и характеристиках человека или группы. Затем такую важную информацию можно использовать для улучшения услуг или продуктов. В результате корпорации сейчас соревнуются в разработке самых мощных, точных и всеобъемлющих инструментов сбора и анализа данных.

Плюсы больших данных

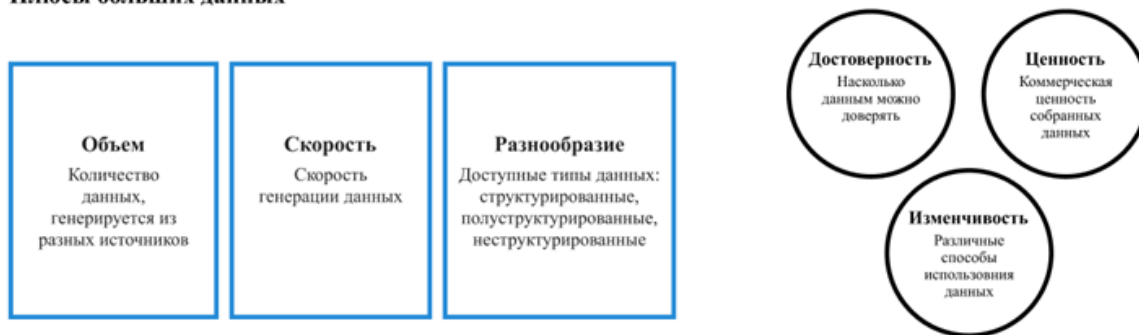


Рисунок 1 Плюсы больших данных

Большие данные и аналитика охватывают несколько технологий, которые функционируют одновременно, помогая учреждениям извлекать максимальную пользу из своей информации. Этими технологиями являются прогнозная аналитика, Hadoop, интеллектуальный анализ данных, машинное обучение и интеллектуальный анализ текста в памяти, а также управление данными.

Прогнозная аналитика: эта технология использует статистические алгоритмы и данные для определения будущих результатов на основе исторических данных. Его цель – помочь учреждениям планировать будущее. Он очень успешен в таких областях, как маркетинг, анализ рисков и обнаружение или предотвращение мошенничества.

Hadoop: программная среда с открытым исходным кодом, способная содержать огромные объемы данных. Он может использовать общедоступное оборудование для запуска приложений и считается незаменимым инструментом при работе с постоянно растущим разнообразием или объемом данных. Поскольку он использует вычислительную модель, основанную на распределении, он может быстро обрабатывать большие объемы данных. Еще одним преимуществом открытого исходного кода является то, что он находится в свободном доступе.

Машинное обучение: машинное обучение лучше всего рассматривать как подмножество ИИ, которое обучает машину различным методам обучения, позволяя ей быстро генерировать модели, которые затем могут анализировать большие наборы данных для получения более быстрых результатов с большей точностью. Этого можно добиться даже в больших масштабах. Если все сделано правильно, учреждения будут распознавать выгодные возможности, избегая при этом рисков, которые трудно измерить количественно.

Интеллектуальный анализ текста. Интеллектуальный анализ текста похож на интеллектуальный анализ данных, но с некоторыми отличиями. Здесь цель состоит в том, чтобы оценить текст, доступный в Интернете, в том числе в электронных книгах и разделах комментариев, чтобы получить информацию, которая ранее была недоступна. Интеллектуальный анализ текста тесно связан с естественной обработкой языка (*NLP*). Он позволяет просматривать документы, блоги, электронные письма и каналы социальных сетей, чтобы делать новые и интересные открытия.

Аналитика в памяти: при анализе данных системной памяти вы получите информацию, чтобы быстро на нее реагировать. Аналитика в памяти значительно сократит задержку обработки и подготовку данных. Новые сценарии могут быть протестированы для разработки новых моделей. Это позволяет учреждениям оставаться гибкими, принимая более разумные долгосрочные решения.

Управление данными. Недостаточно владеть большими объемами данных. Эта информация также должна быть систематизирована и иметь высокое качество. Здесь на помощь приходит управление данными. Оно позволяет постоянно поступающим и исходящим данным в организациях подвергаться процессам, которые делают их более расшифровываемыми и удобными для использования. Как только это произойдет, данные можно будет применять новыми и прибыльными способами. Это семь технологий, которые охватывают аналитику и большие данные. Однако для достижения желаемых результатов они должны беспрепятственно функционировать вместе с помощью искусственного интеллекта. По мере роста и развития вычислительной мощности ИИ будет соответственно ускоряться. Текущее состояние отрасли указывает на то, что она уже добилась значительных успехов и готова к более широкому внедрению. Это не совпадение, что обширный сбор данных и ИИ появились одновременно; одно не может существовать без другого.

Достижения в области машинного обучения (*МО*) привели к созданию нового мира, в котором данные можно использовать способами, которые раньше считались невозможными. Эта технология может быстро обрабатывать видео, текст, изображения и даже голоса. ИИ анализирует, чем больше этих данных, тем эффективнее он становится. Есть три основных способа, которыми ИИ приносит пользу большим данным: улучшенная аналитика данных, более невероятная скорость обработки и устранение проблем с данными.

Расширенная аналитика данных. Эффективное управление большими данными – одна из самых серьезных проблем, с которыми сталкиваются организации. Языки типа SQL использовались для извлечения нужных данных в течение нескольких лет. После этого потребовалось много времени и энергии, чтобы собрать ключевые идеи, которые часто включали старые методы, неэффективные. Это изменилось, поскольку теперь предпочтение отдается машинному обучению и искусственному интеллекту.

Более высокая скорость обработки. Когда дело доходит до обработки данных, скорость решает все. Хотя люди все еще привыкли анализировать данные и управлять ими, искусственный интеллект работает быстрее. Это может помочь людям в анализе данных, что приводит к более быстрому пониманию и способности принимать важные стратегические решения, которые могут вывести организации на новые высоты.

Устранение проблем с данными. Много проблем и проблем связано со сбором, управлением и обработкой данных. Самый большой связан с качеством полученной информации. Ни одна организация не хочет тратить много времени, денег и ресурсов на получение данных, которые в конечном итоге бесполезны. По этой причине алгоритмы машинного обучения сейчас используются для очистки и подготовки информации. Есть три дополнительных способа, которыми большие данные помогают корпорациям, а именно за счет более быстрого принятия решений, новых услуг и продуктов и снижения затрат.

Быстрое принятие решений: аналитика на основе оперативной памяти позволяет учреждениям анализировать данные и принимать решения быстрее, чем их конкуренты. Это дает им значительное преимущество, которое в конечном итоге может увеличить долю рынка.

Новые услуги и продукты: вы можете разрабатывать услуги с учетом их потребностей, если вы знаете, чего хотят ваши клиенты или клиенты, используя аналитику. В свою очередь, это приведет к большей лояльности к бренду и прибыли.

Снижение затрат: такие технологии, как облачная аналитика и искусственный интеллект, значительно снизят затраты, особенно при оценке и хранении новых источников данных. Кроме того, он может анализировать существующие бизнес-процессы, чтобы определять уникальные способы снижения затрат при одновременном повышении эффективности.

Розничная торговля. Поддержка клиентов сильно изменилась за последнее десятилетие. Потребители стали более искушенными и теперь хотят, чтобы розничные продавцы знали, чего они хотят и когда. Используя аналитику и большие данные, розничные продавцы могут сделать это, поскольку у них будет доступ к обширным хранилищам данных, которые можно будет оценить, чтобы предсказать тенденции и дать рекомендации для новых продуктов.

Банковское дело. Финансовые учреждения всегда имели доступ к большим объемам данных, но понять их было совсем другое дело. Теперь они могут брать эти неструктурированные данные, использовать ИИ и аналитику для их организации, а затем оценивать информацию, чтобы предоставлять более качественные банковские услуги, максимально повышая эффективность своих операций и защищая учетные записи клиентов от киберугроз.

Производство. Перед компаниями в производственном секторе стоит сложная задача по добыче сырья и производству готовой продукции, которую можно продавать с целью получения прибыли. Как вы понимаете, для этого необходимо выявить и решить проблемы, связанные с механическими неисправностями, проблемами с цепочками поставок и приложениями для перемещения. По этой причине аналитика данных незаменима для этого сектора, поскольку она может помочь выявить проблемы заблаговременно, одновременно повышая эффективность повседневных операций и снижая затраты.

Правительство. Управление и руководство нацией с миллионами, десятками миллионов или сотнями миллионов граждан никогда не было легкой задачей. Правительства должны поддерживать стабильность и действовать в рамках бюджета, не ставя под угрозу производительность и качество жизни своих граждан. Кроме того, правоохранительные органы должны защищать гражданских лиц от преступных элементов и преследовать тех, кто нарушает законы. Военные должны защищать от внешних противников. Для достижения всего этого требуется целостное представление, которое может обеспечить обширная аналитика данных. Правительства будут лучше понимать своих граждан, их потребности и желания, выявляя при этом потенциальные угрозы, которые могут привести к дестабилизации.

Здравоохранение. Те, кто работает в медицинском секторе, будь то врачи или медсестры, должны регулярно принимать трудные решения, которые являются вопросом жизни и смерти. Таким образом, наличие доступа к своевременной и точной информации имеет решающее значение. Сегодня работники медицинской отрасли принимают лучшие решения для пациентов и лучше осведомлены о логистике и доступности жизненно важных лекарств.

Естественные науки. Наука – это основа, которая делает возможными современные технологии. Однако исторически научные исследования были дорогостоящими и медленными, особенно в медицинском секторе. Нередки случаи, когда испытания заканчиваются неудачей по разным причинам. К счастью, появление аналитики, IoMT (*Интернет медицинских вещей*) и искусственного интеллекта открыло двери для значительного ускорения научных исследований.

Заключение.

Хотя мир больших данных и искусственного интеллекта может показаться сложным, они необходимы для того, чтобы предприятия оставались конкурентоспособными. Однако

внедрение надежных смешанных систем больших данных и искусственного интеллекта сопряжено с определенными трудностями.

Большие данные и ИИ неразрывно связаны. Успех последнего зависит от успеха первого, а также помогает предприятиям раскрыть потенциал своих хранилищ данных ранее недоступными или обременительными способами.

Список литературы

[1] Катькало В. С. Корпоративное обучение для цифрового мира / под ред. В. С. Катькало, Д. Л. Волкова. - М.: АНО ДПО «Корпоративный университет Сбербанка», 2017. - 200 с.

[2] В. Майер-Шенбергер, К. Кукьер. Большие данные. Революция, которая изменит то, как мы живем, работаем и мыслим, 2013. – 310 с.

HOW BIG DATA AND AI CAN DELIVER RESULTS

T. V. Kazak

Head of the Department of Engineering Psychology and Ergonomics, doctor of psychological sciences of the Republic of Belarus, doctor of psychological sciences of the Russian Federation, Corresponding Member of the International Academy of Psychological Sciences, Professor

A.V. Svorob

Student of the Department of Engineering Psychology and Ergonomics

A.N. Vasilkova

Assistant of the Department of Engineering Psychology and Ergonomics, master

Department of Engineering Psychology and Ergonomics

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

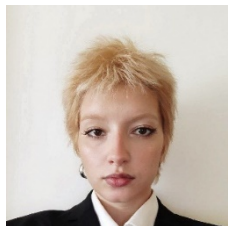
E-mail: a.vasilkova@bsuir.by

Annotation. Big data is a prerequisite for AI and a major driver of AI recognition speed and accuracy. With the development and widespread use of the Internet of Things, the amount of data generated has increased exponentially with a large annual increase. In addition to increasing the number, the dimension of the data was also expanded [105]. These large volumes of multidimensional data make the data more complete and more sufficient to support

Keywords: big data, artificial intelligence, business, analytics, machine learning.

УДК 004.738.5:

ПОТЕНЦИАЛЬНОЕ ВЛИЯНИЕ WEB3 НА B2B РЫНОК



Д.Ю. Маркевич

Студент 4 курса специальности
"Электронный маркетинг"
инженерно-экономического
факультета БГУИР



О.Н. Шкор

Старший преподаватель,
магистр экономических наук,
доктор философии в области
экономики

Белорусский государственный университет информатики и радиоэлектроники
Г. Минск, Республика Беларусь
E-mail: shkor@bsuir.by markevich.darya@gmail.com

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.

Д.Ю. Маркевич

Родилась в 2002 году в Минске. В 2019 году закончила ГУО «Гимназия №10 г. Минска». В этом же году поступила в УО «БГУИР», была зачислена на платную форму обучения по специальности «Электронный маркетинг» инженерно-экономического факультета.

Аннотация. В данной статье обсуждается, как Web 3, следующее поколение Интернета, основанное на технологии блокчейн, обладает потенциалом произвести революцию на рынке B2B. Web 3 может обеспечить более безопасную, эффективную и прозрачную платформу для транзакций, что может значительно снизить затраты и повысить эффективность. В статье предлагается, чтобы рынок B2B принял Web 3 и подготовился к изменениям, которые он принесет.

Ключевые слова: web 3.0, блокчейн, децентрализованный интернет, B2B

Введение.

В этой статье исследуется потенциальное влияние Web 3, следующего поколения Интернета, основанного на технологии блокчейн, на промышленный рынок B2B. В статье обсуждается, как Web 3 может обеспечить более безопасную, эффективную и прозрачную платформу для транзакций, которая может значительно снизить затраты и повысить эффективность. В статье предлагается, чтобы рынок B2B принял Web3 и подготовился к изменениям, которые он принесет.

Промышленный рынок B2B является важнейшей частью мировой экономики, и за последние годы он претерпел значительные изменения в связи с внедрением цифровых технологий. Однако ожидается, что появление Web3 окажет еще более глубокое влияние на рынок, чем любая предыдущая технология. В этой статье мы рассмотрим потенциальное влияние Web3 на B2B рынок и то, как компании могут подготовиться к нему.

Web3 это следующее поколение Интернета, основанное на технологии блокчейн. Эта технология разработана таким образом, чтобы быть более децентрализованной, безопасной и прозрачной, чем нынешняя сеть. [1] Web3 обладает потенциалом произвести революцию на рынке B2B, предоставляя более эффективную, безопасную и прозрачную платформу для транзакций. Одним из основных преимуществ Web3 является то, что он обеспечивает более безопасную платформу для транзакций. [2] Технология блокчейн, лежащая в основе Web3, гарантирует, что все транзакции записываются в децентрализованный реестр, который невосприимчив к взлому и фальсификации. Это означает, что значительно снижается риск мошенничества и кибератак, что крайне важно на рынке B2B.

Более того, Web3 может сделать транзакции более эффективными. [3] Нынешнему рынку B2B часто препятствуют медленные, дорогостоящие и громоздкие процессы. Однако Web3 может упростить эти процессы, предоставляя смарт-контракты, которые являются само исполняющимися контрактами и автоматически выполняются при выполнении определенных условий. Это может значительно сократить время и затраты, связанные с B2B-транзакциями.

Кроме того, Web3 может повысить прозрачность на промышленном рынке B2B. Децентрализованный реестр, лежащий в основе Web3, гарантирует, что все транзакции видны всем вовлеченным сторонам.[5] Это может помочь уменьшить количество споров, укрепить доверие и повысить эффективность. Это также может помочь в управлении цепочками поставок, обеспечивая видимость перемещения товаров и транзакций в режиме реального времени.

Кроме того, Web3 обладает потенциалом для создания новых бизнес-моделей, основанных на децентрализованных сетях. [6] Это могло бы привести к расширению возможностей выхода на рынок для малого бизнеса и стартапов, что усилило бы конкуренцию и инновации. Web3 также может обеспечивать одноранговые транзакции, что может еще больше снизить затраты и повысить эффективность.

Однако вместе с этими преимуществами возникают и проблемы. Чтобы полностью внедрить Web3, предприятиям необходимо инвестировать в необходимую инфраструктуру, включая системы планирования корпоративных ресурсов (ERP) с поддержкой блокчейна. Это требует значительных первоначальных инвестиций, но долгосрочные выгоды могут перевесить затраты. [4]

Более того, предприятиям необходимо адаптироваться к меняющейся нормативно-правовой среде. Технология блокчейн все еще находится на ранних стадиях своего развития, и отсутствует четкость в регулировании. Однако предприятия, которые умеют ориентироваться в этой среде, могут получить значительное конкурентное преимущество.

В случае использования B2B побеждает реальность. Проекты Web3 в настоящее время получают огромную долю шумихи и венчурных инвестиций, но сосредоточены на проектах B2C сомнительной долгосрочной полезности. До тех пор, пока не потребуются базовые функции, связанные с безопасностью, прозрачностью, KYC / AML, хранением записей и управлением транзакциями, руководители предприятий вряд ли будут рисковать своими активами в решениях на основе токенов. Между тем, внедрение частных разрешенных бухгалтерских книг и интеллектуальных промышленных контрактов будет продолжать расти во всех отраслях и географических регионах, и эти технологии предлагают реальную, осязаемую ценность. [2]

В заключение можно сказать, что Web3 обладает потенциалом произвести революцию на рынке B2B. [7] Это может обеспечить более безопасную, эффективную и прозрачную платформу для транзакций, что может значительно снизить затраты и повысить эффективность. Таким образом, B2B рынок должен принять Web3 и подготовиться к изменениям, которые он принесет. Предприятия, которые смогут адаптироваться и опережать события, будут иметь хорошие возможности для достижения успеха в быстро меняющемся бизнес-ландшафте.

Список литературы

- [1] Tomasz Tunguz. «Sizing the Web3 B2B Software Market» Электронный доступ: <https://tomtunguz.com/sizing-web3-b2b-market/>
- [2] William Fox «Web3 for B2B: Hype Versus Reality». Электронный доступ: <https://www.spiceworks.com/tech/innovation/guest-article/web3-for-b2b-hype-versus-reality/>
- [3] Nazhim Kalam. «Web 1.0 vs Web 2.0 vs Web 3.0» Электронный доступ: <https://enlear.academy/web-1-0-vs-web-2-0-vs-web-3-0-e428cfe09dde>
- [4] Web3: The hype and how it can transform the internet Электронный доступ: <https://www.weforum.org/agenda/2022/02/web3-transform-the-internet/>
- [5] Khullani M. Abdullahi. «B2B Growth in Web3». Электронный доступ: <https://khullani.medium.com/b2b-growth-in-web3-9bd9145ce99f>
- [6] Kaspersky. «Cryptography Definition». Электронный доступ: <https://www.kaspersky.com/resource-center/definitions/what-is-cryptography>
- [7] Rakesh Sharma. «Decentralized Finance (DeFi) Definition» Электронный доступ: <https://www.investopedia.com/decentralized-finance-defi-5113835>

WHAT IS WEB 3.0 AND WHAT PROBLEMS CAN IT SOLVE

O.N.SHKOR

*Senior Lecturer at the
Department of Economics of*

D.Y.Markevich
Student of BSUIR

Belarusian State University of Informatics and Radioelectronics s. Minsk, Republic of Belarus, Senior Lecturer at the Department of Economics, shkor@bsuir.by

Annotation. This article discusses how Web 3, the next generation of the Internet based on blockchain technology, has the potential to revolutionize the B2B industrial market. Web 3 can provide a more secure, efficient and transparent platform for transactions, which can significantly reduce costs and increase efficiency. The article suggests that the B2B industrial market should adopt Web 3 and prepare for the changes it will bring.

Keywords: web 3.0, blockchain, decentralized Internet, B2B

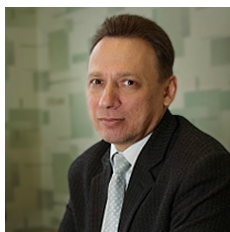
УДК 004.932.2

АЛЬТЕРНАТИВНЫЕ МЕТОДЫ КЛАССИФИКАЦИИ ПРИ РЕШЕНИИ ЗАДАЧ РАСПОЗНАВАНИЯ ЭНДОСКОПИЧЕСКИХ СНИМКОВ



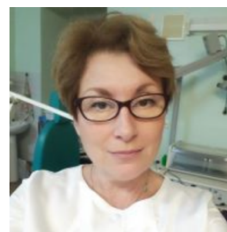
Р.В. Козарь

Аспирант БГУИР, инженер-программист ООО «Е-Ком Технологии»
pozitr@n.kozarroman@gmail.com



А.А. Навроцкий

Заведующий кафедрой информационных технологий автоматизированных систем БГУИР, кандидат физико-математических наук, доцент



Н.С. Конойко

Заведующий фониатрическим отделением консультативной поликлиники РНПЦ оториноларингологии

Р.В. Козарь

Окончил Белорусский государственный университет информатики и радиоэлектроники. Аспирант БГУИР. Работает в ООО «Е-Ком Технологии» в должности инженера-программиста. Проводит научные исследования данных медицинских изображений, полученных методом оптической эндоскопии для дальнейшего использования их в машинном обучении.

А.А. Навроцкий

Заведующий кафедрой информационных технологий автоматизированных систем Белорусского государственного университета информатики и радиоэлектроники. Кандидат физико-математических наук, доцент. Основные научные направления исследований «Компьютерное моделирование электронных приборов, систем и устройств СВЧ».

Н.С. Конойко

Врач-оториноларинголог-фониатр, заведующий фониатрическим отделением консультативной поликлиники РНПЦ оториноларингологии. Основные направления научной деятельности – «Коррекция голосоречевых нарушений при дисфункции мягкого нёба» и «Методы диагностики и реабилитации пациентов с нарушениями голосовой и дыхательной функций при парезах и параличах гортани различной этиологии».

Аннотация. В данной работе предлагаются альтернативные методы классификации данных на медицинских изображениях, полученных методом оптической эндоскопии посредством модификации исходного классификатора.

Ключевые слова: классификация, эндоскопические снимки, распознавание, связанный каскадный классификатор.

Введение.

В настоящее время все также важна и актуальна задача распознавания данных с различных источников для последующего принятия решений. В качестве данных в большинстве случаев используются изображения. Корректность принимаемых решений в огромной степени зависит от того, насколько качественно и точно будет классифицирован распознаваемый объект. Подходы как к распознаванию, так и к классификации могут различаться в зависимости от того, как сильно меняется вид и тип распознаваемого изображения. В общем случае кластеризация – это распределение пикселей изображения на несколько групп. Полученные группы представляют собой объекты (области) изображений. Решение такой задачи как никогда актуально для анализа визуальной информации: а именно – эндоскопических снимков. В данной статье рассматриваются особенности классифицирования подобных изображений. Следует сказать, что крайне хорошо себя зарекомендовала методика распознавания Виолы-Джонса.

Основная его особенность – это выделение примитивов в каждой области распознавания через пространство прямоугольных признаков Хаара, поскольку оно хорошо описывает

характерные особенности объектов, которые связаны с перепадами яркости вдоль вертикального и горизонтального направления. Однако есть такие задачи, в которых распознаваемый объект будет обладать крайне большим количеством перепадов вдоль диагональных направлений. Подобным объектом и будет являться эндоскопический снимок.

При распознавании подобных объектов важно сохранить максимально большое количество деталей, которые прямым образом влияют на принимаемое решение. Применение классического метода Виолы-Джонса с прямоугольными каскадами Хаара при распознавании эндоскопического снимка и его классификации имеет один существенный недостаток: будучи посчитаны поверх яркостного изображения подобные признаки не устойчивы к изменению освещенности. Конечно, есть вариант использования предварительной нормализации, однако однозначного решения у данной проблемы нет.

Модификация классификации для метода Виолы-Джонса.

Несмотря на то, что каскадная структура классификатора может существенно повысить скорость распознавания – у нее есть несколько недостатков.

Во-первых, информация, которая получается на текущем уровне каскада, не может быть передана на следующие уровни [1]. Поэтому решение о том, необходимо ли «отбрасывать» рассматриваемую область, никак не зависит от того, насколько он хорошо был распознан ранее. Поэтому такой подход может привести к построению «слабого» классификатора, который далее работает абсолютно некорректно при любых изменениях на отдельных уровнях [2]. Подобный подход изображен на рисунке 1.

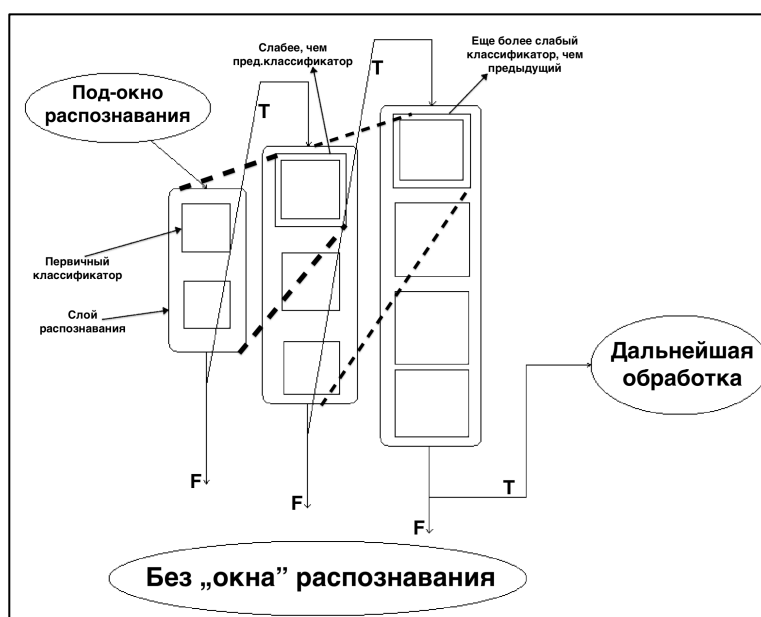


Рисунок 1. Связанный каскадный классификатор

Чтобы решить данную проблему – было предложено использовать в степень уверенности классификатора на предыдущем уровне распознавания в качестве слабого классификатора текущего уровня. В результате этого получится связанный каскадный классификатор, который будет работать более устойчиво.

Второй недостаток подобного каскада – это отсутствие оптимального и эффективного способа построения самих уровней каскада [3]. Причина в том, что при обучении классического каскадного прямоугольного классификатора для каждого из уровней необходимо указать несколько коррелирующих между собой параметров: ложные пропуски, ложные срабатывания и количество классификаторов [4].

Можно использовать различные наборы данных параметров для достижения оптимального результата распознавания, однако качество детектирования все равно может остаться неприемлемым [5].

Для решения данной проблемы в рамках данной работы был предложен новый каскадный классификатор, который немного похож на бустинговый классификатор, однако он способен отклонять отрицательные области после вычисления очередного слабого классификатора. Схема данного классификатора представлена на рисунке 2.

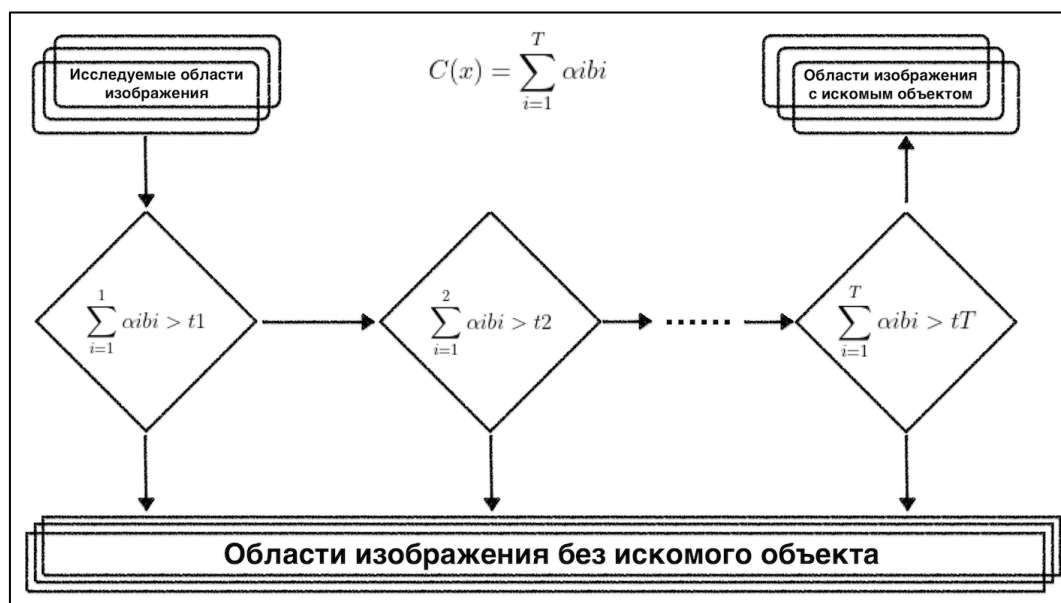


Рисунок 2. Схема модифицированного классификатора

Преимущество модифицированного классификатора достигается за счет сравнения частичных сумм линейных комбинаций слабых классификаторов с пороговыми значениями в процессе вычисления значения основного сильного классификатора.

Процедура построения каскада на основании модифицированного классификатора будет состоять из этапов: построения нового сильного классификатора и поиска необходимых пороговых значений по отсечению ненужных областей распознавания.

Результаты работы модифицированного классификатора.

В ходе написания данной статьи были подготовлены две выборки: положительная и отрицательная. Положительная выборка включала в себя в основном более светлые эндоскопические снимки (которые были сняты при нормальных условиях освещенности) в количестве 150 изображений.

Отрицательная выборка включала в себя 330 изображений, не содержащих необходимых с точки зрения медицины образов (область гортани).

При обучении использовались следующие параметры:

- доля ложных срабатываний при обучении очередного слоя оригинального и модифицированного классификатора 0,23;
- доля верных срабатываний при обучении очередного слоя оригинального и модифицированного классификатора 0,96;
- количество обучаемых слоев 12;
- физический размер классификатора в пикселях 30×30 ;
- набор масштабов классификатора от 1,0 до 3,5 с аддитивным шагом 0,09.

Оценка качества детектирования выполнялась на специальной тестовой обучающей выборке, которая состояла из 73 полноразмерных эндоскопических снимков при всевозможных

условиях освещенности. Результаты тестирования модифицированного классификатора для различных порогов классификации представлены в таблице 1.

Таблица 1 – Результат работы модифицированного классификатора.

№	Порог	<i>True Positive</i>	<i>False Positive</i>	<i>False Negative</i>	<i>Precision</i>	<i>Recall</i>
1	37	9	15	67	0,99	0,31
2	33	13	19	63	0,99	0,34
3	30	18	23	58	0,99	0,48
4	25	23	26	54	0,99	0,56
5	21	28	31	49	0,99	0,65
6	16	31	36	44	0,99	0,65
7	13	36	44	37	0,94	0,72
8	10	40	49	33	0,91	0,79
9	8	44	51	26	0,88	0,85
10	6	51	55	21	0,84	0,87
11	4	57	61	14	0,79	0,89
12	2	64	65	9	0,77	0,91
13	1	70	69	4	0,76	0,92

Результаты тестирования исходного классификатора для различных порогов классификации представлены в таблице 2.

Таблица 2 – Результат работы исходного классификатора.

№	Порог	<i>True Positive</i>	<i>False Positive</i>	<i>False Negative</i>	<i>Precision</i>	<i>Recall</i>
1	37	5	11	73	0,99	0,26
2	33	9	14	69	0,99	0,29
3	30	14	21	64	0,99	0,43
4	25	19	25	60	0,99	0,51
5	21	24	30	55	0,99	0,6
6	16	27	36	50	0,96	0,6
7	13	32	45	43	0,94	0,67
8	10	36	50	39	0,89	0,74
9	8	40	55	35	0,85	0,8
10	6	47	59	27	0,81	0,82
11	4	53	64	20	0,78	0,84
12	2	49	69	15	0,74	0,86
13	1	64	72	10	0,73	0,87

Для сравнения результатов были использованы метрики по точности распознавания и по отзыву распознавания (*Precision/Recall*) (1).

$$recall = \frac{TP}{TP + FN} \tag{1}$$
$$precision = \frac{TP}{TP + FP}$$

Precision интерпретируем как долю объектов, которые были названы классификатором положительными и при этом являются таковыми. *Recall* интерпретируем в качестве показателя, показывающего, какая доля объектов положительного класса из всех распознаваемых объектов положительного класса, была распознана корректно. *Recall* показывает способность классификатора обнаружить необходимый класс в принципе, а *Precision* отражает способность отличать найденный класс от других.

Используя приведенные выше табличные данные, был построен график для обоих классификаторов типа *Precision-Recall*, представленный на рисунке 3.

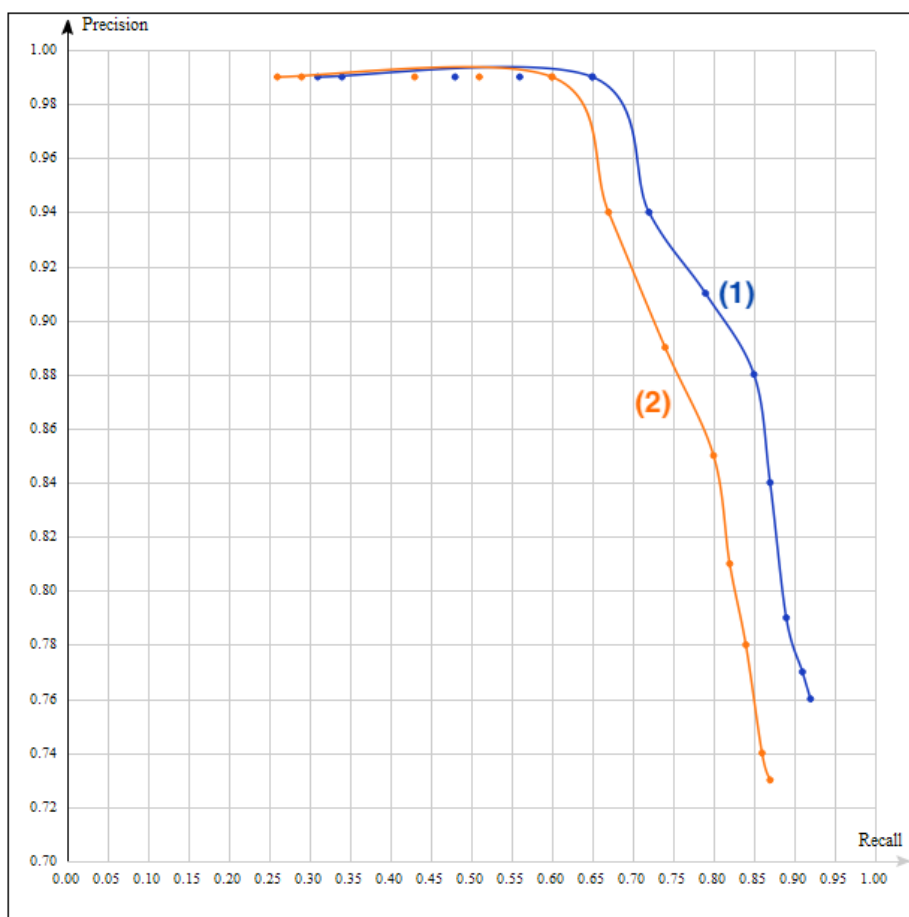


Рисунок 3. *Precision-Recall*-кривые исходного (2) и модифицированного (1) классификаторов

Заключение.

В данной статье был предложен новый модифицированный классификатор для алгоритма Виолы-Джонса для распознавания медицинских снимков, полученных методом оптической

эндоскопии, а также результаты его использования и сравнение с существующим исходным классификатором.

Были представлены результаты реализации нового классификатора для распознавания эндоскопических снимков. Также можно отметить тот факт, что данный новый классификатор может быть применен для любого алгоритма распознавания объектов, в котором есть и используется сканирующее окно и сильный каскадный классификатор.

По представленным на финальном графике результатам видно, что при одном и том же значении полноты точность модифицированного классификатора существенно выше.

Как результат – модифицированный классификатор обеспечивает лучшие показатели в пространстве *Precision/Recall*, что становится особенно важным для распознавания медицинских снимков, полученных методом оптической эндоскопии.

Список литературы

- [1] Xiao, R. Boosting Chain Learning for Object Detection / R. Xiao, L. Zhu, H.-J. Zhang // ICCV '03: Proceedings of the Ninth IEEE International Conference on Computer Vision. – 2003. – P. 710.
- [2] Wu, Bo. Fast rotation invariant multi-view face detection based on real AdaBoost / Bo Wu et al. // In Sixth IEEE International Conference on Automatic Face and Gesture Recognition. – 2004. – P. 81.
- [3] Dollár, P. Integral Channel Features / P. Dollár et al. // BMVC 2009 London Engl. – 2009. – P. 3–4.
- [4] Zhang, C. Multiple-instance pruning for learning efficient cascade detectors / C. Zhang, P. Viola // Adv. Neural Inf. Process. – 2008. – P. 2–4.
- [5] Bourdev, L. Robust Object Detection via Soft Cascade / L. Bourdev, J. Brandt // Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05). – Washington, DC, USA: IEEE Computer Society, 2005. – Vol. 2. – P. 239-241.

ALTERNATIVE CLASSIFICATION METHODS FOR SOLVING THE PROBLEMS OF RECOGNITION OF ENDOSCOPIC IMAGES

R.V. KOZAR

*Postgraduate student of the
BSUIR, software engineer
“E-COM Technologies”*

A.A. NAVROTSKY,

*Candidate of Physical and
Mathematical Sciences
Head of the Department of Information
Technologies of Automated Systems,
BSUIR*

N.S. KONOIKO

*Head of the Phoniatics
Department of the Consultative
Polyclinic of the Republican
Scientific and Practical Center
for Otorhinolaryngology*

*Belarusian State University of Informatics and Radioelectronics, Republic of Belarus
E-mail: pozitr0n.kozarroman@gmail.com*

Abstract. This article proposes alternative methods for classifying data on medical images obtained by optical endoscopy by modifying the original classifier.

Keywords: classification, endoscopic images, recognition, coupled cascade classifier.

УДК 342:004(476)

ГОСУДАРСТВЕННО-ПРАВОВОЕ РЕГУЛИРОВАНИЕ ПРОЦЕССОВ ЦИФРОВИЗАЦИИ В БЕЛАРУСИ



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 научно-методической работы учебно-методического отдела Института информационных технологий БГУИР, магистр управления и права
krysy_eg@bsuir.by

И.П. Сидорчук

Окончила Белорусский государственный университет и аспирантуру в Национальной академии наук Беларуси. Кандидат юридических наук, доцент. Работает в Белорусском государственном университете информатики и радиоэлектроники в должности заместителя директора по научно-методической работе Института информационных технологий. Сферу научных интересов составляют исследования в области нормотворческой деятельности, государственного управления, кадровой политики, оценки регулирующего воздействия, устойчивого развития общества, местного управления и самоуправления, 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. Кандидат технических наук, доцент. Работает в Белорусском государственном университете информатики и радиоэлектроники в должности директора Института информационных технологий. Сферу научных интересов составляют исследования в области цифровой трансформации и электронного правительства, информационно-коммуникационных технологий, вопросы устойчивого развития общества, государственного управления, непрерывного образования и диста.

Е.Г. Крысь

Окончила Белорусский институт правоведения, Институт государственной службы, магистратуру Академии управления при Президенте Республики Беларусь. Магистр управления и права. Аспирант Национального центра законодательства и правовых исследований Республики Беларусь. Работает в Белорусском государственном университете информатики и радиоэлектроники в должности заведующего сектором научно-методической работы учебно-методического отдела Института информационных технологий. Сферу научных интересов составляют исследования в области государственного управления, оценки регулирующего воздействия, непрерывного образования и профессионального развития руководителей и специалистов в условиях цифровой экономики, устойчивого развития общества.

Аннотация. В статье анализируется государственно-правовое регулирование процессов цифровизации в Беларуси. Аргументируется, что в Республике Беларусь создана и функционирует комплексная нормативная правовая база в сфере информатизации, информационных технологий и цифрового развития, которая включает в себя ряд нормативных правовых актов и технических нормативных правовых актов (государственных и отраслевых), которые обеспечивают формирование цифровой инфраструктуры, внедрение новых технологий, а также определяют и расширяют функционал государственных органов и организаций.

Ключевые слова: цифровизация, цифровая трансформация, цифровое развитие, государственно-правовое регулирование вопросов цифровизации, информационное общество, информационное пространство, цифровые технологии.

Введение.

На современном этапе вопросам цифрового развития со стороны учёных и практиков придается первостепенное значение. Исследуются, как экономические основы цифровизации, так и изучается ее инфраструктура, определяются угрозы и риски, переводятся в цифру различные отрасли экономики и сферы жизнедеятельности человека. Все общественные процессы находятся под воздействием цифровизации: большие данные; нейротехнологии и искусственный интеллект; системы распределенного реестра; квантовые технологии; новые производственные технологии; промышленный интернет; компоненты робототехники и сенсорики; технологии беспроводной связи; технологии виртуальной и дополненной реальности и др. которые внедряются и получают всё большее распространение в деятельности государственных органов и иных организаций, в том числе субъектов хозяйствования.

Цифровизация, как справедливо отмечает Д.А. Горулев, базируется на технологических решениях, которые в свою очередь меняют институциональные отношения и обуславливают формирование новых технологических решений и новых институтов, позволяющих выстраивать бизнес отношения в новых условиях [1].

Другой российский ученый говорит уже не с точки зрения бизнеса и экономики, а рассматривает влияние цифровизации на политические процессы и указывает, что политика переносится в диджитал-сферу, что повышает ее открытость и обеспечивает гласность [2]. Подтверждением этого тезиса, в частности, является решение Правительства Беларуси о создании Государственной единой (интегрированной) республиканской информационной системы учета и обработки обращений граждан и юридических лиц в соответствии с Законом Республики Беларусь «Об обращениях граждан и юридических лиц». Доступ пользователей к системе учета и обработки обращений обеспечивается посредством сайта в глобальной компьютерной сети Интернет по адресу: <https://обращения.бел>. Доступ государственных организаций к системе учета и обработки обращений может обеспечиваться с использованием выделенных каналов связи. Ряд ученых обращают внимание на вопросы экономической безопасности и отмечают нарастание рисков и угроз жизненно важным экономическим интересам государства в условиях цифровой трансформации [3]. Причем следует учитывать как внутренние угрозы, например, преступность, коррупцию и теневую экономику, так и внешние к числу которых могут быть отнесены, в частности, геополитические угрозы и транснациональная преступность,

По мнению Головенчик Г.Г., которая одна из первых в республике проводит исследования в области цифровизации, для успешного развития цифровой экономики необходимо наращивать национальные кадровые, интеллектуальные и технологические преимущества, формировать гибкую нормативную базу для внедрения цифровых технологий во все сферы жизни [4].

На значимость регулятивных функций права, возможности обеспечивать с его помощью стабильность, необходимую преобразовательную деятельность индивидуумов и публичных институтов указывают и многие российские ученые [5]. Право не только регулирует общественные процессы, но и, безусловно, задает темп для их трансформации и совершенствования.

В этой связи актуально определить содержание и роль государственно-правовых решений на понимание процессов цифровизации общества и перспектив их дальнейшего развития. Правовые акты создают базу цифровизации, обеспечивают формирование инфраструктуры, внедрение новых технологий, а также меняют функционал государственных органов и организаций. В Республике Беларусь приняты (изданы) сотни нормативных правовых актов (далее – НПА) на уровне различных государственных органов в сфере информатизации,

информационных технологий, кибербезопасности и т.д., которые включают, в том числе технические нормативные правовые акты (государственные и отраслевые).

Среди наиболее значимых решений Парламента следует выделить законы Республики Беларусь «Об информации, информатизации и защите информации» (2008 г. в редакции Закона 2022 г.), «Об электронном документе и электронной цифровой подписи» (2009 г. в редакции Закона 2022 г.), «О защите персональных данных» (2021 г.). Указанные документы определили подходы в работе с различной информацией, возможности ее использования, применения электронных документов, электронной цифровой подписи в электронных документах, а также обеспечения защиты персональных данных, прав и свобод физических лиц при обработке их персональных данных. Кроме того, Закон Республики Беларусь «Об информации, информатизации и защите информации» закрепил перечень субъектов регулирования и управления в области информации, информатизации и защиты информации. К ним, в частности отнесены: Президент Республики Беларусь, Совет Министров Республики Беларусь, Национальная академия наук Беларуси, Оперативно-аналитический центр при Президенте Республики Беларусь, Министерство связи и информатизации, иные государственные органы в пределах их компетенции.

Глава государства в соответствии с Конституцией Республики Беларусь определяет единую государственную политику и осуществляет иное государственное регулирование в области информации, информатизации и защиты информации. Совет Министров Республики Беларусь обеспечивает проведение единой государственной политики; координирует, направляет и контролирует работу республиканских органов государственного управления и иных государственных организаций, подчиненных Правительству Республики Беларусь; утверждает государственные программы, если иное не предусмотрено законодательными актами, и обеспечивает их реализацию; осуществляет иные полномочия. Важные задачи в области информатизации и защиты информации решает Оперативно-аналитический центр при Президенте Республики Беларусь. Эта структура, например, осуществляет государственное регулирование и управление в сфере технической и криптографической защиты информации в соответствии с законодательными актами; разрабатывает проекты нормативных правовых актов, в том числе технические нормативные правовые акты, и принимает (издает) такие акты по вопросам технической и криптографической защиты информации, участвует в разработке проектов нормативных правовых актов по вопросам информатизации; осуществляет иные полномочия в соответствии с законодательством.

Следует более детально остановиться на решениях Главы государства по вопросам цифровизации. Практически ежегодно эти вопросы получали закрепление в декретах, указах, распоряжениях Главы государства. Например, в Концепции национальной безопасности Республики Беларусь, утвержденной Указом Президента Республики Беларусь от 9 ноября 2010 г. № 575 (далее – Концепция национальной безопасности) Президент определил основные национальные интересы в информационной сфере, в том числе формирование и поступательное развитие информационного общества.

Декрет Президента Республики Беларусь от 21.12.2017 № 8 «О развитии цифровой экономики» определил правовые основы функционирования Парка высоких технологий и отнес к числу основных направлений его деятельности не только экспорт информационно-коммуникационных технологий и иных новых и высоких технологий, исключительных прав на результаты интеллектуальной деятельности в сфере новых и высоких технологий, но и разработку и внедрение информационно-коммуникационных и иных новых и высоких технологий в Республике Беларусь. Еще одна важная задача Парка высоких технологий – содействие кадровому обеспечению инновационного развития национальной экономики, развитие образования в сфере информационно-коммуникационных технологий.

В 2021 г. указами Президента Республики Беларусь были утверждены Программа социально-экономического развития Республики Беларусь на 2021-2025 гг. и Государственная

программа инновационного развития Республики Беларусь на 2021-2025 г. Программа социально-экономического развития Республики Беларусь на 2021-2025 гг., в частности, определила направления формирования развитой информационно-коммуникационной инфраструктуры (физической, гибридной и цифровой (виртуальной) на основе: расширения сети стационарного широкополосного доступа с использованием волоконно-оптических линий связи, завершения строительства сети сотовой связи четвертого поколения (4G), создания сети сотовой связи пятого поколения (5G).

Еще одним шагом в развитии цифрового общества стал Указ Президента Республики Беларусь от 7 апреля 2022 г. № 136 «Об органе государственного управления в сфере цифрового развития и вопросах информатизации» (далее – Указ № 136). Согласно Указу. № 136 на Министерство связи и информатизации возложены дополнительные функции государственного управления в сфере цифрового развития. К ним, в частности, относятся: регулирование деятельности по созданию и развитию государственных цифровых платформ, информационных систем и ресурсов (за исключением банковских систем, а также систем, предназначенных для осуществления особого контроля в сфере предотвращения легализации доходов, полученных преступным путем, финансирования террористической деятельности и финансирования распространения оружия массового поражения, и ресурсов, используемых для этих целей); участие в разработке содержания образовательных программ высшего, среднего специального образования и дополнительного образования взрослых для подготовки, переподготовки, стажировки и повышения квалификации кадров в сфере цифрового развития.

Кроме того, на это министерство возложены задачи по разработке в сфере цифрового развития проектов нормативных правовых актов и стратегий; государственных программ. Ряд новых полномочий касается согласования проектов нормативных правовых актов в части вопросов цифрового развития; проектов концепций и стратегий в части вопросов цифрового развития; проектов государственных и иных программ, в том числе отраслевых и (или) межотраслевых программ, программ социально-экономического развития административно-территориальных единиц, региональных комплексов мероприятий, в части мероприятий в сфере цифрового развития и др. Этот документ определил и новые структуры в области цифровизации: Центр цифрового развития, Центр перспективных исследований, «офисы цифровизации». «Офисы цифровизации», в частности, призваны выполнять работы, оказывать государственным органам и организациям услуги, по разработке (доработке) программного обеспечения; сопровождению государственных цифровых платформ и информационных систем.

В целях защиты национальной информационной инфраструктуры от внешних и внутренних угроз издан Указ Президента Республики Беларусь от 14 февраля 2023 г. № 40 «О кибербезопасности». Согласно этому документу в республике создается национальная система обеспечения кибербезопасности, элементами которой являются: Оперативно-аналитический центр при Президенте Республики Беларусь; Национальный центр обеспечения кибербезопасности и реагирования на киберинциденты; центры обеспечения кибербезопасности и реагирования на киберинциденты объектов информационной инфраструктуры государственных органов и иных организаций; оператор электросвязи по взаимодействию Национального центра кибербезопасности, центров кибербезопасности, а также государственных органов и иных организаций; объекты информационной инфраструктуры государственных органов и иных организаций; сети передачи данных, используемые для взаимодействия элементов национальной системы обеспечения кибербезопасности. Основными задачами национальной системы обеспечения кибербезопасности, в частности, являются: достижение максимальной скоординированности действий государственных органов и иных организаций по обнаружению, предотвращению и минимизации последствий кибератак на объекты информационной инфраструктуры; постоянный поиск потенциальных уязвимостей национального сегмента глобальной компьютерной сети Интернет и др. Этот Указ, за исключением некоторых положений, вступает в силу 17 августа 2023 г., т.е. государственным

органам и организациям предоставляется время на проведение соответствующих организационно-правовых, информационных, технических и иных мер по его реализации на практике.

Значимую роль в регулировании процессов развития информационного общества имеет Концепция информационной безопасности Республики Беларусь, утвержденная постановлением Совета Безопасности Республики Беларусь от 18 марта 2019 г. № 1 (далее – Концепция информационной безопасности). В Концепции определяются ключевые понятия в области информации, информатизации и защиты информации, в том числе информационная инфраструктура, кибербезопасность. Ключевым компонентом этого документа являются цели и направления государственной политики, меры по безопасности информационного пространства, информационной инфраструктуры, информационных ресурсов, а также механизмы реализации.

Изменения, происходящие в сфере экономики, на наш взгляд, должны найти отражение как в Концепции национальной безопасности, так и в Концепции информационной безопасности, в частности, актуально дать определения понятий: «цифровая экономика», «цифровая трансформация». Обоснованно также определить термин «цифровые технологии», к числу которых предлагается отнести: большие данные; нейротехнологии и искусственный интеллект; системы распределенного реестра; квантовые технологии; новые производственные технологии; промышленный интернет; компоненты робототехники и сенсорики; технологии беспроводной связи; технологии виртуальной и дополненной реальностей.

Кроме того, следует расширить перечень национальных интересов в экономической сфере, определив направления цифровой трансформации в промышленности, топливно-энергетическом комплексе, сельском хозяйстве, строительстве, транспорте и логистике, финансовом секторе, здравоохранении и др. Для вышеуказанных целей могут использоваться положения СТБ 2583-2020 «Цифровая трансформация. Термины и определения», введенного в действие с 1 марта 2021 г. В указанном документе, например, под цифровой трансформацией понимается – «проявление качественных, революционных изменений, заключающихся не только в отдельных цифровых преобразованиях, но в принципиальном изменении структуры экономики, в переносе центров создания добавленной стоимости в сферу выстраивания цифровых ресурсов и сквозных цифровых процессов» [6].

Не менее важное значение имеют и постановления Совета Министров Республики Беларусь, а также постановления иных государственных органов. Постановлением Совета Министров Республики Беларусь от 2 февраля 2021 г. № 66 утверждена Государственная программа «Цифровое развитие Беларуси» на 2021–2025 гг. Этот документ определяет приоритетные направления развития национальной информационно-коммуникационной инфраструктуры в 2021–2025 гг., в том числе: развитие беспроводного ШПД на основе сети сотовой подвижной электросвязи по технологии LTE (4G); развертывание сетей сотовой подвижной электросвязи пятого поколения (5G); повсеместное внедрение на базе возможностей сетей пятого поколения таких технологических решений, как «умный дом», «умный город» и других, а также дальнейшее развитие облачных технологий; развитие инфраструктуры электронного правительства и ряд других задач. В Государственной программе «Цифровое развитие Беларуси» на 2021–2025 гг. также предусмотрено, что доля специалистов, ответственных за вопросы информатизации в государственных органах и организациях, прошедших обучение в сфере цифрового развития к 2025 г., должна составить 40 %.

В целях достижения указанного показателя актуальна разработка и утверждение в республике специальной программы с условным названием «Кадры для цифровой Беларуси», в рамках которой можно предусмотреть обучение цифровой грамотности дошкольников, школьников, руководителей и специалистов, а также населения различных социальных и возрастных групп. Это позволит предпринять превентивные, дополнительные меры по обеспечению различных отраслей экономики компетентными кадрами, в том числе с учетом потребностей регионов. В рамках обучения необходимо обеспечить набор современных

«цифровых технологий», знание и использование которых позволит решать задачи цифровизации экономики.

Следует отметить и постановление Президиума Национальной академии наук Беларуси от 26 февраля 2018 г. № 17 «Об утверждении Стратегии «Наука и технологии: 2018–2040». В Стратегии отражены ключевые направления развития IT-отрасли и цифровой экономики в Беларуси: обработка больших данных (Big Data), в том числе на основе облачных технологий; интеллектуальные информационные системы; машинное обучение и человеко-машинное взаимодействие; квантовые и оптические технологии; создание цифровых производств, в том числе на основе аддитивных технологий; мехатронные системы и робототехнические комплексы; обеспечение кибербезопасности.

Среди НПА, регламентирующих вопросы цифровизации, имеются и технические нормативные правовые акты (государственные и отраслевые) (далее – ТНПА) (каталог Технических нормативных правовых актов можно посмотреть, перейдя по следующей ссылке: <https://tnpa.by/>). Виды ТНПА определены в статье 26 Закона Республики Беларусь от 17 июля 2018 г. № 130-З «О нормативных правовых актах» (рисунок 1).

К ТНПА в области технического нормирования и стандартизации относятся технические регламенты Республики Беларусь, технические кодексы установившейся практики, государственные стандарты Республики Беларусь, общегосударственные классификаторы Республики Беларусь, технические условия и стандарты организаций.



Рисунок 1. Структура видов ТНПА

ТНПА в области связи и информатизации и информационных технологий в соответствии с Законом о нормативных правовых актах могут быть систематизированы на 5 групп:

1. Технические регламенты:

ТР 2013/027/ВУ, ВУ «Информационные технологии. Средства защиты информации.

Информационная безопасность: технический регламент Республики Беларусь»;

ТР 2013/027/ВУ «Информационные технологии. средства защиты информации. информационная безопасность»;

ТР 2018/024/ВУ «Средства электросвязи. Безопасность»;

и другие.

2. Технические кодексы установившейся практики:

ТКП 223-2020 (33160) «Порядок организации центров обслуживания вызовов с использованием типовых технологий»;

и другие.

3. Государственные стандарты Республики Беларусь (СТБ, ГОСТ) и стандарты предприятий.

В свою очередь стандарты в зависимости от уровня их принятия можно разделить следующим образом:

- международные;
- региональные (включая межгосударственные);
- государственные (включая предварительные);
- стандарты предприятий.

Есть государственные стандарты, применение которых будет способствовать повышению безопасности ИКТ. В их число, в частности, входят следующие стандарты: СТБ 34.101.78-2019 «Информационные технологии и безопасность. Профиль инфраструктуры открытых ключей»; СТБ 34.101.81-2019 «Информационные технологии и безопасность. Протоколы службы заверения данных»; СТБ 2583-2020 «Цифровая трансформация. Термины и определения»; СТБ 34.101.31-2020 «Информационные технологии и безопасность. Алгоритмы шифрования и контроля целостности» и др.

4. Технические условия (ТУ) – утвержденный юридическим лицом или индивидуальным предпринимателем технический нормативный документ, содержащий комплекс требований, включая правила приемки и методы контроля, к конкретным типу, марке, модели, виду реализуемой продукции (оказываемой услуге). Объектом ТУ является конечный продукт, в отличие от стандарта, который может распространяться не только на конечный продукт, но и на отдельные его аспекты – маркировку, правила приемки, отдельные требования (требования безопасности и т. д.). Наиболее широко ТУ применяются в рамках договорных отношений между товаропроизводителями и потребителями продукции, а также торгующими организациями.

5. Общегосударственные классификаторы (ОГКС) – являются составной частью Единой системы классификации и кодирования технико-экономической и социальной информации (ЕСКК ТЭСИ) Республики Беларусь. Основное их предназначение состоит в упорядочении и обеспечении идентичности классификации информации, представляемой различными источниками (Информацию об общегосударственных классификаторах можно узнать, перейдя по следующей ссылке: <https://www.belstat.gov.by/klassifikatory/obschegosudarstvennye-klassifikatory-respubliki-belarus-ispolzuemye-dlya-zapolneniya-gosudarstvennoi-statisticheskoi-otchetnosti/>).

К числу ТНПА можно также отнести руководящие документы: инструкции; технические требования, технические спецификации; руководства; правила; положения; указания; нормы; прочие документы (рекомендации, технологические карты, концепции, программы и др.); методические документы.

С 1 января 2022 г. на территории Республики Беларусь введен национальный общегосударственный классификатор стандартов, разработанный на основе Международного классификатора стандартов в 2015 году. В соответствии с постановлением Государственного комитета по стандартизации Республики Беларусь № 100 от 14 октября 2021 года «Об утверждении, введении в действие и отмене общегосударственных классификаторов Республики Беларусь» с введением в действие ОКРБ 009-2021, отменяется общегосударственный классификатор МК (ИСО/ИНФКО МКС) 001-96 «Межгосударственный классификатор стандартов». Вновь введенный общегосударственный классификатор предназначен для систематизации информации при построении каталогов, указателей, выборочных перечней, библиографических материалов, формировании баз данных по международным, межгосударственным и государственным стандартам и прочим техническим нормативным правовым актам и документам в области технического нормирования и стандартизации. ОКГС содержит обязательные для соблюдения технические требования, направленные на распределение технико-экономической и социальной информации.

Объектами ОКГС являются международные, межгосударственные и государственные стандарты и прочие технические нормативные правовые акты, и документы в области технического нормирования и стандартизации.

Приведенный выше анализ показывает, что в Беларуси за последнее десятилетие создана комплексная нормативная правовая база, включающая НПА различной юридической силы, в том числе ТНПА.

Отдельно в рамках определения мер государственно-правового регулирования следует остановиться на компетенции некоторых государственных органов и иных структур. Выше отмечалось, что полномочия Министерства связи и информатизации Республики Беларусь значительно расширены с учетом современных подходов к цифровизации. В частности, согласно Указу № 136 министерство наделено широкими полномочиями в сфере управления процессами цифрового развития, цифровизации системы государственного управления и всех отраслей экономики. Помимо управления процессами цифровизации за регулятором закреплены нормотворческие, организационные, образовательные, контрольные и другие функции (пункт 1, 2 Указа № 136).

Кроме того, в соответствии с пунктом 2.9 Указа № 136 Министерство связи и информатизации Республики Беларусь осуществляет в порядке, установленном законодательством о контрольной (надзорной) деятельности, контроль за:

- соблюдением нормативных правовых актов, в том числе технических нормативных правовых актов, по вопросам цифрового развития и установленного порядка реализации мероприятий в сфере цифрового развития;
- соблюдением единой технической и технологической политики в сфере цифрового развития и соответствием единой архитектуре государственных цифровых платформ, присоединяемых к ним информационных систем;
- выполнением требований к созданию, развитию, взаимодействию государственных цифровых платформ, государственных информационных систем, не предназначенных для обработки информации, отнесенной к государственным секретам, и обеспечению их совместимости.

Указом Президента от 16 октября 2009 г. № 510 «О совершенствовании контрольной (надзорной) деятельности в Республике Беларусь» (далее – Указ № 510) определен единый порядок проведения контрольной (надзорной) деятельности в Республике Беларусь, а также создание дополнительных условий для развития эффективных форм хозяйствования. Этим же Указом определены перечень контролирующих (надзорных) органов, уполномоченных проводить проверки и сферы их контрольной (надзорной) деятельности. Однако механизм реализации положений Указа № 136 не предусмотрен в Указе № 510. В этой связи предлагается внести в приложение «Перечень контролирующих (надзорных) органов, уполномоченных проводить проверки, и сфер их контрольной (надзорной) деятельности» к Указу № 510 соответствующие дополнения.

Кроме того, с точки зрения соблюдения единой технической и технологической политики и соответствия единой архитектуре государственных цифровых платформ, присоединяемых к ним информационных систем, можно отметить также риски и угрозы, связанные с цифровыми платформами, минимизация которых также может быть отнесена к контрольной (надзорной) деятельности соответствующих органов.

Одной из главных угроз, связанных с деятельностью цифровых платформ, является проблема конфиденциальности персональных данных. Цифровые платформы собирают и обрабатывают большие массивы данные об участниках, отслеживая все их действия и личную информацию, и могут использовать их, например, для продажи в целях рекламы, недобросовестной конкуренции, манипулирования участниками, рынками и даже государствами.

Можно также отметить и ряд других немаловажных проблем: отсутствие четкого и гибкого международного и национального законодательного регулирования деятельности платформенных компаний и неоднозначность судебной практики; проблемы обеспечения безопасности транзакций и многие другие.

Заключение.

На основании изложенного полагаем возможным отметить следующее:

1. В Республике Беларусь создана комплексная нормативная правовая база в сфере информатизации, информационных технологий и цифрового развития, которая включает в себя ряд нормативных правовых актов и технических нормативных правовых актов (государственных и отраслевых).

2. В целях достижения целей, определенных в Государственной программе «Цифровое развитие Беларуси», наряду с решением задач в области совершенствования инфраструктуры, внедрения новых технологий и др. актуальна разработка и утверждение программы «Кадры для цифровой Беларуси», в рамках которой можно предусмотреть обучение цифровой грамотности руководителей и специалистов, а также населения различных возрастных и социальных групп.

3. Процессы цифровизации, в том числе цели, задачи, механизмы, понятия: «цифровая экономика», «цифровая трансформация», «цифровые технологии», следует закрепить как в Концепции национальной безопасности, так и в Концепции информационной безопасности,

4. Министерство связи и информатизации Республики Беларусь не включено в Перечень контролируемых (надзорных) органов, уполномоченных проводить проверки, и сфер их контрольной (надзорной) деятельности, утвержденный Указом № 510. Предлагается дополнить Перечень контролируемых (надзорных) органов, уполномоченных проводить проверки, и сфер их контрольной (надзорной) деятельности (приложение к Указу № 510) новой позицией, согласно которой Министерство связи и информатизации Республики Беларусь включается в число контролируемых (надзорных) органов и осуществляет контрольные (надзорные) функции в следующих сферах:

- контроль за соблюдением нормативных правовых актов, в том числе технических нормативных правовых актов, по вопросам цифрового развития и установленного порядка реализации мероприятий в сфере цифрового развития;

- контроль за соблюдением единой технической и технологической политики в сфере цифрового развития и соответствием единой архитектуре государственных цифровых платформ, присоединяемых к ним информационных систем;

- контроль за выполнением требований к созданию, развитию, взаимодействию государственных цифровых платформ, государственных информационных систем, не предназначенных для обработки информации, отнесенной к государственным секретам, и обеспечению их совместимости.

Список литературы

[1] Горулев, Д.А. Экономическая безопасность в условиях цифровой экономики [Электронный ресурс]. – Режим доступа: cyberleninka.ru/article/n/ekonomicheskaya-bezopasnost-v-usloviyah-tsifrovoy-ekonomiki?ysclid=ld9zqn1ctd136141277. – Дата доступа: 12.03.2023.

[2] Мельцов, В. М. Влияние цифровизации на развитие общественных процессов [Электронный ресурс]. – Режим доступа: <https://cyberleninka.ru/article/n/vliyanie-tsifrovizatsii-na-razvitie-obschestvennyh-protsessov?ysclid=lfkqtj2lhg399252358>. – Дата доступа: 12.03.2023.

[3] Никитенко, П. Г. Экономическая безопасность: теория, методология, практика и / П.Г. Никитенко, В.Г. Булавко. – Минск: Право и экономика, 2009. 483 с.

[4] Головенчик, Г.Г. Цифровая экономика в Республике Беларусь: современные тенденции, вызовы и перспективы [Электронный ресурс]. – Режим доступа: <https://cyberleninka.ru/article/n/tsifrovaya-ekonomika-v-respublike-belarus-sovremennye-tendentsii-vyzovu-i-perspektivy?ysclid=lfktr84jze55531950>. – Дата доступа: 12.03.2023.

[5] Тихомиров, Ю.А. Право и цифровая трансформация [Электронный ресурс]. / Ю.А. Тихомиров, Н.В. Кичигин, Ф.В., Цомартова, С.Б. Бальхаева. – Режим доступа: <https://www.hse.ru/data/2021/07/22/1421019827/%D0%A2%D0%B8%D1%85%D0%BE%D0%BC%D0%B8%D1%80%D0%BE%D0%B2.pdf?ysclid=lfktdsaty549808362>. – Дата доступа: 12.03.2023.

[6]. СТБ 2583-2020 введен в действие с 01.03.2021 «Цифровая трансформация. Термины и определения.» [Электронный ресурс]. – Режим доступа: <https://nasb.gov.by/rus/activity/nauchno-metodicheskoe-obespechenie-razvitiya-informatizatsii/>. – Дата доступа: 12.03.2023.

STATE-LEGAL REGULATION OF DIGITALIZATION PROCESSES IN BELARUS

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

*Institute of Information Technology Belarusian State University of Informatics and Radioelectronics
E-mail: irina_sidorchuk@bsuir.by; okhrimenko@bsuir.by; krys_eg@bsuir.by*

Abstract. The article analyzes the state-legal regulation of digitalization processes in Belarus. It is argued that a comprehensive regulatory legal framework in the field of informatization, information technology and digital development has been created and is functioning in the Republic of Belarus, which includes a number of regulatory legal acts and technical regulatory legal acts (state and industry) that ensure the formation of a digital abstract: infrastructure, introduction of new technologies, as well as define and expand the functionality of state bodies and organizations.

Keywords: digitalization, digital transformation, digital development, state-legal regulation of digitalization issues, information society, information space, digital technologies.

УДК 378

О МЕТОДИКЕ ПРЕПОДАВАНИЯ ПРОГРАММИРОВАНИЯ НА СТАРШИХ КУРСАХ ВУЗОВ



Д.И. Черемисинов

*Ведущий научный сотрудник ОИПИ НАН Республики Беларусь
кандидат технических наук, доцент
cher@newman.bas-net.by*

Д.И. Черемисинов

Окончил Томский государственный университет, кандидат технических наук, доцент. Работает в ОИПИ НАН Беларуси в должности ведущего научного сотрудника, более 10 лет преподавал в Белорусском государственном университете информатики и радиоэлектроники курс «Интерфейсы информационных систем».

Круг научных интересов: программирование, логическое проектирование и тестирование дискретных систем управления, реализация параллельных алгоритмов управления.

Аннотация. Дисциплина «Интерфейсы информационных систем» в учебном плане для подготовки специалистов по информационным технологиям нацелена на формирования объема знаний о взаимосвязи ключевых концепции и принципов программирования. В курсе отражены современные концепции и методические подходы к обучению информатике студентов старших курсов ВУЗов. Материал курса можно считать ядром структуры понятий информатики, так как он расположен в области перекрытия всех четырех ее частей. Одной из важных задач преподавания курса является демонстрация того, что сформированные у студентов на начальных стадиях обучения представления об основных понятиях программирования слишком поверхностны.

Ключевые слова: Преподавание информатики в университетах, информационное взаимодействие, интерфейс, кривая обучения, метод наставничества.

Введение.

Снижение посещаемости занятий отмечалась как проблема для большинства учебных дисциплин, читаемых на старших курсах нашего факультета (Факультет компьютерного проектирования, БГУИР), это наблюдается и в других высших учебных заведениях [1]. Существуют различные гипотезы и мнения, почему это происходит. В этой работе сделана попытка объяснения феномена снижения посещаемости занятий на старших курсах вузах на опыте преподавания дисциплины «Интерфейсы информационных систем», целью которого является продолжение обучения инженеров системотехников программированию на четвертом курсе бакалавриата. Дисциплина читается в БГУИР с 2010 г [1].

Область знаний, связанная с информационными технологиями, принято разделять на четыре основные дисциплины – информатика (computer science), программная инженерия (software engineering), проектирование аппаратных платформ (hardware engineering) и информационные системы (information systems) [3,4]. Профессиональными объединениями разработаны рекомендации по преподаванию программирования и информатики в университетах [3,4], которым должны следовать учебные программы вузов, чтобы квалификация выпускаемых ими специалистов соответствовала международным требованиям. В рекомендациях [3] различаются объем знаний (body of knowledge) выпускников и описание конкретных учебных курсов, посредством которых выпускники овладевают этими знаниями.

В [3] предлагается строить отдельные руководства по разработке учебных планов для подготовки специалистов в каждой из пяти дисциплин. Описание объема знаний по дисциплине «программная инженерия» организовано в виде трехуровневой иерархии. На верхнем уровне

иерархии находятся области преподаваемых знаний, представляющие собой отдельные поддисциплины программной инженерии. Каждая сфера состоит из своей группы тематических модулей, которые обозначаются аббревиатурой сферы с добавлением порядкового номера. Каждый модуль состоит из тем.

В качестве одной из целей обучения в [3] рекомендуется среди прочего формировать у студентов инженерный склад ума, основанный на математическом подходе к решению проблем. Для этого в преподавании дисциплин инженерного профиля задачей обучения является формирование целостной системы взаимосвязанных понятий, составляющих предмет дисциплины [4]. Несмотря на то, что важность математической компетенции выпускника и формирования целостной системы взаимосвязанных понятий неоднократно отмечена в [2], там отсутствует тема или модуль специально посвященная формированию системы понятий программирования.

Дисциплина «Интерфейсы информационных систем» включена в учебный план для подготовки специалистов по информационным технологиям, так как ее целью является формирования объема знаний о взаимосвязи ключевых концепции и принципов программирования, хотя эта дисциплина не является темой или модулем из рекомендаций [3]. В основе метода преподавания лежит тезис о том, что в результате изучения курса в голове студента должна сформироваться некоторая целостная картина, а не пёстрая мозаика из разрозненных фактов.

Взаимодействие как главная тема программирования.

В учебных программах ведущих вузов, готовящих программистов, имеются курсы, целью которых является формирование общего взгляда на объем знаний по информационным технологиям в целом. Часто этот курс называется системным или теоретическим программированием. Например, в МГУ имеется курс «Фундаментальные аспекты системного программирования. Теория и практика разработки системного ПО и компонентов ядра ОС Linux». В этом курсе основой объединения ключевых понятий программирования является задача построения операционной системы (ОС). В БГУ на кафедре технологий программирования читается дисциплина «Системное программирование». Следует отметить, что дисциплина «Системное программирование» не везде используется для формирование общего взгляда, часто это спецкурс для подготовки системных администраторов – высших IT менеджеров. Во многих вузах основой курса «Системное программирование» является задача построения компилятора языка программирования. Эти курсы не дают студентам понимания, как работают программное обеспечение в целом, как прикладные программы используют операционную систему или как аппаратура компьютера, влияет на производительность и правильность программ приложений.

Для решения этой проблемы в университете Карнеги-Меллон разработан вводный курс по компьютерным системам, который называется «Введение в компьютерные системы» [5]. В этом курсе программирование рассматривается более широко, чем в традиционных курсах по организации процесса программирования или логическому проектированию, и охватывает темы архитектуры компьютеров, операционных систем, компиляторов и вычислительных сетей. Авторы курса считают его целью объяснить устойчивые концепции, лежащие в основе всех компьютерных систем, и показать конкретные способы, которыми эти идеи влияют на правильность, производительность, и полезность прикладных программ. Это цель близка направлению нашего курса, однако в университете Карнеги-Меллон он читается в начале обучения программированию, а наш – в конце.

В нашем курсе для БГУИР основой материала для преподавания является не задача, а понятие, именно понятие взаимодействия. Информационное взаимодействие – это воздействие объектов или агентов друг на друга, приводящее к изменению их состояния. Понятие взаимодействия является одним из центральных в кибернетике и информатике. Это понятие связано с понятием интерфейса. Слово интерфейс пришло в профессиональный язык

программистов из английского языка, где его определяют как: Interface is a shared boundary between two separate subsystems (сопряжение, линия раздела, перегородка двух подсистем) или как Interface is a point of interaction between subsystems (точка взаимодействия подсистем). Понятие взаимодействия естественно подходит для объединения как ключевых, так и самых современных идей программирования.

Разделы в совокупности знаний, на которых концентрируется этот курс – это: 1) архитектура и организация ЭВМ, 2) операционные системы, 3) распределенные вычисления. Упор в этом курсе сделан на предотвращение чрезмерного упрощения процесса программирования. В начале обучения эти темы преподаются поверхностно, чтобы сделать материал доступным для начинающих студентов. Основное внимание уделяется не разработке программы, а более простому процессу кодирования. Особенностью нашего курса является более углубленное изучение структуры программного обеспечения с целью демонстрации многогранности связей его частей (рисунок 1). Форма изучения решений, использованных в современных программных системах построена так, чтобы студенты не переоценивали свое владение навыками программирования. Уверенность в достаточности их умения в программировании скрывает необходимость в фундаментальных знаниях, отсутствие которых будут мешать им в поиске решений новых проблем, которые будут возникать в их профессиональной деятельности.



Рисунок 1. Зависимости понятия «мультиагентная система»

Материал курса можно считать ядром структуры информатики, так как он расположен в области перекрытия всех четырех ее частей. Предполагается, что слушатели обладают опытом программирования, т.е. овладели принципом программирования. Стоит задача не учить программированию; а учить правильному подходу к разработке программ. Студенты, которые только начали заниматься профессиональным программированием, должны понять, что существует множество приемов для решения типичных задач.

Состояние набора понятий, связанных с программированием, фиксируется в глоссариях терминов (например [7]). Глоссарий [7] содержит около 12000 понятий. Очевидно, учебный курс не может использовать такое количество понятий. В преподавании курсов, формирующих математический подход к решению проблем, стремятся обойтись небольшим количеством понятий (обычно порядка 30). Глоссарий терминов, связанных с компьютерами, для пенсионеров [8] содержит 37 понятий. Для профессионального обучения в обсуждаемом курсе

выбраны 150 понятий. Довольно, часто отбор терминов в глоссарии осуществляется статистическими методами, путем анализа литературы. В глоссарии обсуждаемого курса отобраны понятия ядра информационных технологий, встречающиеся во всех четырех основных дисциплинах. Порядок обсуждения выбранных понятий формировался так, чтобы создать представление об их внутренней взаимосвязи.

Эксперименты с методом наставничества в информатике.

Можно предположить, что снижение посещаемости занятий происходит из-за стандартной практики преподавания – «мудреца на сцене», обусловленной теоретической, абстрактной природой преподаваемого материала, которая делает лекцию скучной. Предполагается, что решить проблему можно сделав лекции привлекательными, менее сухими и пассивными, чтобы студенты захотели приходить на них. Другая форма лекции, являющаяся более интерактивным способом передачи знаний и обучения, например метод наставничества, возможно, увеличит посещаемость.

Обучение методом наставничества (Peer Instruction) [9] широко использовалось в преподавании различных дисциплин, таких как философия, психология, геология и биология, и показало многообещающие результаты в обучении студентов. На Западе этот подход рассматривается как перспективная область теоретических и практических разработок, способная изменить подходы к обучению, особенно в высшем профессиональном образовании и в корпоративном секторе. В этом случае речь идет об «управляемой» коммуникации посредством использования методик, требующих активной самостоятельной работы и взаимодействия учащихся. Подход был предложен Эриком Мазуром и его коллегами при изучении физики. Методика предполагает вовлечение студентов в активную деятельность по изучению материала дисциплины с последующим объяснением их своим сверстникам [9].

На занятии методом наставничества преподаватель включает в лекцию последовательность вопросов; задаваемых слушателям. Занятия проводятся по четко определенному протоколу преподавания. Каждый вопрос ориентирован на освещении очередного понятия, и от каждого слушателя требуется ответ в форме сигнала определенного формата. Для сигнализации о варианте ответа часто используются специальные интерактивные устройства – кликеры. Одной из важнейших предпосылок успеха метода является задание на чтение перед занятием; то есть, студенты должны заранее знакомиться с темой, затрагиваемой на занятии. Таким образом, предполагается, что класс обладает базовыми знаниями по теме, достаточными для понимания вопросов во время лекции и обсуждения ответов с другими студентами.

Тем не менее, несмотря на область знания и способ подачи материала лекции, и студенты не упускают возможность пропускать их. Сам по себе метод наставничества не увеличивает посещаемость занятий на старших курсах [2]. Важно отметить, что на младших курсах особых проблем с посещаемостью нет.

Применение методики наставничества на старших курсах наталкивается на трудности, тоже связанные со снижением посещаемости занятий. Как уже упоминалось ранее, ключевым условием успеха этой методики является необходимость самостоятельного знакомства обучающегося с темой, затрагиваемой на занятии, до начала занятия. Невозможно предположить, что студент, который пропускает занятия (не имеет мотивации к изучению материала), будет осваивать тему не в учебное время.

Очевидно, проблема не в процессе преподавания (лекторе, способе подачи материала или самом предмете), а в объекте обучения – студентах старших курсов. У этих студентов наблюдается снижение мотивации к обучению.

Кривая обучения.

О понятиях «Кривая обучения» (learning curve) и «Крутая кривая обучения» сейчас часто говорят и пишут в области компьютерных наук. Кривая обучения является графическим представлением того, как результат обучения зависит от его длительности; или зависимостью

производительности работника от опыта работы (чем больше кто-то или что-то выполняет операцию, тем лучше она выполняется). Более широко распространено понимание кривой обучения как математического описание изменения производительности работников при выполнении повторяющихся задач [10].

Кривую обучения можно рассматривать как представление зависимости умственных затрат на обучения чему-либо от времени. Умственные затраты могут быть измерены в числе изученных понятий. Именно такой подход используется в рекомендациях [3]. Число понятий, которыми должен овладеть выпускник, должно постоянно возрастать. Выражение «крутая кривая обучения» определяет кривую обучения как скорость, с которой приобретает навык, поэтому скорость обучения характеризует как способности обучаемого, так и трудность предмета. Число понятий можно считать оценкой трудности предмета, так как это число растет в процессе обучения, крутизна кривой обучения должна уменьшаться. Для демонстрации процесса изучения языка [11] предложена двухмодальная кривая обучения – bipolar learning graph (рисунок 2).

У студентов старших курсов велика уверенность, что обучение специальности уже закончилось – первый пик на кривой обучения на рис. 2. Это уверенность поддерживается тем, что большинство из них уже работают в организациях отрасли производства программного обеспечения. К моменту окончания обучения, по специальности или в смежных областях работают почти 100% обучающихся. Кроме того, они преодолели крутизну первоначального обучения и не сомневаются в том, что специальность программирования легкодоступна (реклама курсов по программированию уверяет, что почти любого грамотного человека можно сделать программистом за пару месяцев). Серьезные знания фундаментальных наук (математики, логики, физики) не требуются на должностях, которые им предложили их работодатели. Они убеждены в том, что обучение происходит быстро, а знания состоят из отдельных фактов. Большинство студентов, особенно слабых, сильно недооценивают время, необходимое для освоения профессии, и считают, что одной хорошей книги достаточно. Они пытаются запоминать отдельные факты, а не считают важным понимание взаимосвязи понятий.

Сомнение в достаточности знаний (спад уверенности после пика на рисунке 2) в дальнейшем по мере профессионального роста появится. Продвижение в области профессиональной деятельности связано с усложнением работы. Чем выше должность (и зарплата) тем сложнее и дороже работа, которую выполняет работник. Работник повышается в должности до тех пор, пока не займёт место, на котором он окажется не в состоянии справиться со своими обязанностями, то есть окажется некомпетентным (принцип Питера [12]). В области информационных технологий повторяемость одинаковых рабочих заданий невысока. Обнаружение собственной некомпетентности происходит не в конце профессиональной карьеры как в традиционных отраслях, а в ее начале. Не приобретая новых знаний, сотрудник рискует «застрять» в должности до тех пор, пока не покинет организацию (то есть не уволится, не умрёт или не выйдет на пенсию). Проблема в том, что мотивация к обучению возникает после окончания обучения в университете.

Курс «Интерфейсы информационных систем» эффективен на стадии обучения, соответствующей промежуточному минимуму на рисунке 2, когда появляется мотивация к повышению уровня образованности вследствие ощущения недостаточности знаний. В процессе обучения в университете у студента осознание недостаточности знаний само по себе не может возникнуть. Следовательно, задачей преподавателя на этом этапе является демонстрация того, что имеющиеся у обучаемых представления об основных понятиях программирования слишком поверхностны.

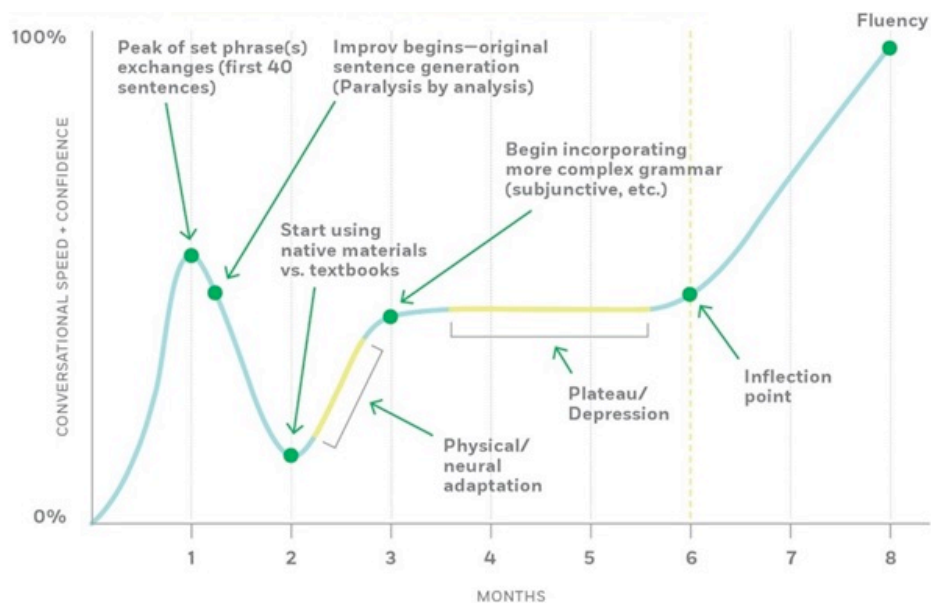


Рисунок 2. Зависимость уверенности в овладении темой от времени обучения из [10]

Повышение мотивации к углубленному изучению программирования.

Многочисленные исследования, обзор которых приведен в [13], продемонстрировали эффективность образовательной технологии обучения «перевернутый класс» – Flipped classroom – (например, методики наставничества) на курсах информатики. В попытке применить активную обучающую методику на лекциях по дисциплине были разработаны вопросы, которые предлагаются на лекциях. Набор состоит из 50 вопросов с несколькими вариантами ответов, на которые студенты отвечают, как индивидуально, так и в группах. Наш опыт преподавания показывает, что активная обучающая методика не повлияла на посещаемость. Отзывы студентов о курсе лекций и до ее применения были в основном положительными.

Неопределенный результат применения методики объясняется несколькими факторами. Во-первых, в классическом виде метод наставничества эффективен для первоначального знакомства с понятиями. Достаточная подготовка перед занятиями является проблемой для студентов старших курсов в применении этой методики, так как обучаемые уже знакомы со всеми понятиями, упомянутыми в глоссарии курса. В результате студенты полагают ненужным готовиться к лекциям, считая, что их знаний достаточно. Во-вторых, метод наставничества ориентирован на учебу друг у друга без вмешательства преподавателя. Роль преподавателя сводится к организации в аудитории дискуссии на тему задаваемых вопросов. В теории методом наставничества учащиеся получают знания посредством наблюдения, изучения, обучения других учащихся и на основе собственного опыта. Проблемой студентов старших курсов является то, что опыт недостаточен, и отсутствует осознание этой недостаточности.

Эффективность вопросов, задаваемых на лекции, сильно зависит от их качества. Вопросы для лекции отличаются по педагогической цели от вопросов для экзамена или домашних заданий [14]. Они должны разрабатываться, чтобы направлять внимание учащихся, стимулировать когнитивные процессы и содействовать формулированию и сопоставлению идей. Вопросы, предлагаемые на лекциях, должны формироваться на основе глоссария курса, направлены на выявление заблуждений и формирование правильных представлений. Очевидно, что все термины глоссария (150 понятий) не могут быть использованы в вопросах. На каждой из 24 лекций времени хватит на обсуждение взаимосвязи не более двух, трех понятий. Разработанные 50 вопросов имеют педагогическую цель демонстрации того, что имеющиеся у обучаемых представления об основных понятиях программирования слишком поверхностны. Они применяются на лекции для того, чтобы снизить уверенность обучаемых в достаточности

их знаний и таким образом повысить мотивацию для восприятия материала лекции. Таким образом, в остальном тактика преподавания остается традиционной. Вопросы к аудитории на лекции имеют цель организации дискуссии, в которой у обучаемых должно появиться сомнение в достаточности их знаний по обсуждаемой проблеме.

Тактика преподавания с использованием вопросов, снижающих уверенность в достаточности знаний, конечно, не может увеличить посещаемость занятий. Однако она помогает удержать тех, кто однажды решил их посетить. При ее применении после нескольких первых лекций число слушателей стабилизируется до конца курса.

Оценка достижения целей курса.

В начале этого курса обучаемые уже знакомы с понятиями, которые будут изучаться. Но этот курс не повторение пройденного. Так как целостность системы знаний в обсуждаемом курсе является важной характеристикой успешности усвоения курса, то нужна специальная методика ее контроля, состоящая из проверки «интеграции знаний». В качестве наиболее просто проверяемого признака целостности знаний используется степень взаимосвязи наиболее важных терминов изученной дисциплины. Предполагается, что чем выше измеряемый показатель целостности системы понятий курса, тем лучше последний усвоен. Проведённая описываемым методом экспериментальная проверка знаний студентов подтверждает обоснованность выдвинутых предположений [15]. В начале курса студенты выполняют тест, в котором задачей является указание взаимосвязей терминов из глоссария курса. Результат теста – это граф, вершинами которого являются понятия глоссария, а ребрами – связи между понятиями. Цель курса состоит в том, чтобы этот граф был односвязным. В конце курса студенты выполняют этот тест повторно. Курс усвоен успешно, если в графе, построенном по результатам повторного теста повышается связность. Этот метод контроля усвояемости курса использовался при чтении курса последние три года. Односвязный граф никогда не возникал в тесте, проводимом в начале курса. В повторном тесте мера связности повышалась у подавляющего большинства студентов.

Заключение.

В дидактике под методом обучения понимают способы достижения цели обучения, способы совместной деятельности преподавателя и студентов. Для студентов старших курсов исследования продемонстрировали эффективность образовательной технологии обучения методом «интеграцией знаний» на курсах информатики. Под «интеграцией знаний» в статье понимается способность представлять всю совокупность связей между понятиями информатики. В данной статье подчёркивается важная роль способности программиста объединять идеи (ability to make links among ideas) при конструировании программы, и что «интеграция знаний» ведёт к глубокому пониманию, поскольку акцентирует связи между научными идеями (the coherence across science ideas).

В работе используется метод, позволяющий получить количественную оценку систематичности концептуального усвоения учебного курса. Метод строится на процедуре учёта связей между понятиями, которые указал студент в ходе проверочного тестирования. В графе результатов тестирования компоненты связности графа группируют связанные по мнению студента между собой понятия в несколько непересекающихся групп. В идеале все термины должны образовывать единую группу, но на практике так не получалось, причём некоторые компоненты связности состоят всего из 2-3 терминов (такие группы уместно считать отдельными фактами, которые студент не связывает с общей картиной). Результаты экспериментов также свидетельствуют, что студенты очень плохо различают типы связей между понятиями: они часто путают классические часть/целое и общее/частное (класс/подкласс), не говоря уже об остальных видах связей.

Список литературы

[1] Черемисинов Д.И. Учебный курс «Интерфейсы информационных систем» в обучении специалистов по информационным и коммуникационным технологиям // Информационные технологии и системы 2016 (ИТС-2016): материалы междунар. науч. конф. (БГУ-ИР, Минск, Республика Беларусь, 26 октября 2016) = Information Technologies

and Systems 2016 (ITS-2016): Proceeding of the Inter-national Conference (BSUIR, Minsk, Republic of Belarus, 26th October 2016) / редкол. : Л. Ю. Шилин [и др.]. - Минск : БГУИР, 2016. – С. 272-273.

[2] Keet, M. C. Experiment with Peer Instruction in Computer Science to Enhance Class Attendance // Huillet, E., Eds. Proceedings 23rd Annual Meeting of the Southern African Association for Research in Mathematics, Science, and Technology Education (SAARMSTE'15) – Maputo, Mozambique, 2018. – P 319-331.

[3] IEEE/AIS/ACM Joint Task Force on Computing Curricula. Computing Curricula 2005. The Overview Report covering undergraduate degree programs in Computer Engineering, Computer Science, Information Systems, Information Technology, Software Engineering. 2005 (<http://www.acm.org/education/curricula.html>).

[4] Рекомендации по преподаванию информатики в университетах / Ред. В. Л. Павлов, А. А. Терехов. – СПб.: Изд-во СПбГУ, 2002. –367 с.

[5] Кознов Д. В. Методика обучения программной инженерии на основе карт памяти // Системное программирование – Т. 3, № 1, 2008. – С. 121-140.

[6] Bryant R. E., O'Hallaron D. R. Computer Systems: A Programmer's Perspective – Prentice Hall, 2002. – 978 p.

[7] Glossary of Computer Related Terms [Electronic resource] / Hohn M. W. 2011. – Mode of access: <http://www.math.utah.edu/wisnia/glossary.html> – Date of access: 1.09.2016.

[8] Glossary of Computer and Internet Terms for Older Adults [Electronic resource] / National Institute on Aging 2016. – Mode of access: <http://www.cheyney.edu/informationtechnology/documents/glossarycomputerterms.pdf> – Date of access: 1.09.2016.

[9] Crouch C. H., and Mazur E. Peer instruction: Ten years of experience and results // American Journal of Physics – v. 69, 2001. – pp. 970-977.

[10] Anzanello M. J., Fogliatto F. S. Learning curve models and applications: literature review and research directions // International Journal of Industrial Ergonomics, 41, 2011. – pp. 573-583.

[11] Ferriss T. Meta Meta-Learning. Frequency: Cramming six months culinary school into 48 hours. // The 4-Hour Chef: The Simple Path to Cooking Like a Pro, Learning Anything, and Living the Good Life – New Harvest; 2012. – pp. 86-87.

[12] Питер Л. Принцип Питера, или, Почему дела идут вкривь и вкось. – АСТ, 2002. – 285 с.

[13] Szabo C., Falkner N., Knutas A., Dorodchi M. Understanding the Effects of Lecturer Intervention on Computer Science Student Behaviour. // ITiCSE-WGR'17, July 3–5, 2017, Bologna, Italy – ACM, New York, NY, USA, 2017. – pp. 105-124.

[14] Beatty I. D., Leonard W. J., Gerace W. J., Dufresne, R. J. Designing effective questions for classroom response system teaching // American Journal of Physics – v. 74(1), 2006. – pp. 31-39.

[15] Еремин Е.А. Количественная оценка целостности системы базовых понятий как мера усвоения учебного материала // Образовательные технологии, 2014, N 4, с.78-98.

ABOUT THE METHODOLOGY OF TEACHING PROGRAMMING AT THE LAST YEARS OF HIGHER EDUCATION

D.I. Cheremisinov

*Leading researcher, United Institute
of Informatics Problems of the
National Academy of Sciences of
Republic of Belarus, PhD of
Technical Sciences, Associate
Professor*

*United Institute of Informatics Problems of the National Academy of Sciences of Belarus, Republic of Belarus
E-mail: cher@newman.bas-net.by*

Abstract. The discipline "Interfaces of information systems" in the curriculum for training specialists in information technology is aimed at creating a body of knowledge about the relationship between key concepts and principles of programming. The course reflects modern concepts and methodological approaches to teaching computer science to senior students of universities. The course material can be considered the core of the structure of computer science concepts, since it is located in the area of overlap of all four of its parts. One of the important tasks of teaching the course is to demonstrate that the ideas formed by students at the initial stages of learning about the basic concepts of programming are too superficial.

Keywords: Computer science education, Information, Interaction, Interface, Learning curve, Peer instruction method.

УДК 339.138.2:004.67

ИСПОЛЬЗОВАНИЕ CRM СИСТЕМ КАК ИНСТРУМЕНТА РАБОТЫ С BIG DATA В МАРКЕТИНГЕ



И. Д. Марковская

Студент инженерно-экономического
факультета БГУИР
markovskayairina02@gmail.com



О.Н. Шкор

Старший преподаватель
кафедры экономики БГУИР
shkor@bsuir.by

И.Д. Марковская

Родилась в 2002 году в Пинске. В 2020 году закончила ГУО «Гимназия №2 г. Пинска». В этом же году поступил в УО «БГУИР», была зачислена на бюджетную форму обучения по специальности «Электронный маркетинг» инженерно-экономического факультета.

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 способы применения Big Data в маркетинге. Описаны основные преимущества сбора и использования большого объёма данных компаниями. Выявлена основная проблема работы с большими данными: правильное структурирование и обработка имеющейся информации. В статье объясняется как полученные данные использовать при запуске E-mail рассылок, таргетированной рекламы и настройке контекстной рекламы. Также в статье рассматриваются CRM-системы как инструмент работы с большими объёмами данных. Чётко сформулированные преимущества использования CRM-систем позволяют понять, кому и как могут быть полезные данные виды систем для управления взаимоотношениями с клиентами.

Ключевые слова: Big Data, большие данные, оптимизация, маркетинг, CRM-системы, данные о клиенте, структурирование данных, связь с клиентом, управление данными, стратегия маркетинга.

Введение.

Как известно, единой целью всех предприятий является получение прибыли в максимальном размере с привлечением минимальных затрат. Для этого необходимо привлекать новых клиентов и поддерживать связь с уже имеющимися постоянными клиентами, повышать лояльность потребителя. Для достижения цели необходимо чётко понимать желания потребителя и предугадывать его поведение. В это помогает постоянный сбор и анализ данных о клиентах. Таким образом, для успешного развития, компаниям приходится работать с большим объёмом данных о потребителях и их взаимодействии с компанией.

Любая компания сегодня располагает крайне разнородными по своей структуре большими данными (Big Data). Их использование, пожалуй, во всём мире только делает первые шаги. Однако, за туманным определением «большие данные» (большие данные – это серия подходов, инструментов и методов обработки структурированных и неструктурированных данных огромных объёмов и значительного многообразия для получения воспринимаемых человеком результатов, эффективных в условиях непрерывного прироста, распределения по многочисленным узлам вычислительной сети [1]) кроется то, что обеспечивает нахождение компаниями удачных управленческих решений и, как следствие, получение лучших позиций на

рынке. Однако из-за обилия и разнородности потоков данных, построить правильные стратегии и проводить грамотные исследования, порой бывает чрезвычайно сложно.

Несмотря на то, что использование больших данных в маркетинге ещё только набирает обороты, уже очевидны некоторые преимущества их использования:

1) Создание наиболее точного портрета целевого потребителя.

Основываясь на тех данных, которые были получены от потребителей можно аналитическим путём выявить схожие характеристики для портрета целевого потребителя, что в дальнейшем можно использовать при разработке стратегии маркетинга для компании.

2) Предсказание реакции потребителей на маркетинговые «сообщения» и предложения того или иного продукта.

Наличие большого объёма данных о потребителях позволяет наиболее точно определить их поведение и реакцию на рекламные сообщения о товаре.

3) Персонализация рекламных сообщений.

Правильная систематизация данных о потребителях позволят провести более детальную сегментацию рынка или даже выделить группы потребителей со схожими потребностями и поведением. Таким образом появляется возможность рассылки рекламных сообщений персонализировано для узкого целевого сегмента, которые с большей вероятностью заинтересуют потребителя, чем шаблонные общие рассылки. Этот способ оптимизации большого объёма данных активно используется в E-mail рассылках.

4) Оптимизация производства и стратегий распределения.

Маркетологи могут использовать большие данные для определения оптимальных каналов размещения собственной продукции, и, соответственно, настройки цепочек поставок.

5) Сохранение большего числа клиентов путём наименьших затрат.

С каждым днём появляется всё больше и больше каналов продвижения. Потенциальные клиенты могут часто переходить с одного канала на другой. Данные, собранные о клиентах, и правильная их интерпретация позволят использовать бюджет экономно и вкладывать его в те каналы, которые приносят максимальную выгоду для компании.

6) Получение лучшего представления о собственном продукте компании и разработка новых товаров.

Постоянное изучение данных о прошлом является основой для вычисления вероятности события в будущем. Если у вас есть огромный объем информации, предиктивный анализ может помочь при внедрении нового продукта или услуги. Основной проблемой при работе с большим объемом данных является способ их структурирования и обработки. Данную проблему решает использование CRM систем. CRM (customer relationship management) – это способ управления взаимоотношениями с клиентами и оптимизации бизнес-процессов [3].

Ключевой составляющей данного подхода является CRM-система – специальное программное обеспечение для организации работы с лидами, отслеживания действий клиентов автоматизации коммуникаций.

CRM-система собирает данные о каждом клиенте в одной панели управления. Это та же таблица с клиентскими контактами, но только при нажатии на имя открывается карточка с подробной информацией о ваших взаимодействиях: от первого СМС и до момента покупки. Здесь видно историю покупок, контроль расходов и напоминания о встречах. С помощью этой информации компания может отслеживать путь покупателя и делать релевантные предложения на каждом этапе. В результате цикл продаж сократиться на 8-14%. CRM помогает определить интересы и предпочтения ваших клиентов. Это позволяет предоставить персонализированный опыт и создавать релевантные маркетинговые кампании. Согласно исследованиям Captterra, компании, которые используют CRM маркетинг, отмечают увеличение удержания клиентов и повышение удовлетворённости на 47% [4].

Главной целью систематизации данных о покупателях является упрощение связи с клиентами. При обращении покупателя автоматически создаётся карточка, отражающая:

- контактную информацию: телефон, почту, ссылки на социальные сети;
- чат с менеджером, письма из почты, звонки, сообщения из сторонних источников;
- действия на сайте, просмотренные вещи и услуги;
- купленные товары, платежи, способы получения и оплаты покупки.

Собранные данные позволяют чётко сформулировать условия для таргетированной рекламы в интернете, а также настройки контекстной рекламы, таким образом можно использовать денежные средства компании с максимальной выгодой и показывать рекламу только тем пользователям, которым она будет действительно интересна.

Качество работы компании повышается из-за уменьшения действий, которые выполняют сотрудники. Менеджерам видна полная информация о человеке в единой базе, расположенной в облаке. Для быстрого ответа не нужно переключаться между десятками вкладок. Сегментируя потенциальных покупателей, организация может сделать персональную рассылку для каждой группы.

Ещё одним преимуществом использования CRM-систем для обработки и структурирования большого объёма данных является формирование системой подробных отчётов со статистикой. С её помощью можно определить причины потери покупателя, выявить какие способы взаимодействия с клиентом более эффективны. Получаемая статистика позволяет также оценить эффективность и уровень нагрузки сотрудников. CRM-системы используются для просмотра действий клиентов на сайте. Получив информацию, можно выстроить маркетинговую стратегию, устранить слабые стороны рекламных кампаний и отредактировать воронки продаж.

При помощи CRM-систем возможна оптимизация рабочих процессов, к целям которой относятся:

- 1) Возможность перенести все процессы на дистанционное управление.

Через сервис руководитель может ставить различные задачи и контролировать их выполнение.

- 2) Коммуникация между всеми сотрудниками компании.

В системе доступна возможность создание открытых и закрытых чатов, проведение видеоконференций.

- 3) Защищённое хранение документации.

Система работает в облаке с закрытым доступом, где право на просмотр и изменения можно настроить для каждого сотрудника. Это защищает данные компании от утечек. CRM поддерживает интеграцию с множеством сервисов. Процесс подключения сторонних программ увеличивает функционал. Немало важное значение имеют CRM-системы и для маркетинговых исследований. На основе уже сформированной статистики можно сделать некоторые выводы полезные для исследования, или подобрать группу покупателей, наиболее подходящую для проведения опроса или фокус группы. Так же наличие всех данных о клиентах позволят провести опрос в максимально сжатые сроки и с минимальными затратами. Исходя из указанных выше преимуществ, очевидно, что с каждым днём количество компаний используемых CRM-системы будет только расти. На такую технологию стоит обратить внимание малому и среднему бизнесу, поскольку это позволит улучшить качество обслуживания клиентов, снизить нагрузку на сотрудников, а также получать большое количество отчётности, на основе которой можно более точно настроить различные виды интернет-рекламы, тем самым разумно распределить бюджет на продвижение и получить максимальную прибыль при минимальных затратах.

Заключение.

Интернет сегодня является основным источником больших данных. Все, начиная от веб-сайтов до аналитики социальных медиа, перехода по ссылкам рекламных объявлений может быть объединено, проанализировано и интерпретировано при помощи различных систем. Одной из таких систем является CRM-система, которая значительно упрощает работу с большими объёмами данных, позволяет структурировать их и анализировать. Данная система позволяет лучше понимать потребителя и успешно разрабатывать стратегии маркетинга. Поскольку

преимущества использования CRM-систем очевидно, то стоит ожидать значительный рост в использовании CRM-систем малым и средним бизнесом.

Список литературы

[1] Радченко И.А. Технологии и инфраструктура Big Data: учеб. пособие / И. А. Радченко, И.Н. Николаев – Санкт-Петербург: Университет ИТМО, 2018. - 55 с.

[2] Big Data и маркетинг // Журнал Деловой Мир [Электронный ресурс]. – Режим доступа: <https://delovoyimir.biz/big-data-i-marketing-5-metodov-ispolzovaniya-s-realnymi-keysami.html>

[3] CRM [Электронный ресурс]. – Режим доступа: <https://sendpulse.by/support/glossary/crm>

[4] Customer Data Management in Marketing: What You Need to Know [Электронный ресурс]. – Режим доступа: <https://www.capterra.com/resources/data-management-in-marketing/>

THE FUTURE OF MARKETING: AUTOMATING CUSTOMER INTERACTION WITH THE HELP OF THE INTERNET

I. D. Markouskaya

Student of engineering and economics at the BSUIR

O.N. Shkor

Senior Lecturer at the Department of Economics at the BSUIR

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

E-mail: markovskayairina02@gmail.com

E-mail: shkor@bsuir.by

Abstract. The article discusses the ways of using Big Data in marketing. The main advantages of collecting and using a large amount of data by companies are described. The main problem of working with big data is revealed: the correct structuring and processing of the available information. The article explains how to use the received data when launching E-Mail marketing, targeted advertising and setting up contextual advertising. The article also discusses CRM-systems as a tool for working with large amounts of data. These systems allow us to automate the work of company employees with customers through good structuring of data and their convenient presentation. The clearly articulated benefits of using CRM systems make it possible to understand to whom and how these types of systems for customer relationship management can be useful.

Keywords: Big data, large amounts of data, optimization, marketing, CRM systems, customer data, data structuring, customer communication, data management, marketing strategy.

УДК 004.855.5

СНИЖЕНИЕ ВЛИЯНИЯ УТЕЧКИ ДАННЫХ НА ТОЧНОСТЬ МОДЕЛЕЙ В МАШИННОМ ОБУЧЕНИИ



В.П. Корячко

Заведующий кафедрой систем автоматизированного проектирования РГРТУ им. В.Ф. Уткина, доктор технических наук, профессор
koryachko.v.p@rsreu.ru



В.И. Орешиков

Доцент кафедры систем автоматизированного проектирования РГРТУ им. В.Ф. Уткина, кандидат технических наук, доцент
vyacheslav.oreshkov@yandex.ru

В.П. Корячко

Окончил Рязанский радиотехнический институт. Область научных интересов связана с применением технологий машинного обучения в системах автоматизированного проектирования, разработкой методов представления знаний с использованием нечёткой логики и мягких вычислений.

В.И. Орешиков

Окончил Рязанскую государственную радиотехническую академию. Область научных интересов связана с использованием технологий искусственного интеллекта в системах автоматизированного проектирования, машинным обучением, проектированием интеллектуальных информационных систем.

Аннотация. Рассмотрена проблема утечки данных в задачах машинного обучения и выявлены связанные с ней факторы, негативно влияющие на точность аналитических моделей на этапе их практического применения. Предложены методы обнаружения утечек данных и их предотвращения. Проведены экспериментальные исследования возможности устранения потерь точности моделей, основанных на машинном обучении, связанных с утечкой данных, на примере нейронной сети в задаче предсказания урожайности зерновых по данным агрохимического обследования почв.

Ключевые слова: интеллектуальный анализ данных, машинное обучение, аналитическая модель, обучающие данные, тестовые данные, утечка данных, нейронная сеть.

Введение.

Одной из основных групп методов обработки данных в технологиях Big Data является интеллектуальный анализ данных (Data Mining), в рамках которого могут быть решены задачи численного предсказания (регрессии), классификации, кластеризации, ассоциации, прогнозирования и т.д. Эти задачи являются основой многих сложных решений в области управления и бизнеса [1].

Наиболее популярным подходом к решению перечисленных задач является машинное обучение (machine learning - ML), а именно построение моделей, основанных на машинном обучении (ML-моделей). В таких моделях происходит автоматическая подстройка параметров в процессе обучения на некоторой выборке данных (называемой обучающей или тренировочной). Модель в процессе обучения «извлекает» из данных скрытые зависимости, закономерности и шаблоны, интерпретация которых человеком (аналитиком, специалистом предметной области) позволяет получать новые, практически полезные знания для целей поддержки принятия решений [2].

Хотя машинное обучение в настоящее время рассматривается как технологическое и инструментальное ядро Data Mining и Big Data, с обучением и практическим использованием DM-моделей связан целый комплекс проблем, игнорирование которых может привести к

некорректно работающим моделям и ошибочным решениям. К основным проблемам относятся эвристический характер многих алгоритмов обучения, возможность переобучения *ML*-моделей, завышенная оценка их точности.

Эвристический характер алгоритмов машинного обучения обусловлен тем, что они часто не имеют под собой строгой математической основы, не гарантируют получения единственного и точного решения, и даже решения вообще, хотя и обеспечивают приемлемый результат в большинстве практически значимых случаев. Действительно, градиентные алгоритмы обучения нейронных сетей могут «застрять» в локальном минимуме целевой функции, а начальные значения весов нейронов устанавливаются случайным образом и, как следствие, единственное решение не может быть получено даже на одном и том же наборе обучающих данных. Вместе с тем, эвристические алгоритмы обучения позволяют получить приемлемое решение даже в условиях неполных и даже частично искажённых данных, когда строгие алгоритмы вообще не работают.

Эффект переобучения (*overfitting*) *ML*-моделей связан со слишком точной подстройкой параметров модели под обучающие данные, что приводит к снижению её точности при практическом использовании на новых данных, не использовавшихся в процессе обучения. Причиной переобучения может стать некорректный выбор архитектуры *ML*-модели (когда число её параметров оказывается сравнимым с числом обучающих примеров) или слишком большое число итераций обучения. Для борьбы с переобучением используются тестовое множество и перекрёстная проверка (крсс-валидация).

Следует отметить, что проблемы, связанные с эвристическим характером *ML*-моделей достаточно хорошо изучены, освещены в литературе, а инструменты для их решения включаются в большинство программных средств анализа данных. Вместе с тем, при решении практических задач в различных предметных областях возникают и другие проблемы, связанные с машинным обучением и практическим применением *ML*-моделей. К одной из таких проблем, на которую в сообществе аналитиков данных стали обращать пристальное внимание в последнее время, относится утечка данных (*data leakage*).

Актуальность.

Утечку данных называют одной из десяти основных проблем интеллектуального анализа данных и машинного обучения [3]. Под утечкой данных в машинном обучении понимается ситуация, когда при построении *ML*-модели используются данные, которые оказываются недоступными на этапе её практического использования. Несмотря на то, что проблема утечки данных способна существенно испортить жизнь аналитикам, в настоящее время она недостаточно широко отражена в литературе

Следует отметить, что утечке в наибольшей степени подвержены данные, формируемые во внешнем окружении организации, доступность которых на зависит от самой организации. Однако, если такие данные игнорировать, потенциально ценная информация из внешнего окружения останется неиспользуемой при обучении модели.

Действительно, эксплуатация любой *ML*-модели производится в двух режимах: обучения (*training*) и предсказания (*prediction*). В процессе обучения выполняется настройка параметров модели с использованием обучающего множества по определённому алгоритму обучения. В режиме предсказания модель формирует (предсказывает) значения целевой переменной для новых наблюдений, которые не использовались на этапе обучения.

Пусть, например, модель обучается на наборе данных, в котором присутствует поле «Возраст клиента». Очевидно, что это важная информация с точки зрения предсказания, например, кредитоспособности заёмщика в кредитном скоринге, или уровня лояльности клиента в маркетинге. Но может случиться так, что в режиме предсказания данные о возрасте окажутся недоступными. В результате информация, используемая в режиме обучения не может быть использован при предсказании. Это и называется утечкой данных. Ожидаемым результатом утечки данных будет снижение точности предсказания модели на практических данных,

относительно точности предсказания на обучающих и тестовых (где «утекшая» переменная присутствует). Причины данного явления интуитивно понятны: при предсказании используется меньше информации, чем при обучении. Такой вид утечки называется «целевой утечкой» (target leakage). Целевая утечка имеет место, когда при построении модели (режим обучения) используется информация, которая не будет доступна в процессе её практического использования (режим предсказания). Целевые утечки необходимо рассматривать в контексте временных периодов доступности данных, а не только той пользы, которые данные могут принести.

Ещё одной причиной, которая может привести к эффекту утечки данных, является их предобработка, а именно нормализация, масштабирование, сглаживание, квантование и другие методы, которые приводят к изменению значений данных. Применение предобработки позволяет сделать процесс обучения более эффективным, а модель более точной но только на обучающих данных. Но при этом «подгонка» модели будет производиться именно к изменённым данным, а не к реальным, которые, вероятнее всего будут использоваться на этапе практического использования. Поэтому качество модели на обучающих данных, окажется переоцененным.

Данный тип утечки в литературе иногда называют «утечкой обучающего/тестового множества» (в англоязычном варианте train-test contamination). Чтобы обнаружить этот вид утечки следует проверять модель на тестовом множестве, которое не было подвержено предобработке. Другой вариант - построить модель на предобработанных данных, а затем оценить её качество на части данных, которые не были подвергнуты предобработке. Например, если в процессе предобработки производится подстановка пропущенных значений, то модель может показать хорошие результаты при тестировании, но плохие при работе с практическими данными.

И, наконец, если проверочные или тестовые данные «протекают» в обучающие, то оценка качества модели при валидации окажется завышенной. Поэтому если валидация модели строится на простом разделении исходного набора данных на обучающее и тестовое множества, то важно позаботиться, чтобы проверочные данные не попали в обучающие, в том числе и на этапе предобработки.

Таким образом, утечка данных является значимой проблемой анализа данных с использованием *ML*-моделей и требует разработки и использования методов обнаружения утечек и минимизации их негативного влияния на результаты.

Подходы к предотвращению утечки данных.

Наиболее очевидным подходом к предотвращению утечки данных является организационный. Т.е. управление данными в процессе их интеллектуального анализа должно выстраиваться так, чтобы не допустить утечки. Здесь можно выделить два подхода:

1. Не использовать при обучении модели переменные, которые потенциально могут оказаться недоступными в процессе её практического использования. Недостатки подхода очевидны: мы изначально получим худшую модель, чем могли бы получить, поскольку сознательно отказываемся от части информации, используемой в процессе обучения.

2. Использовать все доступные переменные, в том числе и те, которые могут оказаться подвержены утечке, построив лучшую модель. Затем, если утечка произойдёт, принять меры к организации сбора «утекших» данных. Недостатками подхода является то, что издержки на сбор данных могут превысить выгоду от их использования, а также то, что такие данные могут оказаться в принципе недоступными или несуществующими. Тривиальным примером такой ситуации является использование целевой переменной в качестве входной при обучении с учителем. Очевидно, что значения целевой переменной в режиме предсказания в принципе не могут быть известны. Другой пример из медицины: если в модели используются данные собранные по выборке ранее наблюдаемых пациентов и среди них есть температура, то новые пациенты просто могут оказаться недоступными для измерения температуры.

3. Восстановление «утекших» данных на основе доступных в обучающей выборке.

Например, заменять при предсказании недостающее значение в новом наблюдении на значение, искусственно сгенерированное из распределения ранее известных наблюдений. Т.е. «утекшее» значение можно рассматривать как пропущенное и восстанавливать его одним из существующих методов. Однако, здесь есть опасность, что со временем распределение данных меняется (скажем, уровень дохода дрейфует вверх или вниз), в результате вместо пропуска мы получим аномалию.

4. Рассматривать утечку данных как неизбежное зло и смириться с ней, если последствия не наносят неприемлемого ущерба.

Одной из главных проблем, связанных с утечкой данных, является то, что её бывает трудно обнаружить и устранить и она приводит, как минимум, к переоценке качества модели.

Экспериментальные исследования.

Исследования процессов утечки данных авторами работы проводились на нескольких обучающих выборках, связанных с решением практических народнохозяйственных задач. Одной из таких задач было предсказание урожайности зерновых культур (ячменя) на основе данных агрохимического обследования почв [4,5] с использованием нейронной сети. Обучающая выборка содержала следующие признаки, представленные в табл. 1.

Таблица 1. Список признаков обучающей выборки

№ п/п	Наименование	Пояснение	Ед. изм.
1	Площадь	Площадь поля	Га
2.	Кислотность	Средняя кислотность почвы поля	pH
3.	Азот	Содержание азота в почве	Мг/100 г.
4.	Калий	Содержание калия в почве	Мг/100 г.
5.	Фосфор	Содержание фосфора в почве	Мг/100 г.
6.	Угол	Средний угол уклона пашни к югу	градус
7.	Уклон	Доля пашни с уклоном к югу	%
8.	Урожайность	Фактическая урожайность поля	ц/га

Признаки со 2-го по 6-й использовались в качестве входных переменных нейросетевой модели. Урожайность - известное значение урожайности, зафиксированное для полей с наблюдаемыми агрохимическими характеристиками. В процессе обучения нейросети урожайность использовалась в качестве целевой переменной. При практическом использовании модели урожайность должна предсказываться сетью для новых полей.

Результаты применения модели позволили решить следующие задачи:

- оценивать будущую урожайность новых полей, для которых собраны данные агрохимического обследования почвы с целью принятия решения о целесообразности включения их в севооборот;

- изучить зависимость урожайности от агрохимических характеристик с целью разработки агротехнологических мероприятий, направленных на повышение урожайности.

Обучение нейросетевой модели производилось с применением аналитической платформы Deductor (в настоящее время Loginom) российской компании BaseGroup Labs (в настоящее время Loginom). В качестве базовой нейросетевой архитектуры использовалась плоскостная сеть прямого распространения с обучением по методу обратного распространения ошибки с оптимизацией параметров алгоритма (крутизны активационной функции, коэффициента скорости обучения и момента) [6]. Граф используемой нейросети представлен на рис. 1.

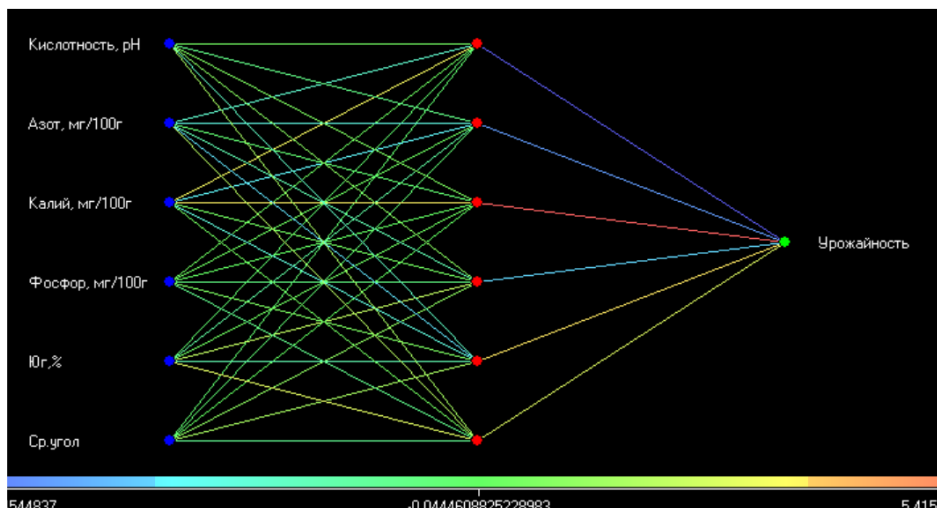


Рисунок 1. Граф нейросети для предсказания урожайности

Обучающий набор данных содержал 150 наблюдений, из которых 90% были случайным образом отобраны в обучающее множество, а 10% в тестовое. Предобработка обучающего и тестового множеств не производилась (т.е. данные использовались для обучения «как есть»). Фрагмент визуализатора с результатами обучения представлен на рис. 2.

№ поля	Площадь, га	Кислотность, pH	Азот, мг/100г	Калий, мг/100г	Фосфор, мг/100г	Юг.%	Ср. угол	Урожайность	Урожайность_OUT
30	81,95	5,1	4	19,59	17,56	33,02	0,7	3,6	2,9
47	128,4	5,9	8	11,98	25,19	46,99	1,15	9,7	9,1
15	191	6	9	17,78	19,74	14,34	1,21	11,6	11,3
19	58,26	6,1	9	14,42	23,28	43,6	1,01	11,6	11,3
17	17,7	4,4	1	6,8	10	46,39	2,32	0,4	0,7
39	21,65	4,7	2	8,55	13,88	7,52	0,91	1,1	1,4
55	24,68	6,1	9	19	20,5	14,47	1,28	11,6	11,4
3	47,52	5,6	6	12,76	16	44,24	0,92	6,3	6,5
40	60,13	5	3	10,25	19,83	5,63	1,03	2,2	2,0
37	21,65	4,7	2	8,55	13,88	7,52	0,91	1,2	1,4
34	14,85	5,1	4	15,59	15,71	62,63	0,82	3,5	3,3
7	18,15	5	3	7,4	6	1,33	1,78	2,2	2,1
33	82,05	4,8	3	10,59	9,75	2,41	0,81	2,3	2,2
20	84,62	5,4	5	8,64	9,33	0,06	1,01	4,8	4,9
26	30,85	6,1	9	17,4	25,07	99,75	1,37	11,6	11,5
36	118,96	5,1	4	15,59	15,71	18,05	0,81	3,4	3,5
35	17,24	4,8	3	10,59	9,75	3,51	0,55	2,1	2,2
53	151,65	4,8	3	10,89	13,23	5,84	1,07	2,2	2,1
23	26	5,9	8	18,58	20,5	0	1,46	9,7	9,8
6	100,06	5,2	4	14,89	10,8	20,05	1,51	3,4	3,5
54	38,23	5,6	6	13,08	25,2	0,08	1,66	6,3	6,4
18	79,67	4,9	3	14,23	14,21	0,9	1,07	2,2	2,3
12	128,73	6,1	9	6,91	16,11	50,15	1,28	5,6	5,7

Рисунок 2. Фрагмент представления с результатами обучения

Для оценивания точности модели можно использовать среднеквадратическую ошибку на выходе сети:

$$E = \frac{1}{N} \sum_{i=1}^N (Y - \hat{Y})^2 = 0,036,$$

где N - число обучающих примеров, Y - наблюдаемое значение целевой переменной, \hat{Y} - значение целевой переменной, предсказанное моделью. Значение, предсказанное сетью, расположено в поле Урожайность_OUT.

После того, как модель была передана в практическую эксплуатацию выяснилось, что в силу организационных проблем данные по признакам 6 и 7 перестали быть доступными. Т.е. при формировании предсказания урожайности для новых полей информация, связанная с уклоном

пашни к югу перестала учитываться. Поскольку для новых полей целевое значение урожайности неизвестно, произвести расчёт среднеквадратической ошибки для оценки того, насколько снизилась точность модели, не представляется возможным. Однако косвенно оценить снижение точности для модели можно с использованием тестового множества, из которого исключены значения «утекших» признаков.

В результате, расчёты показали, что среднеквадратическая ошибка после «утечки» двух признаков составила $E = 0,042$, т.е. увеличилась на 17,4 %. Таким образом качество модели на обучающем наборе оказалось значительно переоцененным. Чтобы компенсировать результаты "утечки" было предложено ввести в модель на этапе предсказания фиктивные значения "утекших" признаков, вычисляемые на основе известных значений обучающего множества. Т.е. когда на вход модели подаётся новое наблюдение, вместо «утекших» значений на соответствующие входы поступают значения, сгенерированные по заданному алгоритму. Используемые алгоритмы и результаты их использования приведены в таблице 2.

Таблица 2. Таблица результатов устранения утечки данных

№ п/п	Наименование	Среднеквадратическая шибка	% ухудшения
1	Без устранения	0,042	17,4
2.	Среднее	0,041	13,8
3.	Медиана	0,040	11,1
4.	Замена наиболее вероятным	0,038	5,5
5.	Замена случайным	0,041	13,8

Таким образом, в данном случае эксперимент показал, что ухудшение точности модели вследствие утечки данных имеет место в любом случае и полностью устранить её последствия предложенными методами не удалось.

Наилучший результат достигнут при замене значений «утекших» признаков на наиболее вероятное значение и составляет $E = 0,038$, т.е. ухудшение точности составляет 5,5%, что, тем не менее даёт определённый выигрыш по сравнению с ситуацией, если на "утечку" никак не реагировать.

Наиболее вероятное значение оценивается по формуле Бернулли:

$$P_n(m) = C_n^m \cdot p^m \cdot q^{n-m},$$

где $P_n(m)$ - вероятность появления значения m на числе наблюдений n , C_n^m - количество сочетаний из n по m , p и q - вероятности появления и не появления произвольного значения ($p=1-q$).

Выводы.

Показано, что при возникновении эффекта утечки данных при обучении и практическом использовании ML -моделей (на примере нейросети) существуют возможность частично компенсировать потери точности модели с помощью замены значений «утекших» признаков при работе модели в режиме предсказания, на значения, сгенерированные на основе доступных значений из обучающих данных. В экспериментальном примере удалось достичь снижения потери точности модели с 17,4% в случае, если никакие меры не предпринимаются, до 5,5% при использовании замены значений "утекших" признаков на наиболее вероятные по обучающему набору.

Следует отметить, что приведённую методику не следует рассматривать как общую, которая будет работать во всех случаях. Это связано с тем, что результаты мероприятий, направленных на предотвращение последствий утечки данных будут зависеть от особенностей самой задачи, решаемой с помощью ML -модели, вида самой модели, природы обучающих данных и характера самих утечек. Тем не менее, приведённый пример показывает, что потери точности

ML-моделей, связанные с утечкой данных могут в определённой степени компенсироваться, поэтому оставлять утечки данных без внимания не следует.

Заключение.

Таким образом в работе произведён анализ причин и последствий такого явления как утечка данных в моделях, основанных на машинном обучении, и предложен ряд методов, которые могут способствовать снижению потери точности моделей при её работе с новыми данными, не использовавшимися в процесс обучения.

В эксперименте, проведённом с использованием нейросетевой модели для оценивания потенциальной урожайности зерновых на основе данных агрохимического обследования почв показано, что потери точности модели, связанные с эффектом утечки данных могут быть частично скомпенсированы путём замены значений признаков, подвергшихся утечке, фиктивными значениями, сгенерированными на основе известных значений признаков обучающего набора данных. Наилучшие результаты были получены при использовании замены на наиболее вероятное значение по обучающей выборке.

Список литературы

- [1] Паклин Н.Б. Бизнес-аналитика: от данных к знаниям (+ CD): учеб. пособие. / Паклин Н.Б., Орешков В.И. 2-е изд., испр. – СПб.: Питер, 2013. – 704 с.
- [2] Корячко В.П. Интеллектуальные системы и нечеткая логика / В.П. Корячко, М.А. Бакулева, В.И. Орешков. – М.: КУРС, 2017. – 346 с.
- [3] S. Kaufman. Leakage in Data Mining: Formulation, Detection, and Avoidance / S. Kaufman, S. Rosset, C. Perlich. KDD'11, August 21 – 24, 2011, San Diego, California, USA.
- [4] Васильев Е.П. Интеллектуальные системы бизнес-аналитики в сельскохозяйственном производстве / Васильев Е.П., Евстропов А.С., Орешков В.И. Проблемы механизации агрохимического обслуживания сельского хозяйства. 2011. № 2. С. 189-208.
- [5] Васильев Е.П. Моделирование урожайности на основе данных агрохимического обследования почв с помощью метода ассоциативного анализа / Васильев Е.П., Орешков В.И. – Вестник РГАТУ. – 2012 – № 4 (16) – С. 8 - 13.
- [6] В.И. Орешков. Интеллектуальный анализ данных: учеб. пособие / Орешков В.И. – Изд-во Рязан. гос. радиотехн. ун-та. Рязань, – 2016. 160 с.

REDUCING THE IMPACT OF DATA LEAKAGE ON MODEL ACCURACY IN MACHINE LEARNING

V.P. Koryachko

*Head of the Department of
Computer-Aided Design Systems,
Ryazan State Radio Engineering
University named after V.F. Utkin,
Doctor of Technical Sciences,
Professor*

V.I. Oreshkov

*Associate Professor of the
Department of Computer-Aided
Design Systems, Ryazan State Radio
Engineering University named after
V.F. Utkin, candidate of technical
sciences*

*Department of Computer-Aided Design Systems
Faculty of Computer Science
Ryazan State Radio Engineering University named after V.F. Utkin
E-mail: vyacheslav.oreshkov@yandex.ru*

Abstract. The problem of data leakage in machine learning problems is considered and the factors associated with it that negatively affect the accuracy of analytical models at the stage of their practical application are identified. Methods for detecting data leaks and preventing them are proposed. Experimental studies have been carried out on the possibility of eliminating the loss of accuracy of models based on machine learning associated with data leakage, using the example of a neural network in the problem of predicting grain yields based on agrochemical soil survey data.

Keywords: data mining, machine learning, analytical model, training data, test data, data leakage, neural network.

УДК 004.65:378

ПРИМЕНЕНИЕ БЛОКЧЕЙНА И БОЛЬШИХ ДАННЫХ В СФЕРЕ ОБРАЗОВАНИЯ



Т.В. Казак

*Заведующий кафедрой инженерной психологии и эргономики, доктор психологических наук Республики Беларусь, доктор психологических наук Российской Федерации, член-корреспондент Международной академии психологических наук, профессор,
kazak@bsuir.by*



А.А. Войтович

Студент кафедры инженерной психологии и эргономики



А.Н. Василькова

*Ассистент кафедры инженерной психологии и эргономики, магистр.
a.vasilkova@bsuir.by*

Т.В. Казак

Научный руководитель, заведующий кафедрой инженерной психологии и эргономики, доктор психологических наук Республики Беларусь, доктор психологических наук Российской Федерации, член-корреспондент Международной академии психологических наук, профессор.

А.А. Войтович

Студент кафедры инженерной психологии и эргономики УО «Белорусский государственный университет информатики и радиоэлектроники».

А.Н. Василькова

Ассистент кафедры инженерной психологии и эргономики УО «Белорусский государственный университет информатики и радиоэлектроники», магистр.

Аннотация. Работа посвящена исследованию возможностей применения технологии блокчейн в образовании с учетом анализа данных большого объема (big data). Работа рассматривает преимущества использования блокчейна для обеспечения прозрачности и безопасности данных студентов, а также для оптимизации процессов оценки и анализа образовательных результатов. В работе также рассматриваются конкретные примеры реализации блокчейна в образовании, анализируются их преимущества и недостатки, а также перспективы дальнейшего развития данной технологии в образовательной сфере. Работа является актуальной в свете необходимости повышения эффективности образовательных процессов и использования современных технологий в образовании.

Ключевые слова: большие данные, блокчейн, образование, аналитика, студенты.

Введение.

Современное общество сталкивается с быстрым развитием информационных технологий, которые активно применяются во многих областях деятельности. Среди таких областей особое место занимает образование, где технологии могут применяться для улучшения качества образовательного процесса, повышения доступности обучения, а также для обеспечения более эффективной работы учебных заведений.

Одной из самых перспективных технологий, которые могут быть применены в сфере образования, является блокчейн. Эта технология позволяет создавать безопасные и прозрачные системы хранения и передачи данных, что может быть весьма полезно в образовательных учреждениях. Блокчейн может помочь улучшить процессы аттестации и сертификации,

повысить безопасность хранения личных данных студентов и улучшить управление финансами учебных заведений.

В совокупности с блокчейном, большие данные (big data) также могут стать важным инструментом для оптимизации образовательных процессов. Анализ больших данных может помочь выявить закономерности и тенденции в обучении, что в свою очередь позволит оптимизировать учебные программы и методы преподавания. Кроме того, большие данные могут использоваться для создания персонализированных образовательных программ, которые будут адаптированы под потребности каждого конкретного студента.

Цель данной научной работы заключается в рассмотрении возможных способов применения блокчейна и больших данных в сфере образования и анализе их эффективности. В ходе исследования будут рассмотрены примеры реализации таких технологий в образовательных учреждениях и оценены их результаты. Работа может быть полезна для руководителей учебных заведений, преподавателей и других специалистов, работающих в области образования, которые заинтересованы в внедрении новых технологий и повышении качества образовательных услуг.

Основная часть.

Big Data - это множество данных, которые по своему объему, скорости и разнообразию выходят за рамки традиционных методов их обработки и анализа. Эти данные могут поступать из различных источников, включая социальные сети, электронную почту, сенсоры, мобильные устройства и другие.

Образовательные учреждения также сталкиваются с большим объемом данных, которые необходимо обработать и анализировать. Эти данные могут включать в себя информацию о студентах, преподавателях, курсах, успеваемости и т.д. Важно, чтобы эти данные были не только безопасными, но и доступными и понятными для использования.

Как блокчейн может помочь в обработке Big Data в образовании?

Блокчейн может помочь в обработке Big Data в образовании, предоставляя надежный и безопасный способ хранения и обмена данными. Кроме того, блокчейн может обеспечить прозрачность и улучшить качество данных. Хранение данных. Одним из ключевых преимуществ блокчейна является его способность обеспечить безопасное хранение данных. Блокчейн предоставляет децентрализованную систему хранения, которая позволяет хранить данные на нескольких узлах сети, что делает данные более устойчивыми к атакам и сбоям.

Кроме того, блокчейн использует криптографические методы для защиты данных, что делает их невозможными для подделки или изменения. Это особенно важно в образовании, где данные оценок и документов могут иметь серьезные последствия для учащихся и их будущей карьеры.

Прозрачность данных. Еще одним преимуществом блокчейна в обработке Big Data является его способность обеспечить прозрачность данных. Блокчейн предоставляет надежный и прозрачный способ хранения данных, который позволяет всем пользователям просматривать данные в режиме реального времени. Это особенно важно для образовательных учреждений, которые должны обеспечивать прозрачность в отношении данных о студентах и преподавателях. Например, блокчейн может использоваться для хранения и обмена данными о курсах, стипендиях, финансовой помощи и т.д., что позволяет студентам и преподавателям получать доступ к актуальным данным в режиме реального времени.

Улучшение качества данных. Блокчейн может также помочь улучшить качество данных в образовательных учреждениях. Блокчейн может использоваться для создания системы проверки данных, которая гарантирует точность и целостность данных.

Кроме того, блокчейн может использоваться для создания системы контроля качества данных, которая позволяет обнаруживать и исправлять ошибки и неточности в данных. Это особенно важно для данных, которые используются для принятия решений, таких как данные о

студентах, которые используются для принятия решений о наборе студентов, преподавателях и курсах.[1]

Применение блокчейна в образовании.

Существует множество способов, которыми блокчейн может быть использован в образовании для обработки Big Data. Некоторые из них описаны ниже:

Хранение академических записей. Блокчейн может использоваться для хранения академических записей студентов, что позволяет иметь быстрый и надежный доступ к актуальным данным о студентах. Это также облегчает процесс передачи данных между учреждениями и работодателями.

Оценка учебных достижений. Блокчейн может использоваться для оценки учебных достижений студентов и преподавателей. Например, блокчейн может использоваться для создания системы оценки, которая гарантирует точность и целостность данных.

Управление финансами. Блокчейн может использоваться для управления финансами в образовательных учреждениях. Например, блокчейн может использоваться для учета стипендий и финансовой помощи, а также для управления бюджетом учреждения. Это позволяет учреждениям лучше контролировать свои финансы и предоставлять студентам более эффективную финансовую поддержку. Улучшение процесса набора студентов. Блокчейн может использоваться для улучшения процесса набора студентов. Например, блокчейн может использоваться для создания системы проверки достоверности информации, предоставляемой студентами при подаче заявлений на поступление. Это помогает улучшить качество данных и предотвратить мошенничество.

Управление исследованиями и разработками. Блокчейн может использоваться для управления исследовательскими проектами и разработками в образовательных учреждениях. Например, блокчейн может использоваться для создания системы учета и отслеживания результатов исследовательских проектов, что позволяет учреждениям эффективнее использовать свои ресурсы и улучшить качество своих исследовательских проектов.

Создание децентрализованных систем управления образованием. Блокчейн может использоваться для создания децентрализованных систем управления образованием, что позволяет учреждениям и преподавателям более эффективно управлять образовательным процессом. Например, блокчейн может использоваться для создания системы учета и контроля учебных планов, расписаний и учебных материалов.

Возможные риски.

Несмотря на многочисленные преимущества блокчейна в образовании, существуют и некоторые потенциальные риски, которые следует учитывать.

Низкая скорость обработки данных. Обработка больших объемов данных в блокчейне может занимать много времени и быть более медленной, чем в других системах хранения данных. Это может означать, что обработка данных может занимать больше времени, что может привести к необходимости использования дополнительных ресурсов, таких как более мощные компьютеры. Необходимость обновления технологий. Внедрение блокчейна в образование требует значительных затрат на технологическое обновление, включая обновление сетевой инфраструктуры и обучение персонала. Это может быть особенно сложно для бюджетных учреждений, которые могут не иметь достаточных ресурсов для такого обновления.

Проблемы конфиденциальности и безопасности данных. Блокчейн предоставляет высокий уровень конфиденциальности и безопасности данных, но это не означает, что система полностью защищена от кибератак и других нарушений безопасности. Недостаток регулирования и стандартизации. Блокчейн является относительно новой технологией, и еще не существует единого набора стандартов и правил для ее использования в образовании. Это может привести к различным интерпретациям и проблемам взаимодействия между учреждениями.

В мире уже реализованы несколько проектов, в которых блокчейн используется для улучшения образовательных процессов. [2]

Несколько примеров:

– Образовательная платформа OpenEDGAR OpenEDGAR - это образовательная платформа, которая помогает людям изучать программирование и IT-технологии. Платформа использует блокчейн для создания надежной системы проверки знаний студентов. Каждый раз, когда студент успешно завершает урок, результаты сохраняются в блокчейне, что позволяет подтвердить достижения студента и обеспечить более прозрачную систему оценки;

– Платформа для создания транскриптов в Университете Кентукки Университет Кентукки в США использует блокчейн для создания цифровых транскриптов для своих студентов. Транскрипты сохраняются в блокчейне и могут быть доступными только студенту или организации, которую он уполномочит. Это обеспечивает более высокий уровень конфиденциальности и защиты данных;

– Платформа для проверки дипломов в России. Там блокчейн используется для создания платформы, которая помогает работодателям проверять подлинность дипломов выпускников образовательных учреждений. Каждый диплом сохраняется в блокчейне, что обеспечивает более высокий уровень безопасности и надежности;

– Платформа для сбора данных в Китае. Там разработана платформа, которая использует блокчейн для сбора данных об образовании и карьерных достижениях студентов. Эта платформа помогает студентам и работодателям лучше понимать рынок труда и требования к навыкам.

Заключение.

В современном обществе важно использовать новые технологии и инновации в различных сферах, включая образование. Одной из таких технологий является блокчейн, который, в сочетании с большими данными, может значительно улучшить качество образования.

Применение блокчейна в образовании может помочь в решении различных проблем, таких как подделка документов, недостаток прозрачности и отслеживаемости успеваемости студентов, а также проблемы взаимодействия между участниками образовательного процесса. Благодаря блокчейну, студенты могут получить цифровые сертификаты, которые могут быть легко проверены и подлинны. Это может быть особенно важно для работодателей, которые могут быть уверены в том, что кандидат имеет необходимые знания и умения. Блокчейн также может помочь в повышении прозрачности в образовательном процессе, что может быть особенно важно для родителей и других заинтересованных сторон. Благодаря технологии блокчейн, можно отслеживать успеваемость студентов и следить за ходом их обучения.

Кроме того, блокчейн может существенно улучшить взаимодействие между участниками образовательного процесса. Например, студенты могут иметь возможность оставлять отзывы о преподавателях, что может помочь улучшить качество обучения. Также, блокчейн может помочь в управлении финансовыми ресурсами, что может уменьшить затраты на бумажную работу и сократить время на обработку данных. Использование больших данных в образовании также имеет большой потенциал. Благодаря анализу больших данных, можно определить наиболее эффективные методы обучения и оценить результативность образовательных программ. Кроме того, большие данные могут помочь в предсказании успеваемости студентов и их потенциальных проблем.

Таким образом, использование блокчейна и больших данных в образовании имеет огромный потенциал для улучшения качества образования и повышения его эффективности. Несмотря на то, что некоторые проблемы могут возникнуть при внедрении этих технологий, такие как необходимость дополнительных инвестиций и обучения персонала, преимущества, которые они могут принести, значительно перевешивают эти проблемы. Однако, важно понимать, что использование технологий в образовании не может заменить человеческий фактор, такой как опыт преподавателей и межличностные отношения между участниками образовательного процесса.

Список литературы

- [1] Дон Тапскотт, Алекс Тапскотт. Технология блокчейн. То, что движет финансовой революцией сегодня, 2017. - 448 с.
[2] Виктор Майер-Шенбергер, Кеннет Кукьер. Большие данные. Революция, которая изменит то, как мы живем, работаем и мыслим, 2013. – 310 с.

HOW BIG DATA AND AI CAN DELIVER RESULTS

T.V. Kazak

Head of the Department of Engineering Psychology and Ergonomics, doctor of psychological sciences of the Republic of Belarus, doctor of psychological sciences of the Russian Federation, Corresponding Member of the International Academy of Psychological Sciences, Professor, Belarusian State University of Informatics and Radioelectronics, Minsk, Republic of Belarus.

A.A. Voytovich

Student of the Department of Engineering Psychology and Ergonomics.

A.N. Vasilkova

Assistant of the Department of Engineering Psychology and Ergonomics, master.

*Department of Engineering Psychology and Ergonomics
Belarusian State University of Informatics and Radioelectronics, Minsk, Republic of Belarus
E-mail: a.vasilkova@bsuir.by*

Annotation. The work is devoted to the study of the possibilities of using blockchain technology in education, taking into account the analysis of large-scale data (big data). The work examines the benefits of using blockchain to ensure the transparency and security of student data, as well as to optimize the evaluation and analysis of educational results. The paper also considers specific examples of blockchain implementation in education, analyzes their advantages and disadvantages, as well as the prospects for further development of this technology in the educational field. The work is relevant in the light of the need to improve the efficiency of educational processes and the use of modern technologies in education.

Keywords: big data, blockchain, education, analytics, students.

УДК 629.3.06+656.13+612.821

РАЗРАБОТКА ПОДХОДА И ОБОБЩЕННОГО АЛГОРИТМА АДАПТИВНОГО ВЗАИМОДЕЙСТВИЯ СИСТЕМ ВЕРХНЕГО УРОВНЯ ВЫСОКОАВТОМАТИЗИРОВАННЫХ ТРАНСПОРТНЫХ СРЕДСТВ



В.В. Савченко

Начальник НИЦ Бортовых систем управления мобильных машин ОИМ НАН Беларуси, кандидат технических наук, доцент
uus@tut.by



В.А. Дубовский

Ведущий научный сотрудник НИЦ Бортовых систем управления мобильных машин ОИМ НАН Беларуси, кандидат технических наук
vdubovsky.email@gmail.com

В.В. Савченко

Окончил Минский радиотехнический институт. Область научных интересов связана с исследованиями безопасности функционирования систем «человек-машина», разработкой методов и систем верхнего уровня для высокоавтоматизированных и беспилотных транспортных средств.

В.А. Дубовский

Окончил Белорусский политехнический институт. Область научных интересов связана с исследованием проблем взаимодействия человека с техническими системами.

Аннотация. Статья посвящена междисциплинарному исследованию высокоавтоматизированных транспортных средств в контексте решения задачи передачи управления транспортным средством от автоматизированной системы вождения водителю, когда транспортное средство выходит из домена штатной эксплуатации с использованием методологии самоорганизации и синергетики. Рассмотрены структура системы верхнего уровня для передачи управления водителю и обобщенный алгоритм адаптивного взаимодействия систем верхнего уровня.

Ключевые слова: высокоавтоматизированное транспортное средство, домен штатной эксплуатации, коммуникационная платформа, облачные сервисы, передача управления водителю, самоорганизация, синергетика, системы активной безопасности, системы-ассистенты водителя, системы связи.

Введение.

По проблеме самоорганизации за последние десятилетия опубликовано большое количество работ, выполненных учеными разных стран. Чаще всего это работы фундаментального характера, рассматривающие самоорганизацию как предмет или инструмент исследования на примере многоуровневых иерархических систем как естественного происхождения, так и антропогенных.

Сказать, что система «самоорганизующаяся», - считает У. Эшби, - значит придавать этому слову два совершенно разных значения. Первое из них относится к системе, все части которой в начале работы отделены друг от друга (так что поведение каждой из них не зависит от состояния других частей), а затем эти части работают таким образом, что между ними устанавливаются некоторые связи. Такая система изменяется от системы с «разделенными частями» до системы со «связанными частями». Во втором значении «самоорганизация», как и в первом случае, означает «переход от неорганизованной системы к организованной». Степень организации может быть различной. Тогда возможен переход от системы с низкой степенью организации к системе с высокой степенью организации.

Таким образом, различают три типа процессов самоорганизации [1]:

- процессы самозарождения организации, т.е. возникновение из некоторой совокупности целостных объектов определенного уровня новой целостной системы со своими специфическими закономерностями;
- процессы, благодаря которым система поддерживает определенный уровень организации при изменении внешних и внутренних условий ее функционирования;
- процессы, связанные с совершенствованием и саморазвитием таких систем, которые способны накапливать и использовать прошлый опыт.

Новые подходы к определению понятия «самоорганизация» предлагает синергетика [2,3]. Рассмотрим основные свойства биологических самоорганизующихся систем. Одним из таких свойств является ее целенаправленное (целесообразное) поведение. Под целесообразностью понимается общая характеристика поведения сложных динамических систем, направленного на достижение определенного конечного результата. Понятие целесообразности означает, что акт или поведение можно считать направленным к достижению «цели», т.е. конечному условию, при котором система устанавливает определенное временное или пространственное соотношение с другой системой. Тогда термин «целесообразность» употребляется как синоним цели, контролируемой обратной связью.

Цели самоорганизующейся системы могут быть самыми различными. К ним может относиться собственное сохранение системы, ее рост, а в некоторых случаях и ее разрушение. Обычно самоорганизующаяся система характеризуется наличием не одной, а нескольких целей. В некоторых случаях цели вступают в противоречие друг с другом. Тогда выбирается одна главная цель системы, выполнению которой подчиняются остальные ее цели. Подсистемы также могут иметь свои цели, однако, доминирующей является главная цель системы. Выполнение поставленных целей обеспечивается ресурсом системы, необходимым для их достижения. Цели функционирования системы предъявляют требования к критериям, по которым производится сличение и которые определяют целесообразность поведения системы. Существенным признаком самоорганизации является обособленность самоорганизующейся системы от окружающей среды [4]. Выделившись из однородной окружающей среды, система должна сохранять свою целостность, обуславливаемую наличием и разветвленностью связей между частями системы. Важная особенность самоорганизации - исключительное разнообразие обратных связей на всех уровнях системы. Действия самоорганизующейся системы выбираются с помощью обратной связи. Обратная связь приобретает особое значение именно для самоорганизующихся систем потому, что такого типа регулятор автоматически компенсирует любые возмущения, даже те, природа которых неизвестна.

Структурным основанием самоорганизации является множественность элементов и разветвленность связей между ними. Большое число связей на всех уровнях системы ведет к возникновению целостности и развитию направленного на оптимизацию системы глубокого взаимодействия элементов, являющегося функциональным основанием самоорганизации. Множество элементов, входящих в систему, большое количество прямых и обратных связей между элементами, функциональное разнообразие элементов определяют закон необходимого разнообразия, действующий в самоорганизующихся системах [5,6]. Согласно этому закону, любое разнообразие выбора k требует переработки количества информации, равного k . Иными словами, чтобы сделать выбор из множества вариантов мощностью k , необходим полный перебор этих вариантов. Для усечения перебора система должна обладать предварительной информацией. Таким образом, сущность обучения состоит в том, что система предварительно накапливает и обобщает информацию об аналогичных ситуациях в количестве, например, k^* единиц, пригодную для большого числа случаев, чтобы в дальнейшем при выборе значительно сократить перебор, переработав $k - k^*$ единиц информации. Видно, что чем больше накопленный предыдущий опыт k^* , тем меньше количество информации следует переработать при выборе и тем меньше времени необходимо для принятия решения.

Чтобы обеспечить большее число вариантов для выбора, система должна обладать значительным числом состояний, которое обуславливается наличием достаточно большого числа элементов, входящих в систему, т.е. увеличить число k^* . В данном случае речь идет о независимых состояниях или о таких, которые не поддаются простому аналитическому описанию. Большое число элементов необходимо также для того, чтобы обеспечить возможность флюктуаций сигналов. [5].

Исследовав, известные из литературных источников, принципы (свойства) самоорганизующихся систем, представляется целесообразным разбить их на четыре класса:

- определяющие целенаправленное поведение систем;
- обеспечивающие существование отбора элементов и выбора необходимых из них для формирования более сложных блоков;
- обеспечивающие организацию структуры системы;
- позволяющие осуществлять приспособляемость системы к условиям ее функционирования.

«Синергетика – это междисциплинарное направление научных исследований, которое изучает закономерности и принципы, лежащие в основе процессов самоорганизации в системах разной природы: физических, химических, биологических, технических, социальных и других.» [7].

Постановка задачи.

Предмет исследования: передача управления транспортным средством от автоматизированной системы вождения человеку; объект исследования: высокоавтоматизированные транспортные средства (ВТС). Ключевым аспектом функционирования ВТС является безопасность. Рассматриваемая проблема системного синтеза, опирается на синергетические законы единства процессов самоорганизации и управления, что позволяет впервые поставить и решить междисциплинарную прикладную задачу взаимодействия систем верхнего уровня ВТС с единой целевой функцией – повышения безопасности функционирования транспортных средств (ТС).

Основные тренды по направлению.

В соответствии со стандартом SAE J 3016, разработанным Обществом автомобильных инженеров (Society of Automotive Engineers, SAE International) к ВТС относятся ТС 3-го и 4-го уровня автоматизации [8]. 3-й уровень автоматизации ТС предполагает, что в определенных условиях эксплуатации (в домене штатной эксплуатации, в описании должны быть четко перечислены конкретные условия, при которых ВТС предназначено для эксплуатации в автоматизированном режиме) управление ТС может полностью осуществляться системой автоматизированного вождения (САВ) без какого-либо участия водителя, но при этом водитель должен быть готов взять управление ТС на себя в любое время в случае сбоя системы. 4-й уровень автоматизации ТС аналогичен 3-му уровню, отличаясь лишь тем, что водитель не обязан брать управление ТС на себя, даже в случае сбоя системы, хотя он может это сделать при желании.

Вышесказанное означает, что при эксплуатации ВТС в домене штатной эксплуатации водитель может передать управление САВ и заняться делами, не связанными с вождением (просмотр видео, чтение, набор текста, отдых и т.д.), вследствие чего его осведомленность о дорожной ситуации и психофизиологическое состояние могут не быть оптимальными для управления ТС. Это повышает риски безопасности, связанные с человеческим фактором при возникновении необходимости незапланированной передачи ему управления ТС, особенно в сложных дорожных ситуациях [9 – 12].

Проблема передачи управления водителю в ВТС является сложной междисциплинарной проблемой, которая к настоящему времени еще не решена полностью [13]. Для ее решения необходим непрерывный анализ функционального состояния и профессионально важных качеств человека-водителя, характеристик ТС в динамике, состояния САВ и внешней среды,

включая информационные потоки интеллектуальных транспортных систем (ИТС). В настоящее время поиск решений данной проблемы ведется в направлении разработки наиболее эффективных способов предупреждения водителя о необходимости передачи ему управления ТС и соответствующих интерфейсов, которые позволяли бы достаточно быстро и плавно приводить функциональное состояние водителя к нужному уровню [14 – 17].

Решение междисциплинарной проблемы перехода от автоматизированного режима управления ВТС к ручному, когда ВТС выходит из зоны среды штатной эксплуатации, предполагает реализацию ряда взаимосвязанных задач [18 – 20]: автоматический анализ информационных потоков в ИТС и взаимодействие бортовых информационно-аналитических комплексов с внешними системами [21, 22], классификация признаков выполнения водителем ВТС алгоритмов деятельности или их фрагментов и классификация релевантной информации для водителей в ВТС, включая кросс-модальные информационные потоки, мониторинг профессионально важных качеств, когнитивного и функционального состояния водителей, непосредственно во время выполнения алгоритмов деятельности при ручном управлении ВТС, и выявление их динамики, финишная обработка информации осуществляется облачными сервисами [23].

Системы верхнего уровня ВТС (рисунок 1) это: 1. системы активной безопасности – основные, антиблокировочная система тормозов (ABS); антипробуксовочная система (ASR); система курсовой устойчивости (ESP); система распределения тормозных усилий (EBD); система экстренного торможения (BAS); система обнаружения пешеходов (PDS); электронную блокировку дифференциала (EDS) и др.; 2. системы ассистенты водителя (Advanced Driver Assistance System (ADAS) – основные, адаптивный круиз-контроль (ACC); система автономного экстренного торможения автомобиля (AEB); система предупреждения о сходе с полосы (LDW); система мониторинга состояния водителя (DSM); критическое сближение с пешеходом (PCW); опасное сближение (FCW); съезд с полосы движения (LDW); стоящий впереди автомобиль начинает двигаться (FVSA); в зоне обнаружения находится транспортное средство (FPW); автоматическая запись видео до и после аварии (DVR); превышение допустимой скорости (SLR) и др.; 3. системы связи – коммуникационная платформа Cellular Vehicle-to-Everything (C-V2X) позволяет выполнять интеграцию и консолидацию данных из различных источников, для использования различными типами бортовых систем высокоавтоматизированных транспортных средств, включая поддержку автономного вождения в домене штатной эксплуатации, передавать информацию водителю при ручном режиме управления, передавать информацию для актуализации специализированных баз данных для их анализа, в том числе в реальном масштабе времени, которая используется для определения момента времени перехода высокоавтоматизированных транспортных средств от автоматизированного режима управления к ручному; в состав коммуникационной платформы C-V2X входят следующие узконаправленные связи V2V (Vehicle-to-Vehicle), V2I (Vehicle-to-Infrastructure), V2N (Vehicle-to-Network), V2P (Vehicle-to-Pedestrian); 4. спутники глобальных навигационных спутниковых систем (ГНСС), например, ГЛОНАСС, «Галилео», BeiDou, GPS; 5. облачные сервисы и хранилища – для решения, в контексте рассматриваемой проблемы, задач, связанных с обработкой Big Data, сбор данных для анализа; хранение и актуализация данных; выполнение облачных вычислений и анализ в соответствии с заданными алгоритмами; передача требуемых данных на бортовые системы ВТС; визуализация и использование результатов обработки для решения других, например, научно-исследовательских, задач.

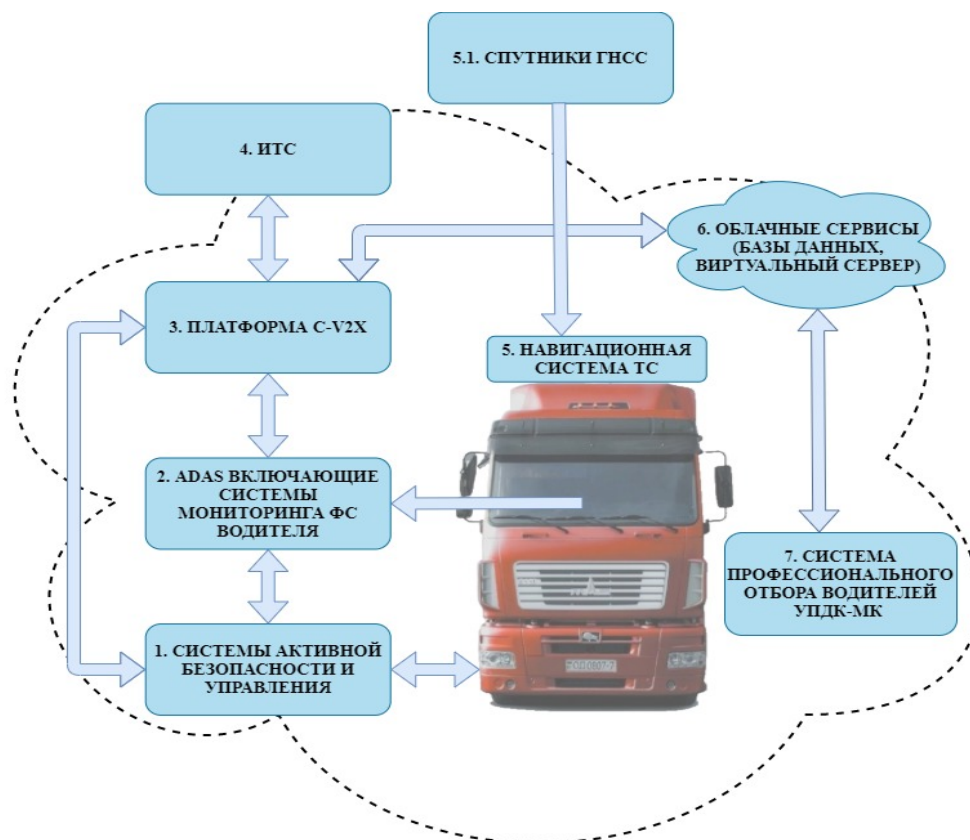


Рисунок 1. Структура системы верхнего уровня для передачи управления водителю в ВТС: 1 – системы активной безопасности и управления ТС; 2 – ADAS, включающие системы мониторинга ФС водителя; 3 – коммуникационная платформа C-V2X; 4 – ИТС; 5 – навигационная система ТС; 5.1 – спутники глобальных навигационных спутниковых систем (ГНСС); 6 – облачные сервисы (базы данных, виртуальный сервер); 7 – стационарная система профессионального отбора водителей (универсальный психодиагностический комплекс УПДК-МК)

Коммуникационная платформа C-V2X является основной информационной системой, обеспечивающей передачу разнородных данных как уже консолидированных в бортовых системах ВТС, так и «сырых» данных для их последующей консолидации облачными сервисами, с целью последующей обработки и анализа для реализации различных функционалов, в том числе появляется возможность использовать сложные сценарии, в бортовых комплексах ВТС, например, поддержка движения как в автономном режиме в домене штатной эксплуатации, так и в режиме ручного управления. Проблема передачи управления водителю, при переходе ВТС к ручному режиму управления, в настоящий момент находится на стадии поисковых исследований. C-V2X позволяет ставить и решать задачи по прототипированию и апробации системы передачи управления водителю на определенном ВТС, с использованием ранее разработанных подходов, концепции, моделей, методов и способов. Рассмотренные архитектуры развертывания C-V2X позволяют обосновывать технические требования к коммуникационной платформе в зависимости от ее функционала, для реализаций современных автомобильных коммуникационных технологий. Обобщенный алгоритм адаптивного взаимодействия систем верхнего уровня ВТС (рисунок 2) ориентирован на получения наиболее достоверного результата в реальном масштабе времени на основе анализа доступной информации о процессах функционирования ВТС и водителя, непрерывного обновления исходных данных, базируется на

концептуальной модели системы «водитель-автомобиль-дорога-среда» [24] и структуре системы верхнего уровня для передачи управления водителю в ВТС.

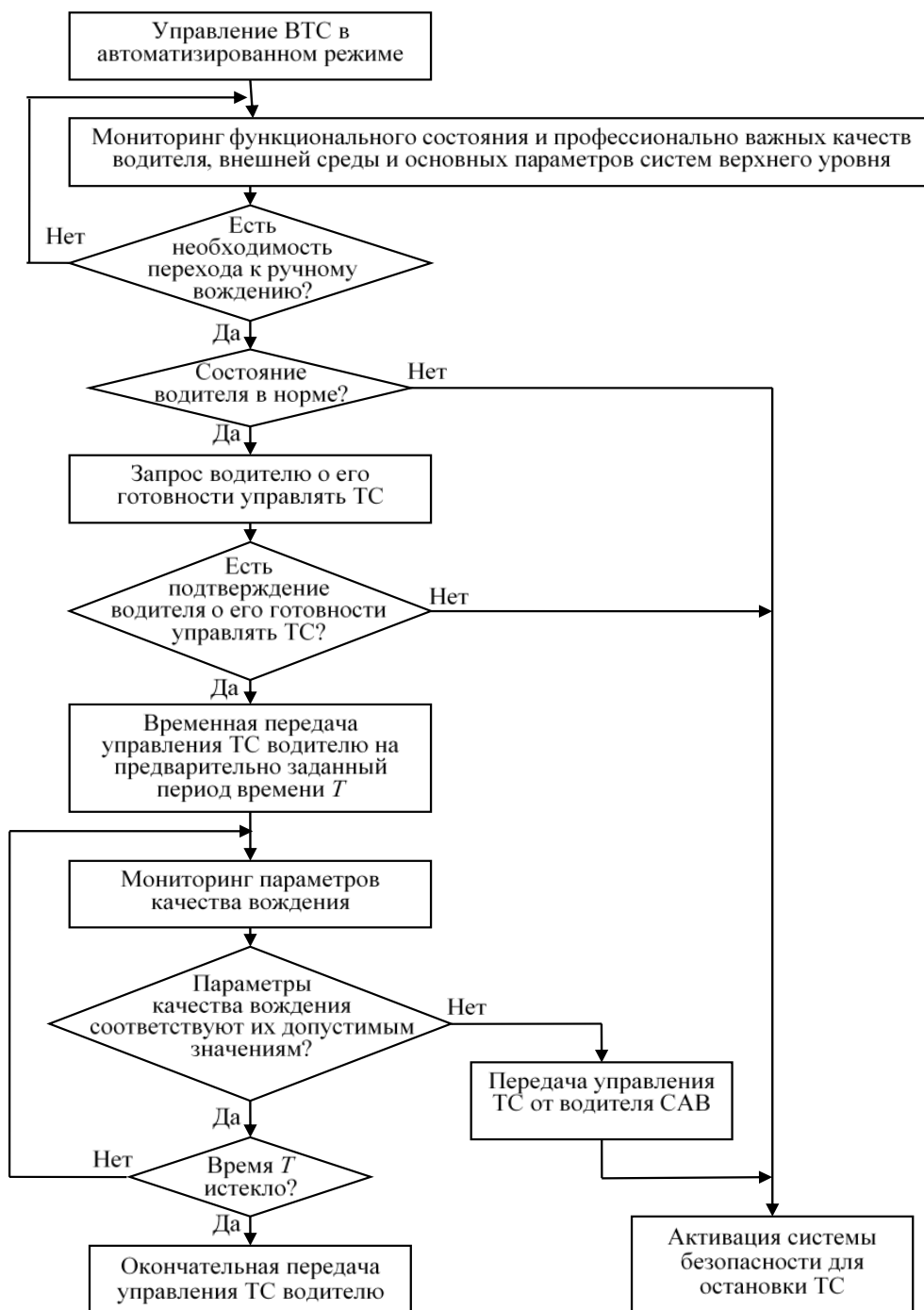


Рисунок 2. Обобщенный алгоритм адаптивного взаимодействия систем верхнего уровня ВТС

Заключение.

С позиций междисциплинарного подхода рассмотрен вариант прикладной реализации подхода к анализу и синтезу систем передачи управления водителю в ВТС, когда транспортное средство выходит из домена штатной эксплуатации. Для реализации задачи используются системы верхнего уровня ВТС и оригинальные методы, способы, алгоритмы мониторинга

бортовых систем управления, водителя, включая основные индивидуальные особенности, в контексте специфики алгоритмов деятельности и внешней среды.

Список литературы

- [1] Самоорганизация // БСЭ. - 3-е изд. - Т. 22. - С. 544.
- [2] Хакен Г. Синергетика. - М.: Мир, 1980. - 404 с.
- [3] Хакен Г. Синергетика: Иерархии неустойчивостей в самоорганизующихся системах и устройствах. - М.: Мир, 1985. - 419 с.
- [4] Биологическая кибернетика // Под ред. А.Б. Когана. - М.: Высшая школа, 1977. - 408 с.
- [5] Винер Н. Кибернетика, или управление и связь в животном и машине. - М.: Сов. радио, 1968. - 326 с.
- [6] Эффективность систем автоматического управления // Энциклопедия кибернетики: В 2 т. - 1975. - Т.2. - С. 511.
- [7] Синергетика [Электронный ресурс]. – Режим доступа: <https://gtmarket.ru/concepts/6876/>. – Дата доступа: 03.03.2023.
- [8] SAE On-Road Automated Vehicles Standards Committee. J3016. Taxonomy and definitions for terms related to on-road motor vehicle automated driving systems. Washington, DC: SAE International; 2014.
- [9] P. Bazilinsky, S.M. Petermeijer, V. Petrovych, D. Dodou, J.C.F. de Winter Take-over requests in highly automated driving: A crowdsourcing survey on auditory, vibrotactile, and visual displays / *Transportation Research Part F* 56 (2018) 82–98.
- [10] Collet C. and Musicant O. Associating Vehicles Automation with Drivers Functional State Assessment Systems: A Challenge for Road Safety in the Future. *Frontiers in Human Neuroscience*. 2019; 13:131. DOI: 10.3389/fnhum.2019.00131.
- [11] Russell H.E.B., Harbott L.K., Nisky I., Pan S., Okamura A.M., Gerdes J.C. Motor learning affects car-to-driver handover in automated vehicles. *Science Robotics*. 2016; 1(1): 1-9.
- [12] Lu Z., Happee R., Cabral C.D.D., Kyriakidis M. and Winter J.C.F. Human Factors of Transitions in Automated Driving: A General Framework and Literature Survey. *Transportation Research Part F: Traffic Psychology and Behaviour*. 2016; 43(11): 183–198.
- [13] Merat N., Jamson A.H., Lai F.C.H., Daly M. and Carsten O.M.J. Transition to manual: Driver behaviour when resuming control from a highly automated vehicle. *Transportation Research Part F*. 2014; 27: 274–282.
- [14] W. Morales-Alvarez, O. Sipele, R. Léberon, H.H. Tadjine and C. Olaverri-Monreal. Automated Driving: A Literature Review of the Take over Request in Conditional Automation *Electronics* 2020, 9, 2087.
- [15] Saito T., Wada T. and Sonoda K. Control Authority Transfer Method for Automated-to-Manual Driving Via a Shared Authority Mode. *IEEE Transactions on Intelligent Vehicles*. 2018; 3(2): 198–207. DOI: 10.1109/TIV.2018.2804167.
- [16] Gold C., Damböck D., Lorenz L. and Bengler K. Take over! How long does it take to get the driver back into the loop? *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*. 2013; 57(1): 1938–1942. DOI: 10.1177/1541931213571433.
- [17] Gold C., Körber M., Lechner D. and Bengler K. Taking Over Control from Highly Automated Vehicles in Complex Traffic Situations: The Role of Traffic Density. *Human Factors*. 2016; 58(4): 642–652. DOI: 10.1177/0018720816634226.
- [18] Савченко, В.В. Проблема передачи управления водителю в высокоавтоматизированных транспортных средствах: метод мониторинга восприятия семантически бинарной релевантной информации водителем / В.В. Савченко // *Механика машин, механизмов и материалов*. – 2019. – №2(47). – С. 14 – 19.
- [19] Савченко, В.В. Концепция передачи управления водителю в высокоавтоматизированных транспортных средствах / В.В. Савченко, С.Н. Поддубко // Докл. Нац. акад. наук Беларуси. – 2020. – Т.64, №5. – С. 624 – 631. <https://doi.org/10.29235/1561-8323-2020-64-5-624-631>
- [20] Дубовский В.А., Савченко В.В. Метод передачи управления транспортным средством от автоматизированной системы вождения человеку. Доклады БГУИР. – 2020. – № 18(7). – С. 40 – 46.
- [21] Савченко В.В. Поддубко С.Н. Подход к разработке метода передачи управления транспортным средством водителю бортовыми системами в автоматическом режиме / В.В. Савченко, С.Н. Поддубко // Труды НГТУ им. П.Е. Алексеева. – 2018. – №2(121). – С. 181–187.
- [22] Савченко В.В. Информационные потоки в высокоавтоматизированных транспортных средствах // «BIG DATE and Advanced Analytics. BIG DATA и анализ высокого уровня», материалы 6-ой Международной научно-практической конференции 20 – 21 мая 2020 г., Часть 2 // г. Минск, Республика Беларусь. – С. 122 – 128.
- [23] Дубовский В.А., Савченко В.В. Метод контроля профессионально важных качеств водителей транспортных средств / В.А. Дубовский, В.В. Савченко // Актуальные вопросы машиноведения: сб. науч. тр. / Объедин. ин-т машиностроения НАН Беларуси; редкол.: С.Н. Поддубко [и др.]. – 2019. – Вып. 8. – С. 159 – 161.
- [24] Дубовский В.А., Савченко В.В. Концептуальная модель системы «водитель-автомобиль-дорога-среда» // Доклады Национальной академии наук Беларуси. – 2019. – Том 63, №1. – С. 112-120.

DEVELOPMENT OF APPROACH AND GENERALIZED ALGORITHM FOR ADAPTIVE INTERACTION OF HIGH-LEVEL SYSTEMS OF HIGHLY AUTOMATED VEHICLES

V.V. Savchenko

*Chief of the Research and Engineering Center
«Onboard Control Systems of Mobile Machines»,
PhD of Technical Sciences, Associate Professor*

V.A. Dubovsky

*Head Lead Researcher of the Research and
Engineering Center «Onboard Control Systems of
Mobile Machines», PhD of Technical Sciences*

State Scientific Institution “The Joint Institute of Mechanical Engineering of the National Academy of Sciences of Belarus”, Minsk, Republic of Belarus.

E-mail: uus@tut.by

Abstract. The article is devoted to the interdisciplinary study of highly automated vehicles in the context of solving the problem of transferring control of the vehicle from an automated driving system to the driver, when the vehicle leaves the normal operation domain using self-organization methodology and synergetics. Structure of upper level system for transfer of control to driver and generalized algorithm of adaptive interaction of upper level systems are considered.

Keywords: highly automated vehicle, standard operation domain, communication platform, cloud services, driver control transfer, self-organization, synergetic, active security systems, driver assistant systems, communication systems.

УДК 004.738.5

ИНТЕРАКТИВНОЕ WEB 3.0 ПРИЛОЖЕНИЕ КАК ИНСТРУМЕНТ КОНКУРЕНТОСПОСОБНОСТИ



В.А. Стельмах
Студент инженерно-
Экономического
факультета БГУИР



О.Н. Шкор
Старший преподаватель
кафедры экономики БГУИР

Белорусский государственный университет информатики и радиоэлектроники, Республика Беларусь.
E-mail: vladislav.stelmakh02@gmail.com, shkor@bsuir.by

В.А. Стельмах

Родился в 2002 году в Минске.. Студент БГУИР

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).

Аннотация. В данной работе рассматривается, как использование интерактивных WEB 3.0 приложений может помочь компаниям, которые предоставляют ИТ-услуги, повысить свою конкурентоспособность на рынке.

WEB 3.0 – это следующее поколение Интернета, которое характеризуется более высокой интерактивностью, интеграцией, доступностью и поддержкой сложных алгоритмов машинного обучения. В контексте ИТ-услуг приложения WEB 3.0 могут помочь компаниям сделать свои сервисы более интерактивными и полезными для пользователей, повысить эффективность работы и наладить более тесные отношения с клиентами. В статье также освещаются примеры приложений WEB 3.0 в различных сферах бизнеса, таких как финансы, образование и здравоохранение. Технологии WEB 3.0 предлагают компаниям более широкий спектр возможностей для решения бизнес-задач и повышения качества обслуживания.

Ключевые слова: WEB 3.0, ИТ-услуги, интерактивные приложения, сервисы, эффективность, конкурентоспособность, блокчейн, искусственный интеллект, безопасность, децентрализация, умные контракты, финансы, здравоохранение, образование, медицинская сфера, инструменты анализа информации.

Введение.

В последние годы большое количество компаний на рынке ИТ-услуг стремятся к использованию более новых технологий и инструментов для повышения своей конкурентоспособности среди конкурентов. Одним из таких инструментов является концепция интерактивного WEB 3.0 приложения. В данной статье мы рассмотрим, как использование такого приложения может помочь компаниям в сфере ИТ-услуг повысить свои шансы на успех и стать более конкурентоспособными.

Прежде чем погрузиться в детали темы для нашей статьи, вспомним, что такое WEB 3.0 – это следующее поколение интернета, которое характеризуется более высокой степенью интерактивности, чем предшествующие версии. WEB 3.0 также отличается большей интеграцией и доступностью различных данных, а также поддерживает более комплексные алгоритмы машинного обучения, что позволяет программному обеспечению становиться наиболее интеллектуальными.

В контексте ИТ-услуг WEB 3.0 приложения могут помочь компаниям сделать свои сервисы более интерактивными и полезными для пользователей. Например, компания может создать интерактивный сервисный портал, который обеспечивает пользователей лучшим пользовательским опытом и удобством в использовании своих сервисов. Он может также позволить пользователям получать очень подробную информацию об услугах компании и проводить операции, которые ранее могли быть доступны только через телефон или личную встречу между двумя сторонами [1].

Важно также отметить, что WEB 3.0 приложения могут помочь в повышении эффективности работоспособности компаний. Например, создание приложения для управления проектами или инструмента для мониторинга производительности сотрудников может ускорить процессы и сократить время на выполнение задач. Кроме того, WEB 3.0 приложения могут помочь компаниям создавать более тесные отношения с клиентами. Пользователи могут получать персонализированный контент и уведомления, а также иметь возможность взаимодействовать со своими персональными данными, что может увеличить их доверие к компании и усилить связь между ними [2].

Разработка WEB 3.0 приложения может требовать больше времени и усилий, но в конечном итоге это может привести к улучшению конкурентоспособности и росту компании. Существуют также готовые решения для создания WEB 3.0 приложений, которые могут значительно сократить время и расходы на разработку. Чтобы успешно внедрить WEB 3.0 приложение, компания должна определить свои конкретные цели и потребности пользователей, а также использовать современные технологии, такие как блокчейн и искусственный интеллект. Это может потребовать сотрудничества с опытными разработчиками, которые могут помочь в осуществлении идеи. В целом, WEB 3.0 приложения могут предоставить компаниям ИТ-услуг отличный инструмент для улучшения взаимодействия с пользователем и повышения эффективности бизнеса. Они предоставляют доступ к большому количеству данных и позволяют анализировать их с помощью специальных алгоритмов, что помогает в принятии рациональных решений и оптимизации работы компании [3].

Одной из главных особенностей WEB 3.0 приложений является возможность создания децентрализованных приложений (DAPPS), которые запускаются на базе концепции блокчейн. Это позволяет уйти от централизованной модели управления и защищает пользовательские данные от возможного взлома. Также WEB 3.0 приложения могут обеспечить определенную степень анонимности и безопасности для пользователей, что важно для повышения доверия к компании и её продуктам [4].

Одним из интересных примеров WEB 3.0 приложений является система управления и продажи отходов компании «Tracr», которая основывается на блокчейне. Она позволяет отслеживать происхождение отходов от производства до утилизации, что способствует более эффективному управлению производственными отходами и более точной оценке стоимости утилизации [5].

В целом, WEB 3.0 приложения представляют собой мощный инструмент для оптимизации бизнеса, который может быть использован в различных сферах деятельности, от образования до финансов. Многие компании уже начали внедрять WEB 3.0 технологии, чтобы повысить эффективность своих бизнес-процессов и обеспечить более высокий уровень безопасности для своих пользователей [6].

Например, в сфере финансов WEB 3.0 приложения могут быть использованы для создания децентрализованных бирж и платформ, которые позволят пользователям проводить сделки напрямую, минуя посредников и сокращая время и стоимость операций.

В сфере образования WEB 3.0 технологии могут быть использованы для создания децентрализованных платформ для обмена знаниями и опытом, а также для создания персонализированных образовательных программ и курсов.

Кроме того, WEB 3.0 приложения могут быть использованы в сфере здравоохранения для создания защищенных платформ для обмена медицинской информацией и поддержки децентрализованных исследований. В целом, WEB 3.0 технологии будут играть все более важную роль в современном мире, помогая компаниям повышать эффективность своих бизнес-процессов и обеспечивать более высокий уровень безопасности для своих пользователей [7].

Таким образом, можно сделать вывод, что концепция WEB 3.0 в ближайшее будущее сможет стать очень важным двигателем прогресса в различных сферах бизнеса, начиная от медицинской сферы и заканчивая финансовой сферой деятельности. WEB 3.0 позволяет создавать более удобные, универсальные и гибкие инструменты для обработки и анализа информации. Анализ данных и машинное обучение на основе WEB 3.0 позволяют компаниям предоставлять более точные и эффективные решения, улучшающие бизнес-процессы и повышающие уровень сервиса.

Принципы WEB 3.0 также соответствуют понятию «умных контрактов», которые могут быть использованы в различных сферах, включая финансы, здравоохранение и даже логистику. Компании могут использовать «умные контракты» для автоматизации процессов и сокращения затрат на сопровождение транзакций и выполнение правовых процедур.

В целом, WEB 3.0 предоставляет компаниям более широкий спектр возможностей для решения бизнес-задач и улучшения качества предоставляемых услуг, что полезно для компаний, предоставляющих ИТ-услуги .

Список литературы

- [1]. What is Web 3.0 (Web3)? Definition, guide and history [Электронный ресурс]. – Режим доступа: <https://www.techtarget.com/whatis/definition/Web-30>.
- [2]. What is Web 3? [Электронный ресурс]. – Режим доступа: <https://www.lifewire.com/what-is-web3-5271876>.
- [3]. Disruptive Blockchain Technology Use Cases 2022 [Электронный ресурс]. – Режим доступа: <https://blockgeeks.com/guides/disruptive-blockchain-technology-use-cases-2022/>.
- [4]. The new wave of web 3.0 dapps on stacks 2.0 [Электронный ресурс]. – Режим доступа: <https://medium.com/@kylellicott/the-new-wave-of-web-3-0-dapps-on-stacks-2-0-615089672ed3>.
- [5]. Tracr [Электронный ресурс]. – Режим доступа: <https://www.tracr.com>.
- [6]. Web 3.0: what is it, and what are its applications? [Электронный ресурс]. – Режим доступа: <https://medium.com/@amounymelissa/web-3-0-what-is-it-and-what-are-its-applications-dca3e889d60f>.
- [7]. Decentralized IDs (DID's) – Empowering Users in the Decentralized World. [Электронный ресурс]. – Режим доступа: <https://medium.com/@hsn80saman/decentralized-ids-dids-empowering-users-in-the-decentralized-world-a5f5547c0519>.

INTERACTIVE WEB 3.0 APPLICATION AS A COMPETITIVE TOOL

V.A. STELMAKH

*Student of engineering and
economics at the BSUIR*

O.N. SHKOR,

*Senior Lecturer at the
Department of Economics at the
BSUIR*

*Belarusian State University of Informatics and Radioelectronics Minsk, Republic of Belarus
E-mail: a.ausievich@gmail.com, shkor@bsuir.by*

Abstract. The article discusses how the use of interactive WEB 3.0 applications can help IT service companies increase their competitiveness in the market.

WEB 3.0 is the next generation of the internet, which is characterized by higher interactivity, integration, accessibility, and support for complex machine learning algorithms. In the context of IT services, WEB 3.0 applications can help companies make their services more interactive and useful for users, improve operational efficiency, and build closer relationships with customers. Developing a WEB 3.0 application may require more time and effort, but it can lead to improved competitiveness and growth. The article also highlights examples of WEB 3.0 applications in various spheres of business such as finance, education, and health care. WEB 3.0 technologies offer companies a wider range of opportunities to solve business problems and improve service quality.

Keywords: WEB 3.0, IT services, interactive applications, services, efficiency, competitiveness, blockchain, artificial intelligence, security, decentralization, smart contracts, finance, healthcare, education, medical field, information analysis tools.

УДК 616.71

АВТОМАТИЗИРОВАННАЯ СИСТЕМА КОНТРОЛЯ «ДИСПЕТЧЕР ЗДОРОВЬЯ»



А.В. Гордиевич

Инженер-программист ОАО «АГАТ-системы управления» - управляющая компания холдинга «Геоинформационные системы управления»
alinagordievich78@gmail.com



А.Д. Скрипка

Студент физического факультета БГУ



П.В. Камлач

Заместитель декана ФКП, доцент кафедры электронной техники и технологии БГУИР, кандидат технических наук, доцент
kamlachpv@bsuir.by

А.В. Гордиевич

Окончила Белорусский государственный университет информатики и радиоэлектроники. Область научных интересов связана с разработкой методов и алгоритмов автоматизированных систем сбора и обработки информации для контроля за состоянием здоровья.

А.Д. Скрипка

Студентка четвертого курса Белорусского государственного университета. Область научных интересов связана с разработкой методов и алгоритмов автоматизированных систем сбора и обработки информации для контроля за состоянием здоровья.

П.В. Камлач

Окончил Белорусский государственный университет информатики и радиоэлектроники. Область научных интересов связана с разработкой и моделированием работы систем медицинского назначения.

Аннотация. Разработана автоматизированная система для сбора, обработки и передачи данных о состоянии здоровья. Основное назначение системы – оценка состояния сердечно-сосудистой системы пациента. В системе в непосредственной близости к пациенту и удаленно осуществляются измерения и оценка показателей здоровья, представление информации от датчиков и результатов визуально на экранах, передача данных современными техническими средствами для удаленного использования. Получены рекомендации по выбору элементной базы для создания автоматизированной системы медицинской диагностики «Диспетчер здоровья», которая может быть использована в практике медицинского учреждения.

Ключевые слова: Автоматизированная система для контроля состояния здоровья, частота сердечных сокращений, передача данных, экстренная помощь, элементная база, датчик, микроконтроллер.

Введение.

В современном мире существует множество устройств для оценки состояния здоровья человека. С каждым годом потребность в таких устройствах увеличивается. Приоритетной является задача проведения всесторонней и своевременной диагностики пациента с использованием мобильных автоматизированных систем, выполненных из совокупности микропроцессорных устройств и специализированных программных средств.

В современном мире мы не можем представить медицинскую диагностику без использования современных технологий, в первую очередь – автоматизированных систем. Одной из таких систем является «Диспетчер здоровья» – комплекс, созданный для более точного и быстрого определения диагноза пациента, выполненный из совокупности микропроцессорных устройств и специализированных программных средств.

Актуальность.

Современный рынок постоянно предоставляет устройства для расширения возможной медицинской диагностики (беспроводной стетоскоп, мобильный тонометр, фитнес-трекеры и т.п.). Все они обладают рядом полезных свойств и усовершенствований, но имеют и разного рода недостатки. К положительным качествам указанных и подобных им средств можно отнести их большое разнообразие и доступность на рынке, портативное исполнение, оперативное предоставление полезных показателей здоровья. Из минусов, имеющих место в большинстве случаев, надо указать на высокую стоимость качественных приборов, невысокую точность и низкую достоверность измерений, выполняемых приборами широкого потребления, отсутствие полноты и требуемой информативности результатов, которые необходимы врачу для постановки точного диагноза и выработки экстренного решения для оказания помощи, отсутствие или ограниченность технических возможностей для подключения таких интеллектуальных средств обработки и распределения информации как смартфоны и другие гаджеты, удаленные приборы диагностики, наличие высокой сложности и стоимости обеспечения связи для оперативного диалога между пациентом и медицинскими работниками [1].

Решение ряда из указанных проблем возможно на основе создания мобильной автоматизированной системы для контроля за сердечно-сосудистой деятельностью пациента с целью выработки компетентного решения об оказании необходимой помощи.

Описание микропроцессорной системы для сбора, обработки и передачи данных о состоянии здоровья.

Микропроцессорная система «Диспетчер здоровья» предназначена для сбора, обработки и передачи данных о состоянии здоровья. Она включает в себя:

- портативное устройство для сбора и передачи данных, которое включает набор датчиков для измерения параметров состояния органов и систем пациента и средства связи для доставки полученных данных близко расположенным и удаленным пользователям – специалистам;
- комплект средств обработки и визуализации для специалистов, располагаемый в непосредственной близости от пациента;
- комплект средств обработки и визуализации для специалистов, располагаемый удаленно от пациента;
- удаленные диагностические приборы;
- средства связи с глобальной сетью *Internet*.

Микропроцессорная система «Диспетчер здоровья» приведена на рисунке 1.

В системе предусматривается возможность подключения широкого перечня датчиков, а также использование различных средств обработки данных и доступных пользователю коммуникационных средств. В состав средств связи предполагается включать средства для непосредственной и удаленной связи между отдельными частями системы (датчиками, установленными на теле пациента, средствами обработки, отображения), средства для обмена данными об измеренных датчиками параметрах пациента с находящимися в шаговой оперативной доступности и удаленными приборами их обработки, средства доведения результатов удаленной обработки к устройствам контроля состояния здоровья и оказания помощи.

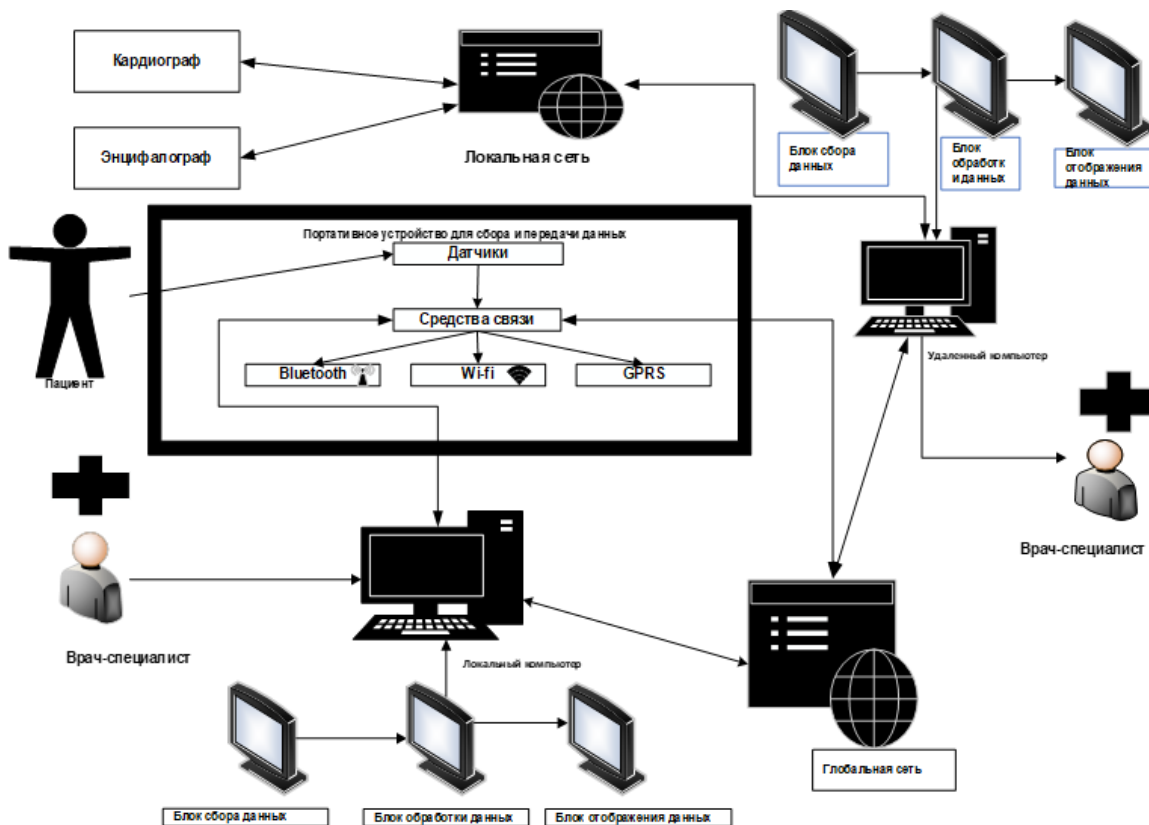


Рисунок 1. Микропроцессорная система для сбора, обработки и передачи данных о состоянии здоровья

На рисунке 2 приведена более подробная схема портативного устройства для сбора и передачи данных.

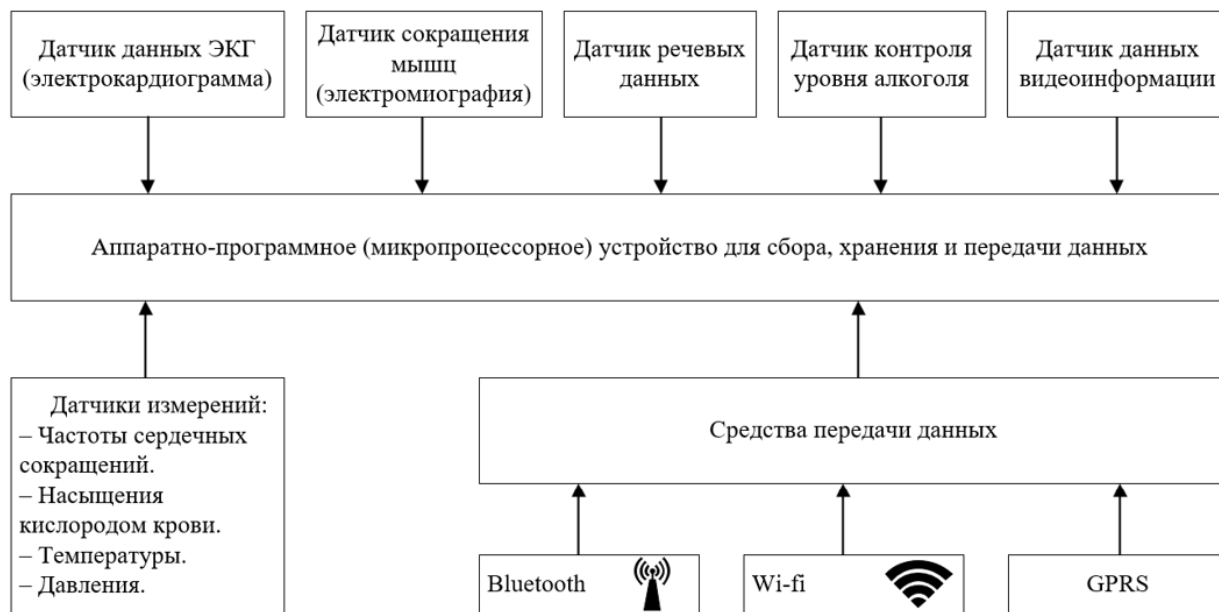


Рисунок 2. Схема портативного устройства для сбора и передачи данных

Для использования системы с целью диагностики сердечно-сосудистой деятельности пациента в ее составе предполагается подключение к пациенту следующие типов датчиков:

- датчик данных для съема электроэнцефалограммы (ЭЭГ);
- датчик видеoinформации о внешнем виде пациента;
- датчик речевых данных;
- датчик данных электрокардиограммы (ЭКГ);
- датчик контроля уровня алкоголя;
- датчик контроля за сокращениями мышц (датчик электромиографии);
- датчик уровня насыщения крови кислородом,
- датчик температуры тела;
- датчик кровяного давления (тонометрия);
- датчик частоты сердечных сокращений (пульсометр).

В состав средств передачи данных предлагается включить:

- модуль *Bluetooth*;
- модуль *Wi-Fi*;
- модуль *GPRS (General Packet Radio Service)*.

Для организации связи между компьютерами может использоваться глобальная сеть Internet или средства локальной вычислительной сети типа Ethernet.

Для подключения средств связи к датчикам в составе портативного устройства предлагается использовать аппаратно-программное (микропроцессорное) устройство для сбора, хранения и передачи данных.

На рисунке 3 схематически показано подключение датчиков к пациенту.

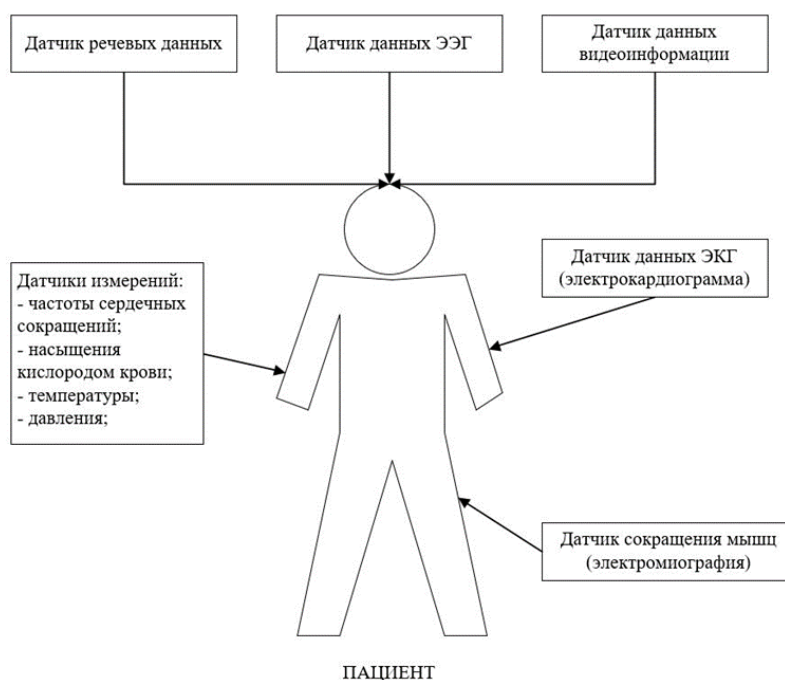


Рисунок 3. Схема подключения датчиков к пациенту

Удаленными датчиками для пациента могут быть такие приборы как энцефалограф и кардиограф, которые показаны на рисунке 1.

Характеристики модулей, предлагаемых для передачи данных.

Технология, используемая для построения модулей *Bluetooth*, является беспроводной. Применяется для дистанционной передачи данных на коротких расстояниях между различными

- обладает 14 портами ввода и вывода, *SPI, I2C, UART*, 10-бит АЦП;
- поддерживает внешнюю память до 16 МБ;
- необходимое питание от 2,2 до 3,6 В, потребляемый ток до 300 мА, в зависимости от выбранного режима.

Важной особенностью микроконтроллера является отсутствие пользовательской энергонезависимой памяти на кристалле.

Bluetooth-модуль *HC-05* – данный модуль является одним из лучших решений для организации двусторонней связи по *Bluetooth* [3]. Он может работать как в режиме ведущего «*master*» (осуществлять поиск *Bluetooth*-устройств и инициировать установку связи), так и ведомого устройства «*slave*» (ведомое устройство). Внешний вид модуля *Bluetooth*-модуль *HC-05* приведен на рисунке 5.

Технические характеристики модуля *HC-05*

- чип *Bluetooth: HC-05(BC417143)*;
- диапазон частот радиосвязи: 2,4–2,48 ГГц;
- мощность передачи: 0,25–2,5 мВт;
- чувствительность: –80 дБм;
- напряжение питания: 3,3–5 В;
- потребляемый ток: 50 мА;
- радиус действия: до 10 метров;
- интерфейс: последовательный порт;
- режимы: *master, slave*;
- температура хранения: –40...85 °С;
- рабочий диапазон температур: –25...75 °С;
- габариты: 27 x 13 x 2,2 мм.

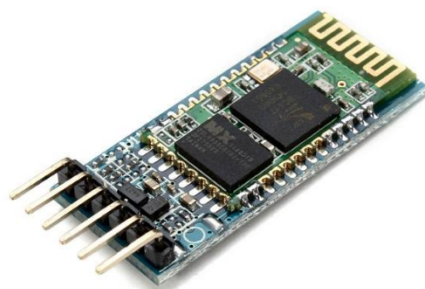


Рисунок 5. Внешний вид *Bluetooth*-модуля *HC-05*

Для обмена Смс-сообщениями и данными по GPRS может быть использован модуль *A6* фирмы *AI-THINKER* [4]. Устройство разработано в 2016 году. Устройство отличается низким потреблением энергии и малыми размерами. Внешний вид устройства приведен на рисунке 6.

Технические характеристики:

- диапазон напряжений 4,5 – 5,5 В;
- питание 5 В;
- диапазон рабочих температур от –30 °С до 80 °С;
- максимальное потребление тока 900 мА;
- *GPRS Class 10*;
- поддержка протоколов *PPP, TCP, UDP, MUX*.



Рисунок 6. Модуль GPRS

В качестве средств отображения могут быть рекомендованы к использованию стандартные мониторы персональных компьютеров, которыми должны оснащаться рабочие места меди-ков-специалистов.

Рассмотрим подробнее варианты для исполнения предложенных ранее датчиков.

Пульсоксиметр *МН-ЕТ МАХ30102* [5]. Внешний вид датчика представлен на рисунке 7.



Рисунок 7. Внешний вид пульсоксиметра *МН-ЕТ МАХ30102*

В корпусе *МАХ30102* реализована полнофункциональная схема сенсорного модуля для создания портативных систем пульсоксиметрии с высокими требованиями к точности измерений. Устройство имеет миниатюрные размеры, добиться которых удалось без ущерба для оптических или электрических характеристик. Для интеграции в полнофункциональную носимую измерительную систему потребуется минимум дополнительных внешних компонентов.

Процентное содержание кислорода в крови в данном случае определяется неинвазивным методом через кожу (о чем свидетельствует обозначение «*Sp*»), как процентное отношение насыщенного кислородом гемоглобина (HbO_2) к общему содержанию гемоглобина ($HbO_2 + Hb$), определяемых с помощью фотодетектора, ИК и красного светодиода *МАХ30102*.

Подсистема измерения SpO_2 включает схему компенсации внешней засветки (КВЗ), сигма-дельта-АЦП и патентованный цифровой фильтр. КВЗ имеет внутреннюю схему блокировки сигнала для устранения внешней засветки и расширения эффективного динамического диапазона. АЦП программируется во всем диапазоне измерений 2-16 мкА. КВЗ позволяет блокировать сигнал внешней засветки величиной до 200 мкА.

В *MAX30102* имеется встроенный датчик температуры для калибровки температурной зависимости подсистемы измерения SpO_2 . Датчик температуры имеет разрешение $0,0625^{\circ}C$.

Выходные данные *MAX30102* сравнительно нечувствительны к длине волны ИК-светодиода, тогда как длина волны красного светодиода имеет решающее значение для правильной интерпретации результатов измерений. Используемый *MAX30102* алгоритм для измерения SpO_2 позволяет компенсировать ошибки, возникающие с изменением температуры окружающей среды.

Модуль измерения ЭКГ на базе *AD8232*. Внешний вид датчика представлен на рисунке 8.

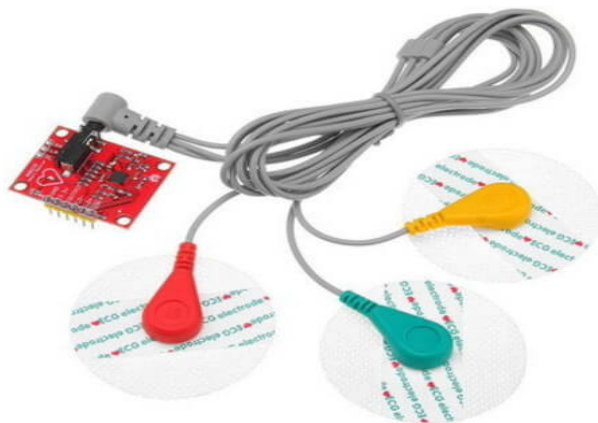


Рисунок 8. Внешний вид модуля измерения ЭКГ на базе *AD8232*

Датчик ЭКГ *AD8232* разработан компанией *Analog Devices*, он на 50% компактнее и использует на 20 процентов меньше энергии, чем аналогичные устройства других производителей. Он состоит из платы *AD8232*, набора электродов и кабеля для их подключения к плате *AD8232* [6].

Преимуществом данного датчика является то, что он предназначен для получения, усиления и фильтрации слабых биопотенциальных сигналов в условиях сильных помех. *AD8232* включает в себя двухполюсный фильтр высоких частот и операционный усилитель, который позволяет использовать технологию многополюсной низкочастотной фильтрации для удаления шума линии и других помех. Благодаря этому датчик можно напрямую подключать к осциллографу через контакт *OUTPUT*. Данный датчик снимает показания ЭКГ по методу двухполюсных отведений, при котором фиксируется разность потенциалов между двумя точками электрического поля. Эти данные можно использовать для мониторинга работы кардиосистемы пациента.

Датчик *MyoWare* измерения параметров сокращения мышц. Внешний вид датчика представлен на рисунке 9.



Рисунок 9. Внешний вид датчика *MyoWare*

Плата *MyoWare*, входящая в состав датчика, действует путем измерения выпрямленной и фильтрованной электрической активности мышцы. Выходное напряжение изменяется от 0 до напряжения питания в зависимости от активности в выбранной мышце [7].

Датчик уровня алкоголя *MQ-3*. Внешний вид датчика представлен на рисунке 10.

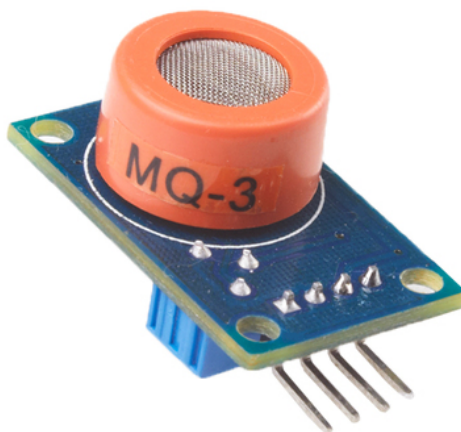


Рисунок 10. Внешний вид датчика *MQ-3*

Датчик *MQ-3* обеспечивает обнаружение паров спирта в выдыхаемом пациентом воздухе. близости от дыхания. Датчик аналоговый, поэтому его точность напрямую зависит от калибровки [8].

Принцип работы модуля основан на нагревании чувствительного элемента до определенной температуры: датчик оставляют на 24 часа, не прекращая подачу питания, внутри нагревательного элемента происходит химическая реакция, после чего вещество в датчике становится чувствительным к парам спирта, так как вступает с ними в реакцию. На выходе модуля появляется напряжение, которое меняется в зависимости от концентрации вещества.

Миниатюрная *IP* видеочамера с микрофоном *Ambertek* для видеонаблюдения [9]. Максимально эффективными для проведения видеонаблюдения за пациентом являются устройства, которые являются незаметными и при этом способны фиксировать не только видео, но и звук. Этим требованиям лучше всего соответствуют миниатюрные *IP* камеры с микрофоном, представленные в линейке *Wi-Fi* видеочамер компании *Ambertek*.

Микрофон даёт возможность получения видео со звуком. Благодаря этому создается полная картина событий, происходящих в месте установки. Кроме того, аудиозапись в некоторых случаях позволяет фиксировать то, что остается вне зоны видимости видеочамеры. Некоторые модели с микрофоном оснащаются датчиком звука. Это позволяет им автоматически начинать съемку при распознавании требуемого сигнала. Миниатюрные *IP* камеры оснащаются встроенным микрофоном.

Работы по созданию автоматизированной системы контроля «Диспетчер здоровья».

В настоящее время в составе работ по созданию микропроцессорной системы обработки, передачи и хранения информации о состоянии здоровья пациента выполнено следующее:

1. Разработана микропроцессорная система для контроля частоты сердечных сокращений во времени на базе платы *Arduino Uno* [10, 11]. Подключение пациента к плате выполнено с помощью датчика *Pulse Sensor*.

2. Выполнено подключение платы *Arduino Uno* к персональному компьютеру с помощью *USB* кабеля, обеспечивающего передачу данных на расстояние до 30 см (рисунок 11).

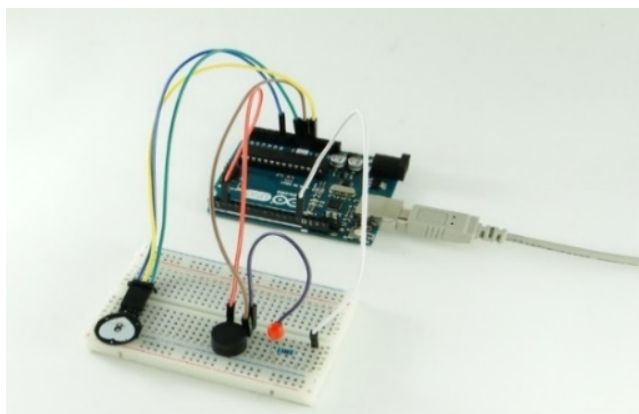


Рисунок 11. Микропроцессорная система для снятия данных о частоте сердечных сокращений

3. Для персонального компьютера разработано программное обеспечение, которое осуществляет прием пульсограммы пациента от платы *Arduino Uno*, ее обработку и визуализацию на экране монитора. Вид пульсограммы при отображении на мониторе приведен на рисунке 12. Программное обеспечение создано на базе языка *Java* в среде разработки *Processing*.

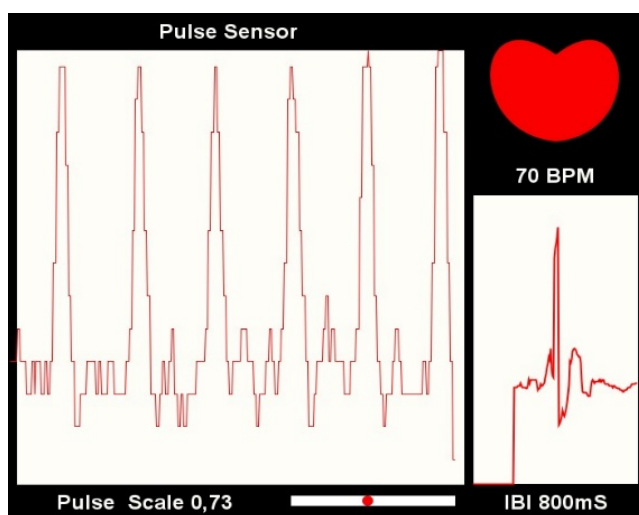


Рисунок 12. Визуализация пульсограммы

4. В компьютере произведена обработка пульсограммы (рисунок 13) с помощью комбинации фильтров Савицкого-Голея и Чебышева [12, 13], обеспечивающих устранение шумов, возникающих при съеме пульсограммы. Для улучшения информативности предоставляемых специалистам результатов на представлении пульсограммы на экране выполнено сглаживание пульсограммы (рисунок 14) с помощью наложения сплайна третьего порядка.

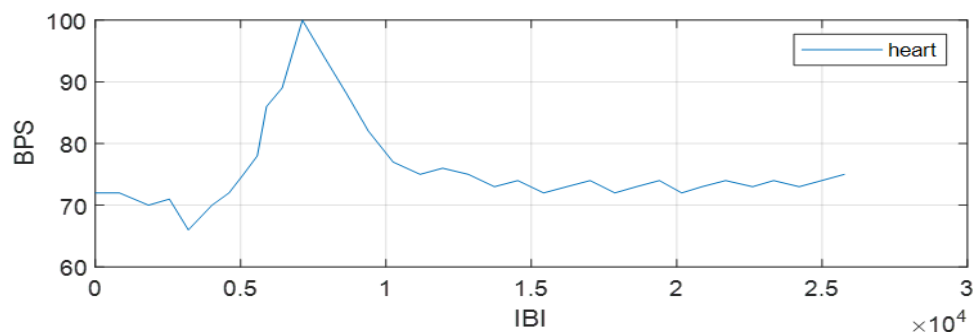


Рисунок 13. График пульсограммы до фильтрации

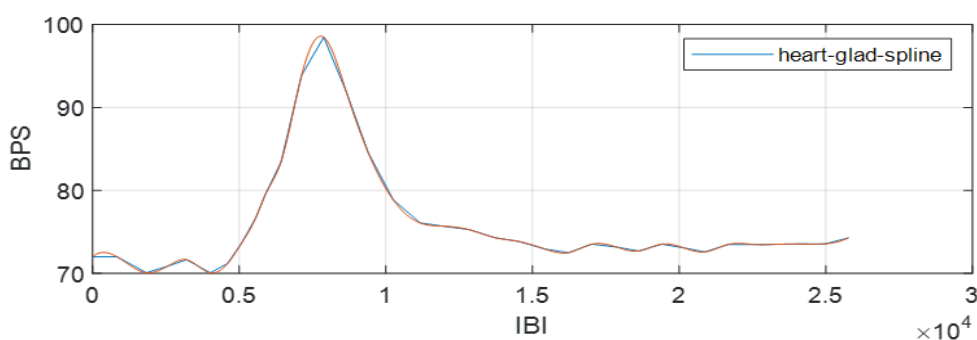


Рисунок 14. Графики пульсограммы после фильтрации и применения сплайна

5. Разработан вариант создания сайта для обеспечения взаимодействия пациента со специалистом на базе *Java SpringBoot*.

Заключение.

В данной работе рассмотрены предложения по разработке микропроцессорной системы сбора, хранения и передачи данных о состоянии здоровья пациента. Рассмотрен также объем и содержание конкретных работ, выполненных по практической реализации рассмотренной системы диагностики с использованием средств передачи данных по *Wi-Fi*. Выбор и оценка элементной базы для автоматической системы медицинской диагностики является важным этапом в разработке и создании системы. Она должна быть надежной, точной и адаптируемой к различным клиническим и эпидемиологическим ситуациям. При правильном выборе и оценке элементной базы можно достичь более точной диагностики и улучшить качество медицинской помощи. В перспективе предполагается дальнейшее практическое развитие системы на описанных принципах. Это позволит при внедрении в медицинскую практику улучшить и ускорить диагностику состояния здоровья пациентов на основе предоставления качественных медицинских данных. Сократится нагрузка на поликлиники, больницы и бригады скорой помощи. Появится возможность удаленного обследования пациентов, тем самым сократиться количество летальных исходов.

Список литературы

- [1] Минченя, В.Т. Автоматизированные системы контроля медикобиологических параметров / В.Т. Минченя, А.Л. Савченко, Р.М. Аси-мов. – Минск: Тесей, 2011. – 168 с.
- [2] Flickenger R. The Wi-Fi Experience: Everyone's Guide to 802.11b Wireless Networking / R. Flickenger. – Que Pub, 2018. – 169 p.
- [3] Bluetooth-модуль HC-05 [Электронный ресурс] – Режим доступа : <https://3d-diy.ru/wiki/arduino-moduli/bluetooth-modul-hc-05/>. – Дата доступа: 10.03.2023.

- [4] GPRS/GSM модуль А6 на базе SIM900А [Электронный ресурс] – Режим доступа : https://amperkot.by/products/gprs_gsm_modul_a6_na_baze_sim900a/24260128.html. – Дата доступа: 10.03.2023.
- [5] Пульсоксиметр МН-ЕТ МАХ30102 [Электронный ресурс] – Режим доступа : <https://compacttool.ru/pulsoksimetr-mh-et-max30102>. – Дата доступа: 10.03.2023.
- [6] Модуль для ЭКГ на базе AD8232 [Электронный ресурс] – Режим доступа : <https://radio-market.by/arduino/41883-modul-dlya-ekg-na-baze-ad8232.html>. – Дата доступа: 10.03.2023.
- [7] Датчики биометрических параметров для Arduino [Электронный ресурс] – Режим доступа : <http://digitrode.ru/articles/1523-biometricheskie-datchiki-sovmestimye-s-arduino.html>. – Дата доступа: 10.03.2023.
- [8] Модуль датчика паров алкоголя MQ-3 [Электронный ресурс] – Режим доступа : https://tixer.ru/catalog/modules/sensor-modules/modul_datchika_parov_alkogolya_mq_3/. – Дата доступа: 10.03.2023.
- [9] Миниатюрные IP камеры Ambertek с микрофоном для видеонаблюдения [Электронный ресурс] – Режим доступа : <https://sagent.ru/informaciya-o-tovarah/skrytyh-videokamer/448-luchshie-mini-wi-fi-videokamery-nablyudeniya-s-mikrofonom>. – Дата доступа: 10.03.2023.
- [10] Петин В.А. Проекты с использованием контроллера Arduino / В.А. Петин. – СПб.: БХВ-Петербург, 2014. – 400 с.
- [11] Соммер У. Программирование микроконтроллерных плат Arduino/Freduino / У. Соммер. – СПб.: БХВ-Петербург, 2012. – 256 с.
- [12] Васковская Л.Ф. Шестиканальный электрокардиограф с фильтрацией электрокардиограмм / Л.Ф. Васковская, А.В. Гордиевич, П.В. Камлач, И.И. Ревинская, // ИЗОБРЕТАТЕЛЬ Международный научно-практический журнал. – 2021. – № 3. – С. 8–14.
- [13] Васковская Л.Ф. Многоканальный электрокардиограф с фильтрацией электрокардиограмм методами медиан, Савицкого-Голея и Чебышева / Л.Ф. Васковская, А.В. Гордиевич, П.В. Камлач // ИНФОРМАЦИОННО- ИЗМЕРИТЕЛЬНЫЕ И УПРАВЛЯЮЩИЕ СИСТЕМЫ Международный научно-технический журнал. – 2021. – Том 19, № 6. – С. 45–54.

AUTOMATED CONTROL SYSTEM «HEALTH SUPERVISOR (DISPATCHER)»

A.V. Hardziyevich

*Engineer-programmer open joint-stock company
«AGAT – Control Systems – Managing Company of Geoinformation Control Systems Holding»*

N.D. Skrypka

*Student
Belarusian State University,
Faculty of Physics*

P.V. Kamlach

Deputy Dean of the Faculty of Computer-Aided Design, Associate professor of the Academic department of electronic technique and technology, PhD, associate professor

Department of Electronic technique and technology

Faculty of Computer Engineering

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

E-mail: alinagordievich78@gmail.com

Abstract. Automated health state data collection, processing and transfer system was developed. The main purpose of the system is to rate the state of the patient cardiovascular system. System provides measurements and evaluations of the health indicators nearby and remote from the patient, shows the data from the sensors on the screens and transfers the data with up-to-date technical means for remote using. Recommendations for choice of elemental base to create the automated medical diagnostic system “Health Supervisor (Dispatcher)”, which can be used in medical establishment practice, were realized.

Keywords: Automated health state control system, heart rate, data transfer, emergency help, elemental base, sensor, microcontroller.

УДК 681.325

ПОЛИНОМИАЛЬНАЯ РЕАЛИЗАЦИЯ СИСТЕМ БУЛЕВЫХ ФУНКЦИЙ В ЗАДАЧАХ ЛОГИЧЕСКОГО ПРОЕКТИРОВАНИЯ



Н.А. Кириенко

Старший научный сотрудник Объединенного
института проблем информатики НАНБ,
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. Функциональные описания проектируемых цифровых устройств могут иметь сотни входных переменных и десятки тысяч уравнений. В связи с этим возникает задача эффективного представления исходных описаний цифровых устройств, которое позволит улучшить качество и сократить время синтеза цифрового устройства. Предлагается использовать полиномиальное представление систем булевых функций, которыми может описываться поведение дискретных устройств.

Рассматриваются представления системы булевых функций в виде полиномов Жегалкина и Рида-Маллера.

Описываются алгоритмы построения полиномов, исходя из различных форм задания поведения логических устройств: систем полностью определенных булевых функций, систем частично определенных булевых функций, таблиц истинности.

Ключевые слова: логическая схема, функциональное описание цифрового устройства, системы булевых функций, полином Жегалкина, полином Рида-Маллера.

Введение.

Исходное описание поведения проектируемых логических устройств удобно представлять формулами, представляющими собой композиции некоторых простых булевых функций, образующих базис синтеза и реализуемых отдельными схемными элементами. Практически важная задача минимизации числа элементов в схеме сводится к нахождению оптимальной композиции.

Наиболее распространенным базисом является трехэлементный базис «конъюнкция, дизъюнкция, отрицание – И, ИЛИ, НЕ». Но с развитием микроэлектронной технологии появились элементы, реализующие многоместные дизъюнкции с исключением.

Структура таких схем описывается формулами, подобными ДНФ, в которых вместо операторов дизъюнкции используются операторы дизъюнкции с исключением, так называемые «И-исключающее ИЛИ (AND/EXOR)» формулы. Иногда такие формулы называют ESOP – exclusive sum of products [1].

Известно, что число термов в таких описаниях значительно меньше, чем в базисе И-ИЛИ-НЕ.

Кроме того, схемы AND/EXOR легче диагностируются. Поэтому развитие методов синтеза легко диагностируемых схем в базисе полиномиальных форм является актуальной задачей.

В настоящей работе представлено решение задачи построения полиномиальных форм для различных исходных заданий булевых функций.

Постановка задачи.

Рассмотрим два вида полиномиального представления булевых функций: полином Жегалкина и полином Рида-Маллера. Полином Жегалкина определяется как многоместная сумма по модулю два попарно различных элементарных конъюнкций (содержащих только положительные литералы) [2].

В виде полинома Жегалкина можно представить любую булеву функцию, эта форма является канонической.

Дизъюнктивная нормальная форма (ДНФ) представляет многоместную дизъюнкцию элементарных конъюнкций (термов). ДНФ называется совершенной, если каждый терм содержит символы всех входных переменных описания.

Полином Жегалкина можно получить из дизъюнктивной нормальной формы путем последовательных эквивалентных преобразований [2]:

$$\begin{aligned} ab\vee a'c\vee bd' &= ab\vee a'c\vee a'bc'd' = ab \oplus a'c \oplus a'bc'd' = \\ &= ab \oplus (a \oplus 1)c \oplus (a \oplus 1)b(c \oplus 1)(d \oplus 1) = \\ &= ab \oplus ac \oplus c \oplus b \oplus ab \oplus bc \oplus bd \oplus abc \oplus abd \oplus bcd \oplus abcd = \\ &= ac \oplus c \oplus b \oplus bc \oplus bd \oplus abc \oplus abd \oplus bcd \oplus abcd. \end{aligned}$$

Символ «'» означает инверсию переменной, после которой он стоит.

Полином Рида-Маллера отличается от полинома Жегалкина тем, что некоторые из литералов могут быть со знаком инверсии (и тогда только с ним). Разработаны компактные векторно-матричные процедуры перехода от СДНФ к полиномам Жегалкина и Рида-Маллера, и обратно [2].

В работе рассматриваются алгоритмы и программы построения полиномиальных форм для различных исходных заданий системы булевых функций, которой описывается поведение логической схемы.

В качестве исходной формы булевых функций могут быть ДНФ, или таблицы истинности.

Алгоритмы построения полиномиальных форм.

Алгоритмы построения полиномов Жегалкина и Рида-Маллера базируются на четырех векторно-матричных процедурах, представленных в [2], реализованных на языке C++, и перечисленных в таблице 1.

Функции 1 и 2, которые представлены в таблице 1, требуют в качестве исходного задания две булевых матрицы, соответствующих системе частичных булевых функций в СДНФ. Функции 3 и 4 требуют в качестве исходного задания булев вектор, который соответствует значениям булевой функции на всех 2^n наборах булева пространства (таблица истинности).

Рассмотрим формы представления систем булевых функций, описывающих поведение логических схем. Для целей выполнения дальнейшей оптимизации и синтеза цифровых устройств актуальными являются описания в форме систем ДНФ, частичных булевых функций, систем булевых функций, определенных на таблице истинности.

Таблица 1. Векторно-матричные процедуры построения полиномов Жегалкина и Рида-Маллера

	Описание процедуры	Выполняемые действия	Вид исходного задания
1	CSOP MinSiJ_D(CBM &BmX, CBM BmF)	Минимизация в классе полиномов Жегалкина системы частичных булевых функций, заданных в совершенной нормальной форме.	Булева $p \times n$ матрица BmX задает p наборов n переменных, а булева $p \times m$ матрица BmF – значения m функций на этих наборах.
2	CSOP MinSiRM_D(CBM &BmX, CBM BmF)	Минимизация в классе общего вида полиномов Рида-Маллера системы частичных булевых функций, заданных в совершенной нормальной форме.	Булева $p \times n$ матрица BmX задает p наборов n переменных, а булева $p \times m$ матрица BmF – значения m функций на этих наборах.
3	void MinSiJ(CBM &BmF)	Построение системы полиномов Жегалкина, реализующей систему полностью определенных булевых функций, заданных в векторной форме.	Булева $m \times 2^n$ матрица BmF задает вектор значений функций на наборах таблицы истинности.
4	CSOP MinFPRM(int Regime, CBM BmF, int &v, int &w, double &T1, double &T2, double &T3)	Поиск оптимальной полярности для FPRM-представления системы полностью определенных булевых функций, заданных в векторной форме	Булева $m \times 2^n$ матрица BmF задает вектор значений функций на наборах таблицы истинности. При Regime=0 ищется система полиномов Жегалкина, иначе – система полиномов Рида-Маллера. v - число термов в системе полиномов Жегалкина, w - число термов в системе полиномов Рида-Маллера. T1 - время (в с) построения системы полиномов Жегалкина, T2 - время (в с) построения системы полиномов Рида-Маллера, T3 - время (в с) оформления разностной нормальной формы РНФ.

В таблице 2 представлены исходное структурно-функциональное описание схемы GenP в виде системы ДНФ и результирующее описание схемы в виде полинома Жегалкина (матричная и символьная формы). Описание представлено на внутреннем языке SF, который используется для описания объектов проектирования в системе логического проектирования FLC-2 [3].

Результирующее описание в матричной форме следует понимать, как «сумму по модулю два» конъюнкций, отмеченных единицей в матрице функций.

Таблица 2. Описание схемы GenP в виде системы ДНФ и полинома Жегалкина

Исходное задание	Полином Жегалкина в матричной форме	Полином Жегалкина в символьной форме
TITLE GenP FORMAT SF AUTHOR Author DATE 03.11.2020 10:28:44 PROJECT Test DCL_PIN EXT INP x0 x1 x2 x3 OUT y0 y1 y2 INTER END_PIN FUNCTION SDF 4 3 8 1001 001 0100 010 1100 010 1010 001 0010 101 0101 100 0011 101 0111 100 END_SDF END_FUNCTION END_GenP	TITLE GenP FORMAT SF AUTHOR Author DATE 03.11.2020 10:28:44 PROJECT Test DCL_PIN EXT INP x0 x1 x2 x3 OUT y0 y1 y2 INTER END_PIN FUNCTION SDF 4 3 6 ---- 001 -1-- 100 --1- 011 ---1 011 -11- 011 -1-1 010 END_SDF END_FUNCTION END_GenP	TITLE GenP FORMAT SF AUTHOR Author DATE 03.11.2020 10:28:44 PROJECT Test DCL_PIN EXT INP x0 x1 x2 x3 OUT y0 y1 y2 INTER END_PIN FUNCTION LOG 4 3 0 y0 = x1; y1 = x2⊕x3⊕x1*x2⊕ x1*x3; y2=1⊕x2⊕x3⊕x1*x2; END_LOG END_FUNCTION END_GenP

В таблице 3 представлено структурно-функциональное описание схемы GenP1 в виде системы частичных булевых функций и результирующее описание схемы в виде полинома Жегалкина (матричная и символьная формы).

Формы представления полиномов Рида-Маллера аналогичны формам представления полиномов Жегалкина с той разницей, что часть переменных представлена в инверсной форме (и только в ней).

Основной задачей алгоритма построения полиномов является преобразование исходного описания логической схемы к формату, воспринимаемому процедурами построения полиномиальных форм. Обращаем внимание, что система частичных булевых функций помечена комментарием /*T*/ в разделе «FUNCTION». Полином Жегалкина помечен комментарием /*G*/. Полином Рида-Маллера помечен комментарием /*RM*/.

Таблица 3. Описание схемы GenP1 в виде системы частичных булевых функций и полинома Жегалкина

Исходное задание	Полином Жегалкина в матричной форме	Полином Жегалкина в символьной форме
TITLE GenP1 FORMAT SF AUTHOR Author DATE 03.11.2020 10:28:44 PROJECT Test DCL_PIN EXT INP x0 x1 x2 x3 OUT y0 y1 y2 INTER END_PIN FUNCTION /*T*/ SDF 4 3 10 1001 0-1 0100 010 1100 010 1010 -01 0010 101 0101 100 0011 1-1 0111 100 0000 --- 0001 --- END_SDF END_FUNCTION END_GenP1	TITLE GenP1 FORMAT SF AUTHOR Author DATE 03.11.2020 10:28:44 PROJECT Test DCL_PIN EXT INP x0 x1 x2 x3 OUT y0 y1 y2 INTER END_PIN FUNCTION /*G*/ SDF 4 3 5 ---- 101 1--- 100 -1-- 111 11-- 100 -1-1 110 END_SDF END_FUNCTION END_GenP1	TITLE GenP1 FORMAT SF AUTHOR Author DATE 03.11.2020 10:28:44 PROJECT Test DCL_PIN EXT INP x0 x1 x2 x3 OUT y0 y1 y2 INTER END_PIN FUNCTION /*G*/ LOG 4 3 0 y0 = 1⊕x0⊕x1⊕x0*x1⊕ x1*x3; y1 = x1⊕x1*x3; y2 = 1⊕x1; END_LOG END_FUNCTION END_GenP1

В процессе реализации алгоритма строится 12 различных функций, представленных в таблице 4. В таблице представлены имена функций, описание построенных форм, значение специальных признаков, соответствующих каждой из функций. Признак «mode» управляет выбором процедуры построения полинома, признак «regime» соответствует выбору типа полинома (Жегалкина, или Рида-Маллера), признак «prizn_sdf» управляет формой представления полинома (матричная, или символьная).

В процессе построения функций выполняется преобразование исходных заданий систем булевых функций к форме, приемлемой для процедур построения полиномов (таблица 1). При этом система частичных булевых функций доопределяется нулевыми значениями в позициях, где она не определена. Для системы ДНФ выполняется перевод в векторную форму (система определяется на всех наборах таблицы истинности).

Таблица 4. Описание функций построения полиномов

Имя функции	Построенная форма	Значение признака «mode»	Значение признака «regime»	Значение признака «prizn_sdf»
polG_chast	Полином Жегалкина в символьной форме для частичных булевых функций	1	0	0
polG_chast_sdf	Полином Жегалкина в матричной форме для частичных булевых функций	1	0	1
polG_SOP	Полином Жегалкина в символьной форме для систем ДНФ	1	0	0
polG_SOP_sdf	Полином Жегалкина в матричной форме для систем ДНФ	1	0	1
polG_tabl	Полином Жегалкина в символьной форме для таблиц истинности	4	0	0
polG_tabl_sdf	Полином Жегалкина в матричной форме для таблиц истинности	4	0	1
polRM_chast	Полином Рида-Маллера в символьной форме для частичных булевых функций	2	1	0
polRM_chast_sdf	Полином Рида-Маллера в матричной форме для частичных булевых функций	2	1	1
polRM_SOP	Полином Рида-Маллера в символьной форме для систем ДНФ	2	1	0
polRM_SOP_sdf	Полином Рида-Маллера в матричной форме для систем ДНФ	2	1	1
polRM_tabl	Полином Рида-Маллера в символьной форме для таблиц истинности	4	1	0
polRM_tabl_sdf	Полином Рида-Маллера в матричной форме для таблиц истинности	4	1	1

Результаты экспериментального исследования. Функции построения полиномов (таблица 4) реализованы на языке C++ и включены в экспериментальную систему логического проектирования FLC2 [3], разработанную в лаборатории логического проектирования ОИПИ НАН Беларуси.

Основная задача исследования – оценить сложность схем в результате их полиномиального представления. Результаты исследования приведены в таблице 5. Исследование проведено на примерах известной серии для оценки алгоритмов (URL: <http://www1.cs.columbia.edu/~cs6861/sis>). Имена схем представлены в столбце 1, параметры схем (n – число входов, m – число выходов) представлены в столбце 2.

Для каждого примера выполнялась процедура построения полиномов Жегалкина и Рида-Маллера. В столбцах 3, 4, 5 представлены числа элементарных конъюнкций (термов) в исходном описании, полиноме Жегалкина, и в полиноме Рида-Маллера соответственно. В столбцах 6, 7, 8 представлены числа литералов в исходном описании, полиноме Жегалкина, и в полиноме Рида-Маллера соответственно. В столбце 9 представлены значения относительных выигрышей при представлении полиномами Жегалкина – отношения числа литералов в исходной форме и в полиномах Жегалкина. В столбце 10 представлены значения относительных выигрышей при представлении полиномами Рида-Маллера – отношения числа литералов в исходной форме и в полиномах Рида-Маллера.

Таблица 5. Исследование сложности полиномиальных форм описания функционирования логических схем

Пример	Параметры схемы n, m	Число термов в исходной форме	Число термов в полиноме Жегалкина	Число термов в полиноме Риди-Маллера	Число литералов в исходной форме	Число литералов в полиноме Жегалкина	Число литералов в полиноме Риди-Маллера	Выигрыш в полиноме Жегалкина	Выигрыш в полиноме Риди-Маллера
1	2	3	4	5	6	7	8	9	10
add6	12, 7	1092	125	125	8748	654	654	13,38	13,38
adm4	9, 8	512	196	196	11790	1050	1050	11,23	11,23
b12	15, 9	431	320	81	1923	1209	347	1,59	5,54
ex7	16, 5	123	100	100	788	563	563	1,40	1,40
f51m	8, 8	256	72	72	8192	264	264	31,03	31,03
intb	15, 7	664	3293	1915	5594	21540	12838	0,26	0,44
life	9, 1	512	183	99	1260	792	596	1,59	2,11
m4	8, 16	256	1429	410	17464	5630	2023	3,10	8,63
m181	15, 9	430	324	81	1921	1225	352	1,57	5,46
max512	9, 6	512	714	772	14544	3548	3792	4,10	3,84
rd53	5, 3	32	17	17	210	45	45	4,67	4,67
rd73	7, 3	147	60	60	876	189	189	4,63	4,63
ryu6	16, 1	112	79	63	624	624	464	1,00	1,34
sym10	10, 1	837	265	265	8370	1300	1300	6,44	6,44
t3	12, 8	152	1827	113	1630	11522	718	0,14	2,27
z4	7, 4	128	28	28	1792	89	89	20,13	20,13
z9sym	9, 1	420	209	172	3780	756	636	5,00	5,94

Заклучение.

Использование полиномиальных представлений систем булевых функций для описания поведения дискретных устройств позволяет значительно сократить сложность описания устройств. Экспериментальные исследования показали, что в некоторых случаях количество литералов сокращается в 30 раз. Только в одном из 17 примеров получен проигрыш. Полиномиальные представления систем булевых функций позволяют легко синтезировать схемы в базисе элементов AND/EXOR, что позволяет сокращать число элементов в схемах. Кроме того, такие схемы считаются хорошо диагностируемыми.

Список литературы

- [1] Representations of Discrete Functions (ed. by T. Sasao and M. Fujita). Kluwer Academic Publishers, Boston/London/Dordrecht. 1996. 332 p.
- [2] Закревский А.Д., Торопов Н.Р. Полиномиальная реализация частичных булевых функций и систем / Минск: Ин-т технической кибернетики НАН Беларуси. 2001. 200 с.
- [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.

POLYNOMIAL IMPLEMENTATION OF SYSTEMS OF BOOLEAN FUNCTIONS IN PROBLEMS OF LOGIC DESIGN

*N.A. Kirienko,
Senior Researcher of UIIP of NAS of Belarus,
PhD of Technical Sciences,
Associate Professor*

*The United Institute of Informatics Problems of the National Academy of Sciences of Belarus, Republic Belarus
E-mail: kir@newman.bas-net.by*

Abstract. Functional descriptions of designed digital devices have hundreds of input variables and tens of thousands of equations. In this regard, the problem arises of effective representation of the initial descriptions of digital devices, which will allow obtaining a sufficiently optimal solution and reducing the time of digital device synthesis. It is proposed to use a polynomial representation of systems of Boolean functions, which, as a rule, describe the behavior of discrete devices.

The representation of the system of Boolean functions in the form of Zhegalkin and Reed-Muller polynomials is considered. Algorithms for constructing polynomials for various forms of specifying systems of Boolean functions are described: systems of disjunctive normal forms (DNFs), partial Boolean functions, and truth tables.

Keywords: logic circuit, functional description of the digital device, systems of Boolean functions, Zhegalkin polynomial, Reed-Muller polynomial.

УДК 004.732

ОРГАНИЗАЦИЯ И УПРАВЛЕНИЕ ЗОНОЙ WI-FI В УЧРЕЖДЕНИИ ОБРАЗОВАНИЯ ДЛЯ РАБОТЫ С BIGDATA



М.В. Романюк

*Инженер-программист отдела сетевых технологий Центра информатизации и инновационных разработок БГУИР, ассистент кафедры информатики БГУИР, магистрант кафедры ПИКС
romanuk@bsuir.by*



3

А.Н. Марков

*Старший преподаватель кафедры ПОИТ, 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 отдела сетевых технологий Центра информатизации и инновационных разработок БГУИР, ассистент кафедры информатики БГУИР, магистрант кафедры проектирования информационно-компьютерных систем БГУИР.

А.Н. Марков

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

Аннотация. Рассмотрена специфика организации беспроводных сетей на базе Wi-Fi в учреждениях образования с учетом генерации больших объемов данных. Приведены данные о Wi-Fi зоне в БГУИР. Рассмотрены перспективы развития беспроводных сетей Wi-Fi в учреждениях образования.

Ключевые слова: организация зоны Wi-Fi, Wi-Fi в учреждении образования, BigData

Введение.

В настоящее время доступ к всемирной паутине перестал быть чем-то сверхординарным и воспринимается скорее, как обязательный пункт всех сфер жизни. Доступ к образовательному контенту, образовательным площадкам, курсам, тестам и другим учебным материалам позволяет расширить сферу охвата образовательного процесса для всех категорий лиц независимо от местоположения, сферы деятельности и возраста. С учетом особенностей контента генерируются большие объемы данных, что неумолимо влечет за собой корректную организацию и управление зоной беспроводных сетей. В классических высших учреждениях образования расширяются возможности использования образовательных технологий дополненной реальности, проведения дистанционных лекционных занятий с приглашенными преподавателями с мировыми именами, расширяются границы потребления контента по учебным дисциплинам и курсам, что влечет за собой передачу больших данных. Организованная зона Wi-Fi в учреждениях образования позволяет не зависеть от “провода” и быть “онлайн” даже в моменты перемещения по кампусу. Таким образом, для организации передачи BigData должны быть организованные высокоскоростные беспроводные бесшовные каналы связи.

Специфика организации беспроводной сети в учреждении образования с генерацией больших объемов данных.

Традиционно беспроводные сети в организациях разделяются на гостевые сети и сети для сотрудников.

В учреждениях образования данная концепция может быть реализована путем выделения общей беспроводной сети для студентов, сотрудников, имеющих учетные данные. Доступ из данной сети должен осуществляться исключительно к сети интернет без доступа во внутреннюю сеть организации. Примером гостевой сети в учреждениях образования может служить международная сеть для образовательных и исследовательских учреждений Eduroam. Сервис eduroam (EDUcation ROAMing) позволяет пользователям (исследователям, преподавателям, студентам, научным сотрудникам) из участвующих в сервисе организаций безопасно получать доступ в Интернет из любого учреждения, имеющего подключение к системе роуминговой идентификации, построенной на базе RADIUS-серверов учреждений-участников. Подробнее принцип работы сервиса eduroam будет описан далее.

Для некоторых сотрудников (администраторов сети, разработчиков, администрации университета) может создаваться беспроводная сеть, из которой возможен доступ к ресурсам внутренней сети. Однако ввиду использования устройств, на которых не может быть подтверждено отсутствие вредоносного ПО, а также с устройств, которые не включены в домен, необходимо перед предоставлением доступа к внутренним ресурсам компании выполнять сканирование сетевого трафика посредством межсетевых экранов. В Республике Беларусь межсетевые экраны для беспроводных сетей должны иметь сертификат соответствия государственному стандарту СТБ 34.101.75-2017, п.7.3, п.7.7 [1].

При подключении пользователей к беспроводным сетям необходимо выполнять аутентификацию и авторизацию пользователей. Традиционно аутентификация и авторизация в корпоративных сетях осуществляется при помощи RADIUS-сервера. Алгоритм подключения к беспроводной сети при этом выглядит следующим образом:

1. Когда пользователь пытается подключиться к Wi-Fi сети, беспроводная точка доступа запрашивает идентификационные данные у пользователя.
2. Пользователь вводит свои учетные данные (обычно это имя пользователя и пароль).
3. Беспроводная точка доступа передает идентификационные данные на Radius-сервер.
4. Radius-сервер проверяет идентификационные данные пользователя, используя базу данных учетных записей (LDAP, AD).
5. Если идентификационные данные верны, то Radius-сервер передает беспроводной точке доступа разрешение на доступ к сети для пользователя.
6. Пользователь получает доступ к сетевым ресурсам.

Таким образом, пользователь не имеет доступа к сети пока не пройдет процессы аутентификации и авторизации.

С технической точки зрения Eduroam работает следующим образом: в каждой стране, которая внедряет данную технологию, есть сервера, хранящие информацию о пользователях. Эти сервера представлены RADIUS-сервером национального сегмента и RADIUS-серверами учреждений, подключенных к сервису. Аутентификация пользователей в любом сегменте Eduroam проводится их родительскими организациями с использованием тех же учетных данных и методов, что и при локальной аутентификации при нахождении в сети таковой, тогда как организация доступа к интернету лежит на посещаемом учреждении. Таким образом, роль системы заключается в переадресации учетных данных пользователей по иерархической системе RADIUS-серверов из посещаемой организации в родительскую, где они могут быть верифицированы. Например, если пользователь, числящийся в системе БГУИР приехал в гамбургский университет и пытается там войти в сеть Eduroam, то запрос отправляется сначала на RADIUS-сервер гамбургского университета, оттуда – на сервер национального сегмента Германии, далее на сервер Eduroam верхнего уровня, потом на сервер национального сегмента Республики Беларусь и после этого на сервер БГУИР, который идентифицирует пользователя и отправляет обратно информацию об авторизации по той же цепочке серверов [2].

С учетом специфики финансирования учреждений образования, при подборе оборудования необходимо исходить из приоритетности зоны покрытия, плотности пользователей, необходимой скорости работы, поддержки роуминга пользователей между точками доступа.

Так, например, в высоконагруженных зонах (в администрации университета, а также в отделах, связанных с активной работой с беспроводной сетью и генерацией большого трафика и данных) могут устанавливаться точки доступа высшего ценового и производительного сегмента с поддержкой роуминга и наиболее современных стандартов беспроводной связи (Cisco, HPE Aruba), в то время как в менее приоритетных зонах могут устанавливаться менее дорогие и производительные точки (D-Link, Ruijie Networks, TP-Link, Huawei).

Wi-Fi в БГУИР.

На начало 2023 года в БГУИР под централизованным контролем специалистов функционирует 94 точки доступа, которые представлены 3 зонами беспроводной сети, построенными с использованием оборудования различных вендоров.

Основная зона Wi-Fi состоит из 79 точек доступа D-Link моделей DWL-8610, DWL-8620, DWL-6610, работающих под управлением беспроводного контроллера D-Link DWC-2000, а также одиночных точек доступа DAP-2699, DAP-2360, которые могут управляться при помощи программного контроллера D-Link CWM-100.

Зона Wi-Fi в администрации университета построена на базе 10 точек доступа Aruba IAP-325. Данные точки доступа управляются при помощи Aruba Virtual Controller, который автоматически размещается на одной из точек доступа для обеспечения отказоустойчивости работы зоны беспроводной сети.

Высоконагруженная зона Wi-Fi, которая обеспечивает работу программистов, администраторов сети на одном из этажей Центра информатизации и инновационных разработок БГУИР построена на базе 5 точек доступа Cisco AIR-AP3802I. Точки управляются посредством виртуального контроллера Cisco Mobility Controller, который также размещается непосредственно на одной из точек доступа и имеет возможность мигрировать на другие выбранные администратором сети точки доступа.

Беспроводные контроллеры необходимы для централизованного управления зоной беспроводной сети: подбора мощности соседних точек доступа при сбоях в сети Wi-Fi, настройки SSID, параметров авторизации, частотных диапазонов, а также для работы роуминга беспроводных клиентов с использованием протоколов 802.11k/v/r.

На всех функционирующих точках доступа создана сеть eduRoam, которая позволяет пользователям (исследователям, преподавателям, студентам, научным сотрудникам) из участвующих в сервисе организаций безопасно получать доступ в Интернет из любого учреждения, имеющего подключение к системе EduRoam.

Перспективы развития Wi-Fi сетей в учреждениях образования с учетом BigData.

Хотя, как говорилось ранее, большая часть университета централизованно покрыта общеуниверситетскими сетями Wi-Fi, на некоторых кафедрах и в структурных подразделениях университета остаются точки доступа, не контролируемые единым центром или сервисом, что потенциально создает угрозу информационной безопасности. На 2023 год запланировано расширение общеуниверситетской сети Wi-Fi с заменой одиночных и устаревших точек доступа на современные Wi-Fi точки с поддержкой стандарта Wi-Fi 6. Использование стандарта Wi-Fi 6 позволит увеличить скорость доступа к сети интернет, увеличить стабильность работы при высокой плотности пользователей. Однако при проектировании расширения зоны беспроводной сети необходимо производить радиообследование помещений, которое включает в себя измерения затухания стен, перекрытий, анализ уже используемых каналов другими точками доступа в прилегающих помещениях. Базово зонирование с учетом радиообследования на анализ затуханий, а также соответствие требований и стандартов в области беспроводных сетей

необходимо проводить с учетом проектного решения. Однако, с учетом специфики финансирования учреждений образования, зачастую, проектные решения не предусмотрены.

В целом, при развитии проводных и беспроводных сетей для доступа к сети Интернет с учетом новых технологий, оборудования, стандартов так или иначе ведет к трансформации доступа к большим данными и в учреждениях образования.

Список литературы

[1] Информационные технологии. Системы обнаружения и предотвращения вторжений. Общие требования: СТБ 34.101.75-2017. – Введ. 01.10.2017. – Минск: Госстандарт, 2017. – 18с.

[2] Принцип работы сети Eduroam: [Электронный ресурс]. URL: <https://eduroam.org/how/>. (Дата обращения: 29.03.2023).

ORGANIZATION AND MANAGEMENT OF A WI-FI ZONE IN AN EDUCATIONAL INSTITUTION FOR WORKING WITH BIGDATA

M.V. Romaniuk

*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 of the department,
Deputy head of the Center of
Informatization and Innovation
Elaborations*

*Center of Informatization and Innovation Elaborations of the Belarusian State University of Informatics and
Radioelectronics, Republic of Belarus
E-mail: romanuk@bsuir.by, a.n.markov@bsuir.by*

Abstract. The specifics of the organization of wireless networks based on Wi-Fi in educational institutions, considering the generation of large amounts of data, are considered. The data about an existing Wi-Fi zone in BSUIR is given. The prospects for the development of wireless networks based on Wi-Fi in educational institutions are considered.

Keywords: organization of a Wi-Fi zone, Wi-Fi in an educational institution, BigData.

УДК 004.65+004.85

BIG DATA И МАШИННОЕ ОБУЧЕНИЕ



Н.И. Потапенко
Старший преподаватель
кафедры инженерной
психологии и эргономики



В.А. Буд-Гусаим
Студент кафедры инженерной
психологии и эргономики



А.Н. Василькова
Ассистент кафедры
инженерной психологии и
эргономики, магистр
a.vasilkova@bsuir.by

Н.И. Потапенко

Старший преподаватель кафедры инженерной психологии и эргономики Белорусского государственного университета информатики и радиоэлектроники.

А.А. Войтович

Студент кафедры инженерной психологии и эргономики Белорусского государственного университета информатики и радиоэлектроники.

А.Н. Василькова

Ассистент кафедры инженерной психологии и эргономики Белорусского государственного университета информатики и радиоэлектроники, магистр.

Аннотация. Большие данные (Big Data) является мощным инструментом для решения и переосмысления проблем в IT сфере, бизнесе, медицине и т.д. Поскольку объем данных продолжает расти с невероятной скоростью, способность правильно их обрабатывать стала важнейшим ключом к созданию богатого источника для обучения алгоритмов. Для работы с Big Data используют различные методы машинного обучения (ML), такие как деревья решений, алгоритмы кластеризации, нейросети, которые способны учиться на обрабатываемых ими данных. Цель статьи – дать представление о том, как организации могут использовать ML и Big Data, чтобы получить конкурентное преимущество в современном мире, управляемом данными.

Ключевые слова: большие данные, машинное обучение, алгоритм, нейросеть, обработка наборов данных, зависимость, самообучение.

Введение.

В современную цифровую эпоху объем генерируемых данных растет экспоненциально, что создает требования наличия мощных и эффективных методов обработки данных. Однако из-за объема и сложности Big Data традиционные методы обработки оказались неточными и неэффективными. Здесь на помощь приходит ML, предоставляющее передовые алгоритмы, которые автоматически выявляют закономерности, делают прогнозы и учатся на основе данных.

Использование ML для обработки Big Data практикуется все чаще с каждым днем. Объединив эти две технологии, организации могут анализировать свои данные, принимать более обоснованные решения и повышать общую эффективность своей работы. Кроме того, самообучающиеся алгоритмы ML постоянно развиваются. Объединяя Big Data и ML, алгоритмы Machine Learning помогают нам справляться с непрерывным потоком данных, а объемы и разнообразие потоков данных прокачивают алгоритмы, делая их более совершенными. В статье рассматриваются проблемы и потенциал этого подхода, приводятся примеры его применения в различных областях и дается обзор последних исследований в этой области. Цель статьи – дать представление о том, как организации могут использовать ML и Big Data, чтобы получить конкурентное преимущество в современном мире, управляемом данными.

Основная часть.

Big Data- это коллекция различных сложных наборов данных. С каждым днем их становится все больше и больше, в результате, их трудно хранить и обрабатывать. Чтобы Big Data превратилась в полезный инструмент для человека, ее обработка должна включать в себя сбор, отбор, хранение, поиск, совместное использование, передачу, анализ и визуализацию этих данных. Работа с большими данными строится вокруг пяти основных принципов (с англ. V's of Big Data: Volume, Velocity, Variety, Veracity, Value):

Объем: количество собираемой компаниями информации действительно огромно, поэтому ее объем становится критическим фактором в аналитике. **Скорость:** практически все происходящее вокруг нас (поисковые запросы, социальные сети и т. д.) очень быстро генерирует новые данные, многие из которых могут быть использованы в бизнес-решениях.

Разнообразие: генерируемая информация неоднородна и может быть представлена в различных форматах, таких, например, как видео, текст, базы данных, числа, сенсорные данные и т. д. **Понимание типов больших данных** является ключевым фактором для раскрытия их ценности. **Достоверность:** данные высокой достоверности содержат много записей, которые ценны для анализа и которые вносят значимый вклад в общие результаты. Данные с низкой достоверностью содержат высокий процент бессмысленной информации, которая называется шумом.

Ценность: возможность трансформировать большие данные в бизнес-решения.

Big Data пополняется из большого количества источников. Основные из них:

- Социальные сети
- Облачные хранилища
- Веб-сайты
- Интернет Вещей (IoT)

Machine Learning (ML) – это форма искусственного интеллекта, которая позволяет анализировать данные, принимать на их основе решения, перекрывать некоторые проблемы в сфере работы, автоматизируя большую часть работы с информацией. Он работает путем изучения данных и выявления закономерностей и требует минимального вмешательства человека. Главной особенностью ML является его возможность самообучаться с помощью обработанных данных. Почти любую задачу, которую можно выполнить с помощью шаблона или набора правил, определяемых данными, можно автоматизировать с помощью машинного обучения. Это позволяет компаниям трансформировать процессы, которые ранее были доступны только людям, например, ответы на звонки в службу поддержки клиентов, ведение бухгалтерского учета и просмотр резюме.

Использование алгоритмов ML для анализа Big Data - логичный шаг для компаний, стремящихся максимизировать потенциальную ценность своих данных. Они могут быть реализованы во всех элементах работы с Big Data, включая маркировку и сегментацию данных, анализ данных и моделирование сценариев. Алгоритмы учатся на данных по мере их обработки, в отличие от традиционных аналитических систем, основанных на правилах, которые следуют четким инструкциям.

Примеры использования Big Data и ML. Сегментация. Для любого бизнеса одним из важнейших ключей для успеха является целевая аудитория. Каждое предприятие должно понимать аудиторию и рынок, на которые оно хочет ориентироваться. Вот почему предприятиям необходимо проводить исследования рынка, которые могут глубоко проникнуть в мысли потенциальных клиентов и предоставить ценные данные. С этим может помочь ML, использующее контролируемые и неконтролируемые алгоритмы для точной интерпретации потребительских моделей и поведения. Средства массовой информации и индустрия развлечений используют ML, чтобы понимать, что нравится и не нравится их аудитории, и нацеливать на нее нужный контент. Например, Netflix, Facebook, Google, Twitter и т.д.

Изучение поведения клиентов. ML также помогает компаниям исследовать поведение

аудитории и создавать прочную структуру своих клиентов. Эта система ML, известная как пользовательское моделирование, является прямым результатом взаимодействия человека с компьютером. Он собирает данные, чтобы захватить мысли пользователя и позволить бизнес-предприятиям принимать разумные решения. Amazon, Uber, LinkedIn и другие полагаются на системы моделирования пользователей, чтобы знать своих пользователей наизнанку и делать соответствующие предложения.

Предсказание трендов. Алгоритмы ML используют Big Data для изучения будущих тенденций и прогнозирования их для бизнеса. С помощью этих данных ML может самостоятельно учиться новому и улучшать свои аналитические навыки каждый день. Таким образом, он не просто вычисляет данные, но ведет себя как интеллектуальная система, которая использует прошлый опыт для формирования будущего. Бренд кондиционеров может полагаться на ML, чтобы прогнозировать спрос на кондиционеры в следующем сезоне и соответствующим образом планировать свое производство.

Помощь в принятии решений. ML использует метод, называемый анализом временных рядов, который позволяет анализировать массив данных вместе. Это отличный инструмент для сбора и анализа данных, облегчающий менеджерам принятие решений на будущее. Предприятия, особенно розничные торговцы, могут использовать этот метод, усиленный машинным обучением, для предсказания будущего с похвальной точностью.

Как это работает?

Структура использования ML с Big Data (MLBiD) состоит из 5 компонентов: Machine Learning, Big Data, пользователь, система, сфера использования. Все эти компоненты взаимодействуют между собой обоюдно. Так, например, Big Data служит «материалом» для ML, оно же генерирует новые данные, которые становятся частью Big Data. При работе с данными ML проходит 3 этапа: предварительная обработка, обучение, оценка. Предварительная обработка. Она преобразует исходные данные в соответствующую для последующих этапов обучения форму. Исходные данные, скорее всего, неструктурированные, неполные, непоследовательные и содержат много «шума». Этап предварительной обработки преобразует такие данные посредством очистки, извлечения, преобразования и объединения данных.

Обучение. На этапе обучения выбирается и настраивается алгоритм машинного обучения для получения желаемых результатов с использованием предварительно обработанных входных данных. Некоторые модели обучения, в частности репрезентативное обучение, также могут быть использованы для предварительной обработки данных. Во время обучения алгоритм изучает закономерности и отношения в данных. Оценка. Алгоритм и его работу необходимо оценить, чтобы определить точность и производительность. Это предполагает тестирование модели на отдельном наборе данных, который не использовался во время обучения. Результаты оценки могут привести к корректировке параметров выбранных алгоритмов обучения и/или выбору других алгоритмов. Однако работая с Big Data объем данных настолько велик, что обрабатывать в памяти все сразу не является возможным. В таких случаях используют алгоритмы, методы которых умеют работать с слишком большими данными. Один из подходов заключается в использовании сред распределенных вычислений, таких как Apache Spark или Hadoop, которые позволяют обрабатывать данные параллельно на нескольких машинах.

Другой подход заключается в использовании метода, называемого «онлайн-обучение», при котором модель постоянно обновляется по мере поступления новых данных. Это может быть полезно при работе с динамическими данными.

Кроме того, существуют такие методы, как извлечение признаков и уменьшение размерности, которые можно использовать для уменьшения размера данных при сохранении важной информации.

Плюсы и минусы.

Минусы:

Дубликаты – на просторах Big Data есть много дублированной информации. Это

происходит, когда несколько выборок имеют одинаковые сущности.

Маркировка – из-за большого количества данных и их источников сложно маркировать наборы данных правильным образом, чтобы при поиске определенной информации не возникало проблем и временных затрат. В результате нужный набор данных может не охватывать все специфические для пользователя контексты.

Хранение - работа с Big Data требует значительного объема памяти для хранения, что может быть дорого и сложно в реализации.

Дискретизация данных - некоторые алгоритмы ML, такие как дерево решений, могут работать только с дискретными данными. Для дискретизации данным в Big Data требуются нестандартные методы. Они трудоёмкие, затратные и непрактичные.

Стоимость - чтобы правильно настроить рабочее ПО необходимо много различных ресурсов. Поэтому использование ML не является практичным решением для всех.

Плюсы:

Автоматизация – ML сокращает время и рабочую нагрузку на человека, выполняя огромную часть аналитической работы без малейшего участия человека. Автоматизация надежна, эффективна и быстра.

Самосовершенствование – обрабатывая данные Big Data алгоритмы ML способны благодаря этим же данным учиться. При этом сложно определить до какого уровня ML сможет совершенствоваться.

Широкая сфера применения - эта технология имеет очень широкий спектр применения. ML играет роль практически во всех областях, таких как образование, медицина, наука, банковское дело и бизнес. Это создает больше возможностей.

Возможность предугадывать тренды – ML может анализировать большие объемы данных и обнаруживать определенные тенденции и закономерности, которые не были бы очевидны людям. Например, для веб-сайта электронной коммерции, такого как Amazon, он служит для понимания поведения пользователей в Интернете и истории покупок, чтобы помочь им найти правильные продукты, предложения и напоминания, относящиеся к ним.

Заключение. В статье обсуждается пересечение Big Data и ML, эти технологии содержат в себе большой потенциал для использования в широком спектре сфер. Big Data предоставляет объемы информации, которую можно использовать для обучения алгоритмов ML. Они же могут создавать закономерности и решения, которые люди не могут идентифицировать самостоятельно. В статье также рассматриваются преимущества и проблемы работы с Big Data и ML, в том числе потребность в мощных вычислительных ресурсах, проблемы с точностью данных и их источниками. В целом в статье подчеркивается огромное значение этих двух областей для преобразования различных отраслей, от здравоохранения до финансов.

Переход на ML может стать большим скачком для бизнеса. Это влечет за собой переопределение рабочих процессов, архитектуры, сбора и хранения данных, аналитики и других модулей.

Пошаговый подход лучше всего подходит для такого перехода. Во-первых, предприятиям необходимо разработать надежную стратегию на основе искусственного интеллекта и ML, которая соответствует их бизнес-целям. Во-вторых качественные данные являются ключом к реализации всего потенциала инструментов ML. Компании должны создать корпоративную культуру вокруг данных. Правильные люди и правильные данные могут иметь огромное значение. Наконец, время имеет решающее значение, и предприятия должны действовать быстро.

Поскольку объем данных со временем продолжает увеличиваться, сбор данных и управление ими становится важнейшей задачей для бизнеса. Управление и определение смысла собранных таким образом данных для улучшения маркетинговой стратегии и увеличения доходов – это большая битва. Таким образом, ML незаменимо для решения задач, связанных с Big Data, и раскрытия скрытых моделей, знаний и идей из наборов данных, чтобы превратить

потенциал последних в реальную ценность для принятия деловых решений и научных исследований. Все минусы использования этих технологий вопрос времени, так как ошибки ML при работе с Big Data говорят о том, что есть еще огромное пространство для совершенствования.

Список литературы

[1] Big Data + Machine Learning = Love// Хабрахабр. URL: <https://habr.com/ru/company/first/blog/692978/> (дата обращения: 19.03.2023).

[2] Chithrai Mani. How Is Big Data Analytics Using Machine Learning?// Forbes Technology Council. URL: <https://www.forbes.com/sites/forbestechcouncil/2020/10/20/how-is-big-data-analytics-using-machine-earning/?sh=602cf9aa71d2> (дата обращения: 20.03.2023).

[3] Big Data от А до Я. Часть 1: Принципы работы с большими данными, парадигма MapReduce// Хабрахабр. URL: <https://habrahabr.ru/company/dca/blog/267361/> (дата обращения: 9.05.2019).

[4] N. Carah and D. Angus, "Algorithmic brand culture: participatory labour, machine learning and branding on social media," *Media, Cult. Soc.*, vol. 40, no. 2, pp. 178–194, 2018

[5] N. Japkowicz and M. Shah, *Evaluating Learning Algorithms: A Classification Perspective*: Cambridge University Press, 2011.

BIG DATA AND MACHINE LEARNING

N.I. Potapenko

*Senior Lecturer, Department of
Engineering Psychology and
Ergonomics*

V.A. Bud-Husaim

*Student of the Department of
Engineering Psychology and
Ergonomics*

A.N. Vasilkova

*Assistant of the Department of
Engineering Psychology and
Ergonomics, master*

Department of Engineering Psychology and Ergonomics

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

E-mail: a.vasilkova@bsuir.by

Annotation. Big Data is the vehicle for solving and rethinking problems in IT, business, medicine, etc. The volume of data search continues to grow at an incredible rate, the ability to choose the right one has become an important key factor in identifying sources for studying algorithms. learn to neutralize mitochondrial data. The purpose of the article is to enable the use of representation as an organization, OD and big data in order to gain a competitive advantage in the important world, to manage the volume of data.

Keywords: big data, machine learning, algorithm, neural network, dataset processing, addiction, self-learning.

УДК 159.96

ОРГАНИЗАЦИЯ ПСИХОПРОФИЛАКТИЧЕСКОЙ РАБОТЫ В ВОЙСКАХ



В.Н. Шведко

*Аспирант кафедры инженерной психологии и эргономики БГУИР
vasilisa14194@mail.ru*



Т.В. Казак

*заведующий кафедрой инженерной психологии и эргономики, член-корреспондент Международной академии психологических наук, доктор психологических наук, профессор
kazak@bsuir.by*

В.Н. Шведко

Окончила БГУ (факультет журналистики) и магистратуру БГУИР. Является аспирантом БГУИР.

Т.В. Казак

Заведующий кафедрой инженерной психологии и эргономики БГУИР, доктор психологических наук Республики Беларусь, доктор психологических наук Российской Федерации, член-корреспондент Международной академии психологических наук, профессор.

Аннотация. В настоящее время под психопрофилактикой понимается комплекс мероприятий, направленных на предупреждение развития психических расстройств путем предотвращения действия психотравмирующих факторов, раннее выявление лиц с признаками нервно-психической неустойчивости (НПН), психических расстройств, предрасположенных к алкоголизму, наркомании, а также предупреждение хронизации психических заболеваний путем проведения лечебно-оздоровительных мероприятий нуждающимся военнослужащим.

Ключевые слова: Психическое здоровье, военная служба, служба по контракту, боевые действия, экстремальная ситуация, боевая обстановка, национальная безопасность, психопрофилактика.

Введение.

Различают три вида психопрофилактики: первичную – предупреждение нервно-психических болезней; вторичную – система лечебно-профилактических и реабилитационных мероприятий по предотвращению рецидивов болезни; третичную – предупреждение хронизации болезни и инвалидности пациента.

Первичная психопрофилактика включает систему мероприятий, предусматривающих изучение вредных воздействий на организм человека, его психического состояния и возможностей предупреждения психических и психогенно обусловленных болезней. Указанные меры имеют целью повышения сопротивляемости психики к негативным факторам среды и предупреждение самого факта возникновения болезни. На данном уровне профилактики решающее значение в охране здоровья военнослужащих имеет совокупность социально-экономических условий, рациональная организация их служебной деятельности, быта, отдыха и т.д. При этом медицинская служба занимает не основные, а рекомендательные (консультативные) позиции.

Особое значение имеет оценка и прогнозирование психического здоровья. Речь идет, во-первых, об определении вероятности развития психического расстройства и, во-вторых, об оценке качества психического функционирования и перспектив поведения военнослужащих в различных условиях их жизни и деятельности. Несомненно, значимым для решения вопроса о первичной психопрофилактике, предупреждения самого факта заболевания, является знание военными врачами и военными психологами клинических проявлений пограничных состояний, что позволит им проводить обоснованные, целенаправленные, индивидуальные профилактические мероприятия.

Вторичная психопрофилактика - совокупность мер, направленных на предупреждение неблагоприятных последствий уже развившейся психической болезни или психосоматического заболевания. Имеется ввиду активное выявление военнослужащих с начальными проявлениями болезненных состояний, раннее начало лечения с использованием комплексной, как правило, патогенетической терапии, что способствует благоприятному течению болезни, более быстрому выздоровлению, препятствует хронизации болезненного процесса.

Большое значение в системе психопрофилактических мероприятий в войсках имеет выявление лиц с нервно-психической неустойчивостью, их учет и организация медицинского наблюдения за ними. Своевременное выявление таких может быть осуществлено только при совместных усилиях командиров, воспитателей, военных психологов и врачей. [1]. Действенность психологической помощи зависит от ее организационно-деятельностных основ. Главной фигурой, определяющей эту работу, является офицер. Однако, несмотря на существующие распоряжения и рекомендации, далеко не все должностные лица могут психологически грамотно организовать и оказывать такую помощь. И тому есть немало причин мотивационного, профессионального, организационного, собственно деятельного характера.

Зачастую на практике психологическая помощь в основном сводится к усилению контроля над подчиненными, стремлению перевоспитать их, к ужесточению наказаний за проступки. Менее всего многие офицеры-руководители стремятся разобраться в причинах нервно-психической неустойчивости подчиненных, перепоручая такого рода вопросы их непосредственным командирам и начальникам, чаще не готовым к их разрешению в силу недостатка опыта и необходимости решения ряда других более понятных им задач.

Исходя из перечисленного выше, необходимо целенаправленное формирование мотивационной основы деятельности офицеров по организации и проведению психологической помощи военнослужащим. Основными направлениями реализации могли бы стать:

- постоянное психологическое просвещение по вопросам психического здоровья и психопрофилактики нервно-психической неустойчивости в условиях воинской деятельности;
- побуждение офицеров к деятельному участию в этой работе за счет придания ей профессиональной, социальной, личной значимости с соответствующим закреплением в различных руководящих документах;
- пропаганда передового опыта личного участия в этой работе и др.

Следующей составляющей психологической помощи военнослужащим является профессиональная подготовленность (компетентность) офицеров к этой работе. Об этом свидетельствует опыт работы военных психологов [2, 3, 4].

Не менее важная сторона деятельностной составляющей - тщательный подбор кадров специалистов, призванных заниматься психопрофилактикой нервно-психической неустойчивости в узком смысле слова или (и) руководить этой работой в целом. Главное ограничение и противопоказание здесь - морально-нравственная сторона. Психопрофилактика нервно-психической неустойчивости (в узком смысле особенно) или руководство ею - вид психологической помощи нуждающемуся военнослужащему, а потому такие специалисты для начала должны быть сами нервно-психически устойчивыми. Кроме того, занятие этой деятельностью предполагает доступ к интимной, сугубо личностной информации об особенностях человека, обычно широко не афишируемой. А потому важно, чтобы специалисты умело и корректно ею пользовались, делились с кем-либо ее частями только в той мере, насколько это диктуется требованиями служебной целесообразности и профессионального долга.

Заключение.

Таким образом, практические основы психопрофилактики нервно-психической неустойчивости представляют собой совокупность мотивационных, профессиональных, организационных и деятельностных предпосылок, учет которых, на наш взгляд, позволит разрабатывать и реализовывать конкретные авторские программы оказания психологической

помощи военнослужащим в конкретных видах профессионального труда.

Список литературы

[1] Съедин С.И., Абдурахманов Р.А. Психологические последствия воздействия боевой обстановки (Учебное пособие). М.: Военный университет. 1992.

[2] Буянов В.И. Психологическая коррекция затруднений социально-психологической адаптации молодых офицеров к условиям службы в воинской части: Дис... канд. психол. наук. - М.: ВА им. Ф.Э.Дзержинского, 1996. - 254 с.

[3] Караяни А.Г., Бархаев Б.П., Первалов В.Ф., Сыромятников И.В. Взаимодействие офицера и психолога в изучении индивидуально-психологических особенностей военнослужащих // Психология и педагогика профессиональной деятельности офицера. Сер. «Библиотека офицера» / Под редакцией Б. П. Бархаева. - Москва, 2006. - С. 141-152.

[4] Султанмуратов Ю.М., Овчинников А.А., Султанова А.Н., Сычева Т.Ю. Индивидуально-психологические особенности военнослужащих, влияющие на успешность адаптации // Сибирский вестник психиатрии и наркологии. - 2017. - № 3 (96). - С. 50-56.

ORGANIZATION OF PSYCHOPROPHYLACTIC WORK IN THE MILITARY

V.N. Shvedko

*Belarusian State University of Informatics and
Radioelectronics, Minsk, Republic of Belarus (style
T-institution)*

T.V. Kazak

*Head of the Department of Engineering Psychology
and Ergonomics, Corresponding Member of the
International Academy of Psychological Sciences,
Doctor of Psychological Sciences, Professor*

*Belarusian State University of computer science and Radio Electronics, Republic of Belarus
E-mail: vasilisa14194@mail.ru; fedora220864@gmail.com*

Abstract. Currently, psychoprophylaxis is understood as a set of measures aimed at preventing the development of mental disorders by preventing the effects of traumatic factors, early detection of persons with signs of neuropsychiatric instability (NPN), mental disorders predisposed to alcoholism, drug addiction, as well as prevention of the chronization of mental illnesses by conducting therapeutic and recreational activities for military personnel in need.

Keywords: Mental health, military service, contract service, combat operations, extreme situation, combat situation, national security, psychoprophylaxis

УДК 004.032.26

ПРОГРАММА ДЛЯ ПОДСЧЕТА МИКРООРГАНИЗМОВ



Коркин Л.Р.

Ассистент кафедры Инженерной психологии и эргономики БГУИР, магистр техники и технологии, аспирант.
l.korkin@bsuir.by



Саевич К.Ф.

Профессор кафедры физикохимии материалов и производственных технологий БГЭУ, доктор биологических наук, профессор
konstantin.saevich@mail.ru

Л.Р. Коркин

Окончил БГУИР, специальность «Инженерно-психологическое обеспечение информационных технологий». Квалификация инженер-системотехник. В 2020 году окончил магистратуру по специальности «Управление безопасностью производственных процессов» с присвоением академической степени магистра техники и технологии.

К.Ф. Саевич

В 1984 защитил диссертацию на соискание учёной степени кандидата сельскохозяйственных наук, в 1992 - доктора биологических наук, с 1998 - профессор. Области научных интересов: Экология, Рациональное природопользование, Безопасность жизнедеятельности.

Аннотация. Для количественного состава микробиоты исследуемого продукта питания используют метод подсчета колоний микроорганизмов, выделенных из продуктов питания. Процесс усложняется тем, что в некоторых случаях метод подсчета не дает стопроцентный результат даже при применении счетчика колоний микроорганизмов. Принцип работы этого счетчика заключается в том, что оператор с помощью электропера наносит точки на дно чашки Петри в места нахождения колоний. Прибор регистрирует факт касания, а результат выводится на цифровое табло. Счетчик регистрирует только факт касания электропера, что не может застраховать конечный результат от ошибок оператора. В других случаях подсчет производится математически, т.е. проводится подсчет в определенной области, как правило, в одной четвертой чашки Петри, а после умножается на четыре. Есть и другие методы определения числа колоний микроорганизмов, но какой бы из существующих методов определения мы бы не выбрали, каждый из них имеет свои недостатки.

Ключевые слова. Машинное обучение, искусственный интеллект, микробиология, колонии микроорганизмов.

Введение.

Чтобы превратить распознавание объектов на изображении из идеи в реальность нам необходимо использовать такой вид машинного обучения как обучение нейронной сети (ее еще называют первыми ступенями развития искусственного интеллекта будущего), которое всегда начинается с подготовки материалов, а в нашем случае это размеченные и подготовленные должным образом изображения колоний (снимки колоний микроорганизмов (СКМ)), которые нужно будет распознавать для последующего подсчета микроорганизмов (количественного состава микробиоты исследуемого продукта питания). Для этого мы создаем систему распознавания СКМ (СРСКМ), которая будет способна выявлять наличие колоний микроорганизмов (КМ) на снимках, а также проводить подсчёт выросших на чашке Петри. Это позволит выявлять некачественные продукты питания, а также поможет повысить продовольственную безопасность и качество жизни людей.

Актуальность.

Так как в настоящее время при контроле качества продуктов питания или лекарственных средства пользуются таксономией бактерий, описанной в серии из пяти томов, вышедших под

общим наименованием «Руководство Берджи по систематической бактериологии». Данный метод является трудоемким и требует больших временных и трудовых затрат. С целью решения задачи повышения точности измерений разрабатывается система распознавания и подсчета колоний микроорганизмов.

Основная часть.

Целью исследования является разработать автоматизированный метод подсчета колоний микроорганизмов, выделенных из продуктов питания, а также разработать автоматизированную систему для проведения количественного микробиологического анализа», которое поможет пользователю подсчитать колонии микроорганизмов, которые были выращены на чашке Петри. Программное обеспечение для проведения количественного микробиологического анализа будет разрабатываться на языке программирования Python.

Для достижения поставленной цели необходимо решить следующие задачи:

- Изучить существующие аналоги систем для проведения количественного микробиологического анализа;
- Конвертацию цветного изображения в бинарное (изображение из RGB (на языке python это BGR) в черно-белое (или двухцветное) изображение (рис 1 а, б, в, г.)) и обучение нейронной сети для подсчета микроорганизмов на чаши Петри
- Подсчет количества колоний микроорганизмов и сохранение результатов подсчета в Excel.

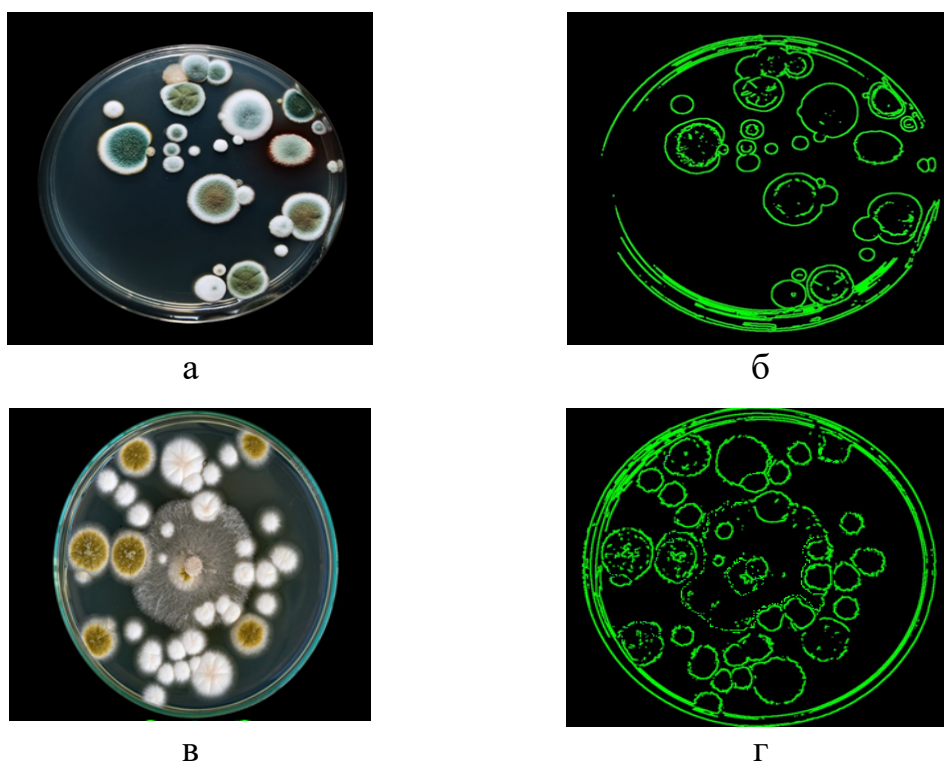


Рисунок 1. Начальное (а, в) и сконвертированное (б, г) изображения из BGR в двухцветное изображение

После того как изображение загрузится в систему, необходимо разбить изображение на набор пикселей. В последующем необходимо произвести форматирование цветного изображения в серое с помощью «grayscale». Для прорисовки контуров использовать «edges». Затем функции «findContours» найдет координаты контуров, а функция «drawContours» создаст найденные контуры на изображении. Далее для разметки данных в файл записываются все

координаты, которые были получены, после преобразования изображения. Во время обучения в файл записываются все координаты от 0 и до 1. Разметка данных представлена на рисунке 2.



Рисунок 2. Разметка данных

После чего необходимо открыть файл с массивом data и в data append вычислить координаты, которые далее будут использоваться для обучения нейросети с учителем.

Далее в «build_labels» вносятся полученные контуры, и делается обработка через нейросеть, чтобы в последствии понять включаемся в метку или нет. Если результат ровняется единице, то – это колония, если ноль – значение распознается как блик. После происходит сравнение и если вероятность больше нуля, то метка добавится в результат. При открытии системы, пользователь видит перед собой главную страницу Web-приложения с полем для добавления изображения.

После загрузки изображения приложение его конвертирует в бинарное для облегчения расчетов микроорганизмов. Также на странице под изображениями появляются кнопки «Рассчитать» и «Очистить». На изображениях добавляются иконки сохранения бинарного изображения и открытия изображений в полноэкранный режим. Загрузка изображения в систему представлено на рисунке 3.



Рисунок 3. Загрузка изображения в систему

После нажатия на кнопку «Рассчитать» в таблице отображается информация о подсчете микроорганизмов. Также в таблице отображается не только подсчет, но и информация о номере изображения и названии. Подсчет колоний микроорганизмов и отображение информации о изображении представлена рисунке 4.

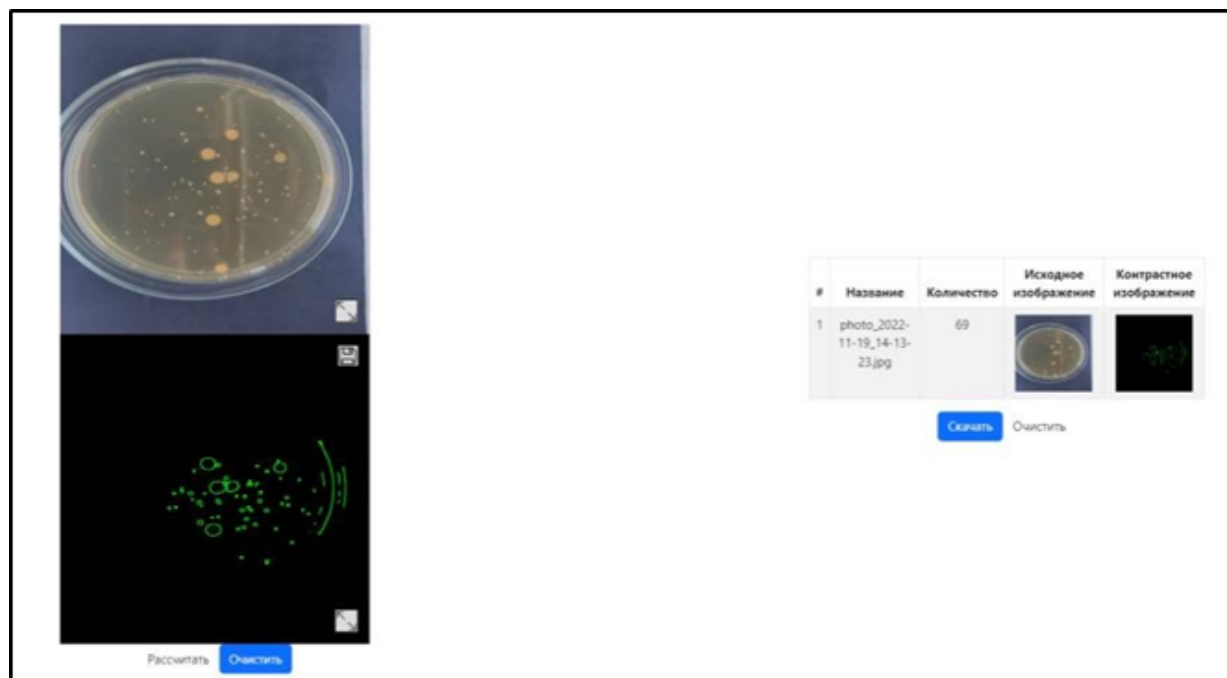


Рисунок 4. Подсчет колоний микроорганизмов и отображение информации о изображении

Сохранение информации из таблицы происходит при нажатии на «Скачать». Пользователь видит скаченный файл в загрузках браузера. При его открытии перед пользователем открывается Excel-файл с информацией о изображении и результат проведенных анализов. Созданная программное обеспечения является Web-приложением, что позволяет быстро перейти к его использованию. А также система требует минимальное количество действий от пользователя. Таким образом разработка и внедрение системы позволяет в значительной мере ускорить подсчет микроорганизмов выращенных на чаше Петри, снизить затраты рабочего времени, а также автоматический подсчет микроорганизмов выросших на чашках Петри, изображенных на снимках, с помощью системы позволит намного точнее и проще проводить исследования количественного микробиологического анализа.

Заключение.

В результате выполнения данной работы становится возможным использование современных технологий для проведения анализа количественного состава микробиоты продуктов питания, что повышает эффективность оценки продовольственной безопасности для населения в целом.

Список литературы

- [1] Хоулт Дж., Криг Н. Определитель бактерий Берджи. В 2-х т, 1997 год – 800 с.
- [2] Язык программирования Python: применение, особенности и перспективы [Электронный ресурс]. –Режим доступа: <https://timeweb.com/ru/community/articles/chto-takoe-python>. – Дата доступа: 08.11.2022
- [3] Нейронные сети на Python: как всё устроено [Электронный ресурс]. – Режим доступа <https://gb.ru/blog/nejronnye-seti-python/>. – Дата доступа: 08.11.2022
- [4] Поиск объекта на изображении с помощью полносвёрточных нейронных сетей [Электронный ресурс]. – Режим доступа <https://www.azoft.ru/blog/poisk-obekta-na-izobrazhenii-s-pomoshchyu-polnosvyortochnyh-nejronnyh-setej/>. – Дата доступа: 08.11.2022

[5] Коркин Л.Р., Коркина М.В., Саевич К.Ф., Ящин К.Д., Подготовка материалов для машинного обучения с помощью виртуальной студии. Седьмая Международная научно-практическая конференция «BIG DATA and Advanced Analytics. BIG DATA и анализ высокого уровня», Минск. – 2021-418с.

MICROORGANISMS COUNTING SYSTEM

L. Korkin

*Assistant of the Department of Engineering Psychology
and Ergonomics of BSUIR, Master of Engineering and
Technology*

K. Saevich

Doctor of Biological Sciences, Professor

Department of Engineering Psychology and Ergonomics

Faculty of Computer Engineering

Belarusian State University of Informatics and Radio Electronics, Republic of Belarus.

E-mail: l.korkin@bsuir.by; konstantin.saevich@mail.ru

Annotation. The method of counting the colonies of microorganisms isolated from food products is used for the quantitative composition of the microbiota of the researched food product. The process is complicated by the fact that in some cases the counting method does not give a 100% result even when it is using a microorganism colony counter. The principle of operation of this counter is that the operator, using an electric pen, puts dots on the bottom of the Petri plate to the location of the colonies. The device registers the fact of touching, and the result is displayed on a digital display. The counter registers only the fact of touching the electric pen, which can't insure the final result from errors of operator. In other cases, the calculation is done mathematically, i.e. counting is carried out in a certain area, usually in one fourth of the Petri plate, and then multiplied by four. There are other methods for determining the number of colonies of microorganisms, but whichever of the existing methods of determination we would choose, each of them has its drawbacks.

Keywords. Machine learning, artificial intelligence, microbiology, colonies of microorganisms.

УДК: 002.6:021

ПЕРСПЕКТИВЫ РАЗВИТИЯ ИТ-АУДИТА СИСТЕМЫ БУХГАЛТЕРСКОГО УЧЁТА ЗА СЧЁТ ВНЕДРЕНИЯ ИКТ



Д.А. Мирзаев
Заведующий кафедры
информационные системы и
цифровые технологии ТФИ,
PhD, доцент
mdilshod@mail.ru



И.Ж. Бозорова
Стажёр-преподаватель
кафедры алгоритмы и
технологии
программирования КарГУ,
i.bozorova667@gmail.com



Ф.Х. Мамасидикова
Студентка магистратуры
кафедры информационные
системы и цифровые
технологии ТФИ,
dizafa@icloud.com

Д.А. Мирзаев

Окончил Ташкентский университет информационных технологий. Область научных интересов связана с разработкой методов и алгоритмов построения информационно-управляющих систем, надёжностью программного обеспечения, мобильными системам связи, организацией учебного и научно-исследовательского процессов в финансовом институте.

И.Ж. Бозорова

Окончила Каршинский филиал Ташкентского университета информационных технологий. Область научных интересов связана с исследованием методов, алгоритмов и технологий программирования, совершенствования системы бухгалтерского учёта в условиях цифровой трансформации, организацией учебного и научно-исследовательского процессов с использованием информационно-коммуникационных технологий.

Ф.Х. Мамасидикова

Студентка магистратуры Ташкентского финансового института. Область научных интересов связана с исследованием экономики и управления в сфере информационно-компьютерных систем.

Аннотация. В статье дается определение ИТ-аудита и рассматривается влияние информационно-коммуникационных технологий (ИКТ) на существующие важные ресурсы в аудите для системы бухгалтерского учета. В качестве основных современных ИКТ рассмотрены технологии на основе искусственного интеллекта, облачного вычисления и Блокчейн технологий. Также в статье рассматривается использование информационных технологий в аудите, даны примеры информационно-коммуникационных средств, применяемых аудиторами. Кроме того, систематизирована задача, решаемая внедрением ИКТ в работу аудиторских компаний, представлен принцип аудита, проводимого с использованием ИКТ.

Ключевые слова: информационные технологии; аудит; бухгалтерский учет; искусственный интеллект; облачные вычисления; Блокчейн; операционный аудит; автоматизация; цифровые технологии

Введение.

В настоящее время информационные технологии играют важную роль в жизни общества, в значительной мере определяя уровень его развития. Они применяются во многих сферах деятельности человека, облегчая при этом выполнение различных задач и операций. Широкое распространение информационные технологии получили и в сфере экономики, в частности в системе бухгалтерского учета и аудита.

Благодаря использованию ИКТ можно увеличить прозрачность деятельности компаний, однако при отсутствии специальных ресурсов по увеличению этой прозрачности применение ИКТ может затруднить возможность идентификации злоупотреблений и защиты от нарушения

прав для лиц, не обладающих необходимой профессиональной квалификацией или лишённых доступа к ИКТ системам поддержки управления компанией.

В связи с распространением использования современных технологий в различных структурах производственных предприятий, в основном в банках, возрастает и незащищённость систем данных. В связи с этим растут риски и угрозы значительных прямых финансовых потерь. В мире наблюдаются быстрый рост количества кибератак в основном на платёжные системы банков. В определении и идентификации характера рисков возникнет потребность в IT-аудите. В деятельности банков прозрачность и надёжность IT-систем взаимосвязаны между собой, потому что информация о важнейших показателях повышения эффективности корпоративного управления, конкурентоспособности, прозрачности бизнеса и инвестиционной привлекательности всегда должна быть в актуальном состоянии. Использование безопасной IT-системы гарантирует надёжность и прозрачность, а также своевременность подготовки бухгалтерской отчётности производственных предприятий, в том числе и банков. Положительные итоги IT-аудита дали возможность проводить финансовый аудит более качественно и оперативно, при этом сэкономить значительные ресурсы производственных предприятий и банков. В компаниях инфраструктура информационных систем является системой организационных структур, обеспечивающих функционирование и развитие информационной среды и средств информационного взаимодействия. IT-аудит – это тщательное изучение, исследование и составление экспертной оценки по деятельности IT-инфраструктуры банков (производственных предприятий) или их подразделений, по соответствию утверждённым стандартам, регламентам и политикам. В основном IT-аудит считается необходимым в случаях:

- внедрения нового программного обеспечения (ПО) и планирования реорганизации в компании;
- для экономии затрат на расходы информационных отделов;
- для выявления «слабого» звена, которое тормозит развитие предприятия и оптимизации бизнес-процессов.

Как правило, основанием для проведения аудита в основных случаях является смена структуры управления или руководителя компаний, которому необходимо получить информацию о текущем состоянии всей структуры. Задача IT-аудита состоит в определении коэффициента полезного действия всей IT-инфраструктуры, в привлечении в порядок информационной системы и устранении возможных технических рисков. В зависимости от деятельности компании (банков и производственных предприятий) и масштаба её IT-инфраструктуры, продолжительность IT-аудита занимает от двух недель до двух месяцев. Для компании любого масштаба IT-аудит считается полезным, даже если её информационная система включает всего несколько компьютерных технологий. После проведения IT-аудита компании получают достоверные отчёты и рекомендации по предотвращению замечаний. Во всех инфраструктурах производственных предприятий и банков IT-аудит используют для развития бизнес-процессов, повышения эффективности затрат и оптимизации работы в целом.

В последнее время в республике Узбекистан проводится политика широкого доступа к информационным ресурсам и большому объёму данных отчётности, что даёт возможность всем бизнес-единицам существенно сократить расходы на понимание деятельности контрагентов и анализ их основных финансовых показателей. Определённые государством приоритетные направления реализуется в политике цифровизации бизнеса. Для осуществления цифровизации необходимо наличие персонала, имеющего соответствующий уровень квалификации в части ИКТ и высокий уровень гражданской ответственности. Навыки использования информационных ресурсов определяют возможности цифровизации для решения текущих и стратегических бизнес-задач компаний.

Возможности, открывшиеся благодаря внедрению современных цифровых ИКТ.

Аудиторская проверка с применением информационных технологий в обязательном порядке включает в себя все процедуры, проводимые при работе аудитора непосредственно в компании. Основными факторами являются установление взаимодействия с лицами, отвечающими за корпоративное управление компаний:

- сотрудники, т.е. персонал, ответственный за корпоративное управление компанией, должен быть доступен в период проведения аудиторской проверки, в том числе и для обсуждения результатов аудита;
- в ходе планирования аудиторской проверки согласовывается подключение участников аудиторской группы к информационной системе аудируемого персонала;
- лица, ответственные за подготовку финансовой, т.е. бухгалтерской отчетности, предоставляют всю необходимую информацию для проверки;
- предварительные документы, подтверждающие осуществленные операции, оперативно предоставляются посредством телекоммуникационных каналов связи;
- Ответы на вопросы, влияющие на достоверность финансовой отчетности, должны предоставляться максимально быстро.

Использование информационных технологий при проведении аудита всегда обеспечивается соответствующим программным обеспечением (например, как IT-аудит, AuditXP, ЭкспрессАудит) [1]. При проведении аудита с использованием информационных технологий в обязательном порядке должны соблюдаться принципы, представленные на рисунке 1. Применение представленных принципов в ходе автоматизации процесса оказания аудиторских услуг позволяет избежать технических ошибок, ускорить процесс коммуникации, повысить качество работы аудитора [2].

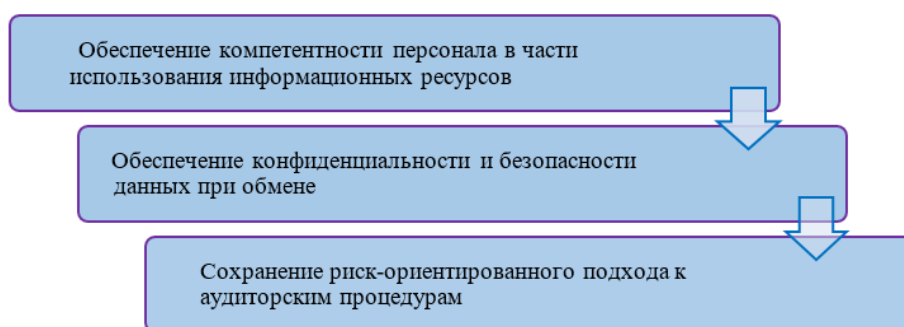


Рисунок 1. Принципы аудита с использованием информационных технологий

Информационные технологии, применяемые при проведении аудита представлены на рисунке 2.

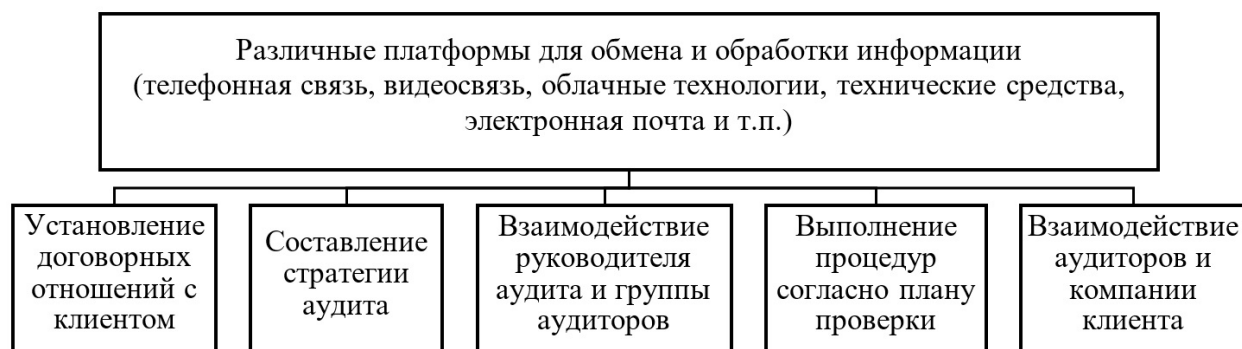


Рисунок 2. Информационные технологии, применяемые при проведении аудита

Ниже рассмотрено влияние использования современных цифровых ИКТ на снижение трудоёмкости и повышение качества аудита. Технологии на основе облачных вычислений, искусственного интеллекта и Блокчейн технологии способны радикально изменить сферу аудита. Технологии, в том числе и, выше перечисленные, приведены на рисунке 3.

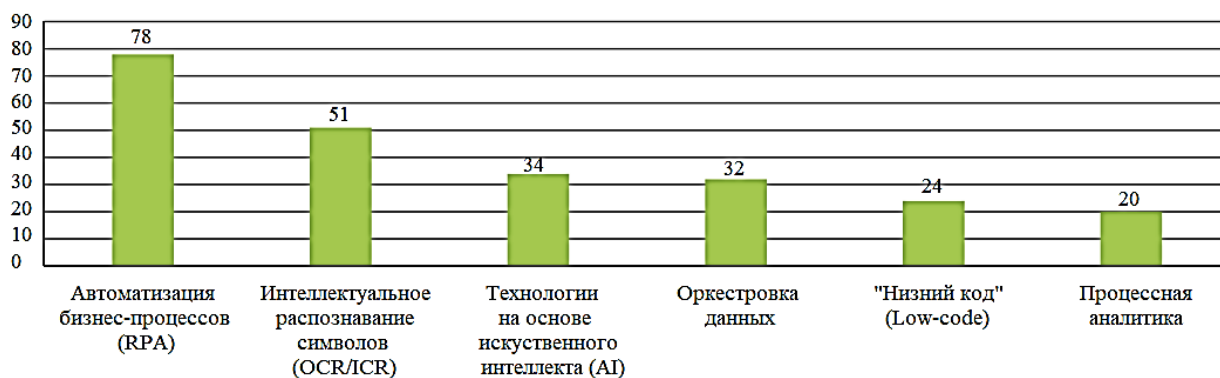


Рисунок 3. Показатели внедрения компаниями современных технологий, %

Технологии на основе облачных вычислений обеспечивают проведение аудита в режиме реального времени и своевременное выявление аномальных транзакций. Облачные вычисления – модель для обеспечения повсеместного, удобного сетевого доступа по требованию к общему ряду вычислительных ресурсов (сетей, серверов, хранилищ, приложений и служб) [3]. Облачные решения внедряются в организации с помощью моделей развертывания – способа доступа к облачным сервисам. Технологии облачного вычисления поддерживают горизонтальное масштабирование, т.е. добавление большего объема данных одного и того же ресурса, и вертикальное масштабирование, которое подразумевает увеличение производительности одного и того же ресурса. Облачные вычисления предоставляют возможности для снижения затрат на инфраструктуру. Масштабирование позволяет наращивать мощность и производительность по мере необходимости. За счёт этого создаётся баланс между снижением затрат на неиспользуемые мощности и альтернативными затратами на недостающие мощности [4]. Блокчейн – это технология хранения данных, которые располагаются в цепочке последовательно связанных блоков. Иными словами, это непрерывный цифровой каталог транзакций, который дублируется и распространяется по системам. Блокчейн технология может стать цифровой бухгалтерской книгой для учёта всех переводов, проводимых между различными контрагентами. Все участники этого процесса, использующие общую базу данных, имеют возможность доступа к идентичным копиям книги. Блокчейн технологии дают возможность обеспечить достоверность и неизменность внесённой информации, стандартизацию финансового учёта и повышение прозрачности [5]. На сегодняшний день Блокчейн считается надёжным, потому что полные копии реестра поддерживаются всеми участниками, то есть даже если один из них потеряет доступ, то реестр по-прежнему остается доступен для всех остальных участников сети. В итоге участники могут изменяться, однако целостность и надёжность сети сохранится. Блокчейн технология способна ускорить многие операции по извлечению данных вручную, которые являются достаточно трудозатратными. С помощью данной технологии аудиторы потенциально могут использовать возможности автоматизированного управления, аналитики и машинного обучения, например, автоматически уведомлять соответствующие стороны о подозрительных транзакциях практически во время их обнаружения [6]. Доступ к аудиторским доказательствам, таким как вспомогательная документация, в частности контракты, соглашения, заказы на поставку и счёта-фактуры, может сократить время проведения аудита. Блокчейн технология содержит историю всех транзакций, которые были произведены и находится в общем доступе.

Данное решение позволяет подтвердить факт совершения транзакции и сохраняет прозрачность проведения операций во всех этапах. Необратимость внесённых изменений каждого блока в цепочке ссылается на предыдущие, что предотвращает удаление или изменение транзакций после их добавления в цепочку. Это действие позволяет проверить достоверность информации и предотвратить повторный учёт доходов или расходов. Данная технология способна значительно упростить проверку наличия транзакций в сфере аудита, поскольку они фиксируются в общую базу блокчейна. Однако это решает исключительно проблему анализа наличия данных транзакций, в то время как аудит также подразумевает оценку того, что зарегистрированные операции подтверждаются соответствующими, надёжными, объективными, точными и поддающимися проверке доказательствами [7].

Технология, основанная на искусственном интеллекте, даёт возможность поиска, извлечения и обработки информации, в том числе проведения интеллектуального анализа данных. Технологии на основе искусственной нейронной сети также позволяют выявлять незаметные на первый взгляд аномалии и формировать прогностическую модель [8].

Технологии на основе искусственного интеллекта, облачных вычислений и Блокчейн технологий являются одними из основных инноваций, способных изменить сферу аудита. Перечисленные технологии, а также возможности, которые они предоставляют, отражены на рисунке 4.



Рисунок 4. Возможности современных цифровых технологий

Заключение.

Аудит подтверждает достоверность финансовой отчётности организации. Данная информация необходима как самой компании, так и её руководству для предотвращения непредвиденных проблем, в том числе попадания к третьим лицам, для уверенности в финансовой устойчивости организаций. В настоящее время организации применяют новейшие и эффективные способы сбора и обработки данных, и компании, осуществляющие аудит, в свою очередь должны быть в состоянии удовлетворить их потребности. В большой мере это означает использование цифровых инструментов при проведении аудита. В этом случае, как было выше отмечено, можно выделить технологии на основе искусственного интеллекта, блокчейна и облачных вычислений.

Существует ряд рисков, которые необходимо учитывать при внедрении цифровых технологий в сфере аудита. В рамках операционных рисков необходимо выделить то, что не полностью протестированные технологии и некачественные данные, а также неудовлетворительный объём выборки приведут в итоге ко множеству ошибок. Непринятие изменений и боязнь сокращения рабочих мест также являются одной из причин негативного отношения к новым цифровым технологиям. Среди рекомендаций по снижению или устранению

рисков при внедрении цифровых технологий можно выделить качественную разработку и многократное тестирование работы технологий, предоставление качественных входных данных и понимание итогового результата внедрения цифровых технологий, обучение сотрудников работе с передовыми технологиям и разработке методов по оценке инвестиционной привлекательности внедрения цифровых технологий. Существующие на рынке практики внедрения цифровых технологий свидетельствуют о возможности эффективной интеграции инноваций в финансовые процессы компаний. Основными преимуществами вышеупомянутых в статье технологий являются гибкость, возможность доступа к данным вне зависимости от местоположения, оперативность управления данными, безопасность обмена информацией и отслеживание действий, связанных с аудитом.

Список литературы

- [1] Кузнецова Е.В. Автоматизация аудита как инструмент контроля дистанционной работы аудиторской организации / Инновационное развитие экономики: тенденции и перспективы. – 2020. – Т. 1. – С. 264-271.
- [2] Изварина Н.Ю. Особенности дистанционного аудита систем менеджмента качества в современных условиях / Н.Ю. Изварина, Р.О. Реутова, Р.А. Бондарев, С.С. Зубарев // Экономика и бизнес: теория и практика. – 2021. – № 10-1 (80). – С. 146-149.
- [3] Khandelwal Y. et al. Pricing strategies of an oligopolist in federated cloud markets. *Cloud Comp.* 2021. –pp. 1–13.
- [4] Нуфарова И.И. Применение облачных технологий в бухгалтерском учёте / Научная дискуссия современной молодёжи: экономика и право. 2016. 175–177 с.
- [5] Пряников М.М., Чугунов А.В. Блокчейн как коммуникационная основа формирования цифровой экономики: преимущества и проблемы / *International Journal of Open Information Technologies.* 2017. 49–55 с.
- [6] Федотова В.В., Емельянов Б.Г., Типнер Л.М. Понятие блокчейн и возможности его использования / *European Science.* 2018. 40–48 с.
- [7] Жидков А.С., Литвинюк А.В. Перспективы использования цифровых технологий в аудиторской деятельности / *Вестник Московского университета им. С.Ю. Витте. Серия 1: Экономика и управление.* 2021. 87–93 с.
- [8] Yusupbekov, N. R., D. A. Mirzaev, and Z. J. Kuziyev. “Analysis of the Current Status of the Theory and Practice of Assessing the Reliability of Software of Automated Information and Control Systems” *Galaxy International Interdisciplinary Research Journal* 10.10 (2022): 418-425.

PROSPECTS FOR THE DEVELOPMENT OF IT-AUDIT OF THE ACCOUNTING SYSTEM THROUGH THE INTRODUCTION OF ICT

D.A. Mirzaev

*Head of the Department of
Information Systems and Digital
Technologies TFI, PhD, Associate
Professor*

I.Zh. Bozorova

*Intern-teacher of the
Department of Algorithms
and Programming
Technologies, KarSU*

F.Kh. Mamasidikova

*Master's student of the Department of
Information Systems and Digital
Technologies TFI*

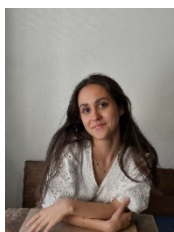
*Department of Information Systems and Digital Technologies
Faculty of taxes and insurance
Tashkent Financial Institute, Republic of Uzbekistan
E-mail: mdilshod@mail.ru*

Abstract. The article defines IT audit and examines the impact of information and communication technologies (ICT) on the existing important resources in the audit for the accounting system. Technologies based on artificial intelligence, cloud computing and Blockchain technologies are considered as the main modern ICT. The article also discusses the use of information technology in the audit, gives examples of information and communication tools used by auditors. In addition, the task solved by the introduction of ICT in the work of audit companies is systematized, the principle of an audit conducted using ICT is presented.

Keywords: information technologies; audit; Accounting; artificial intelligence; cloud computing; Blockchain; operational audit; automation; digital technologies

УДК 004.021:004.75

BIG DATA И ADVANCED ANALYTICS В ТЕЛЕМЕДИЦИНЕ



М.Ю. Чиркова

Студентка инженерно-экономического
факультета БГУИР
marinchi03@gmail.com



О.Н. Шкор

Старший преподаватель кафедры экономики
БГУИР
shkor@bsuir.by

М.Ю. Чиркова

Родилась в 2003 году в Минске. В 2020 году поступила в УО «БГУИР», была зачислена на платную форму обучения по специальности «Электронный маркетинг» инженерно-экономического факультета.

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 данной статье рассматриваются такие понятия как Big Data и Advanced Analytics в сфере телемедицины. Предложен перечень современных методов и тенденций использования технологий больших данных.

Ключевые слова: Big Data, Advanced Analytics, телемедицина, медицина.

Введение

В современных реалиях развитие мира невозможно без информационных технологий. Повсеместное использование компьютеризации привело к тому, что на протяжении многих лет, различные компании и организации постоянно накапливают огромное количество информации. Потребность соответствовать быстрым темпам развития мира привело к созданию такого понятия как Big Data – технологии для управления и извлечения ценности из сложных данных, которые генерируются в больших объемах на высоких скоростях.

Использование Big Data в сфере медицины является новой ступенью развития. Медицинская промышленность постоянно использует большие объемы информации: цифровые медицинские записи, генетические данные, клинические исследования, научные статьи и многое другое. На основе этих данных ученые проводят биомедицинские исследования, разрабатывают новые способы лечения различных заболеваний, а также работают над созданием персонализированной медицины.

Телемедицина

Использование современных технологий и средств телекоммуникаций для дистанционного предоставления врачебных и консультационных услуг подразумевается под термином телемедицина [1].

Некоторые из примеров телемедицинских приложений: Mindspa, iBolit, ICE: In case of emergency и многие другие [2].

Популярность телемедицины растет благодаря улучшению качества обслуживания пациентов, экономии времени и гарантированной безопасности пациентов. Существует

множество способов, которыми Big Data популяризируют телемедицинские услуги, далее представлен некоторый список конкретных примеров:

1. Удаленный мониторинг пациентов – это одна из конкретных форм телемедицины, которая включает в себя использование устройства для сбора и передачи жизненно важных показателей или физиологических данных пациента вне офиса [3]. В определенное время в течение дня пациенты самостоятельно контролируют себя, чтобы получить данные о состоянии своего здоровья, а затем посредством сообщения передают эти данные врачам. Мониторинг здоровья включает в себя использование устройств, чаще всего часов, для предоставления информации: частота сердечных сокращений, кровяное давление, уровень сахара и уровень кислорода в крови. Такой способ контроля здоровья помогает пациентам с заболеваниями сердца высокого риска, сложной беременностью или деменцией. Так, Стэнфордский университет вместе с Apple Watch провели исследование, благодаря которому была обнаружена нерегулярную частота сердечных пульсов у пользователей часов, сигнализирующая о необходимости дальнейшего мониторинга серьезной проблемы с врачами [4].

Скрининг или дистанционный мониторинг здоровья используется для хронических заболеваний пациентов. Так, медицинском университете Вирджинии отслеживал уровень глюкозы у пациентов больных диабетом и предупреждал, если показания были в критической зоне [5]. Дистанционный мониторинг пациентов помогает в оценке опасных для жизни ситуаций и помогает пациентам стать здоровее в долгосрочной перспективе. Кроме того, данные о пациентах в режиме реального времени могут помочь врачам вводить лекарства, изменять дозировки и улучшать здоровье.

2. Трансляция хирургических операций – онлайн наблюдение за операцией, которую проводит высококвалифицированный специалист для начинающих или неопытных хирургов. Во время операции специалисты могут передавать друг другу изображения различного медицинского оборудования (микроскоп, эндоскоп или операционный стол), фиксируемого цифровыми видеочкамерами.

3. Прогнозный анализ позволяет врачам делать «правильный» выбор на основе информации, предоставленной пациентами. Аналитика дает представление о том, какие изменения в организме человека могут произойти и как это можно предотвратить. Например, в случаях, когда лечение лекарством приводит к образованию побочных эффектов, происходит замена лечебных средств. Основываясь на сборе и анализе данных о реакции организма - Big Data, рассчитываются возможные риски и моделируется реакция нового лекарства. Кроме того, оценка проблемы заранее даже после того, как пациент находится вне учреждения по уходу, помогает избежать заболеваний после лечения, таких как сепсис. Например, прогностические аналитические инструменты в Университете Пенсильвании используют данные электронных медицинских карточек и машинного обучения для отслеживания сепсиса за 12 часов до его начала.

4. Облачные электронные медицинские карты - цифровые записи, которые включают в себя информацию пациента о демографии, истории болезни, аллергии и т.д., которые хранятся в облаке. В случае необходимой передачи этих данных другим врачам или учреждениям происходит обмен записей через защищенные информационные системы. Каждая запись представляет собой изменяемый файл, в который врачи могут вносить изменения без бумажной работы и без опасности репликации данных, что позволяет закончить лечение пациента дистанционно.

5. Использование дронов – возможность быстрого сбора информации, транспортировки вакцин. В марте 2021 года изобретатели из Университета Цинциннати совместно разработали новый прототип беспилотника для отправки лекарств или предметов медицинского назначения прямо у порога людей. Дрон оснащен камерами и экраном для обеспечения связи между пациентом и врачом. Дроны оказываются полезными для сбора медицинских данных и обеспечения разведки во время стихийного бедствия. Они могут помочь развернуть

медицинские принадлежности и лекарства в городских и сельских районах, которые в противном случае могут быть недоступны. Дроны могут использоваться для доставки вакцин, транспортировки крови, передачи органов, диагностики и развертывания небольших медицинских устройств [6].

Заключение

Преимуществами больших данных в телемедицине является:

1. Улучшение ухода за пациентами – дистанционное, безопасное.
2. Улучшение исследования - открытие важных медицинских феноменов.
3. Улучшение обучения медиков – онлайн конференции.
4. Снижение затрат для пациентов – заранее выявленные потенциальные проблемы здоровья.
5. Увеличение спасенных жизней – использование дронов.

Таким образом, большие данные и их аналитика в медицине позволяет анализировать большие наборы данных тысяч пациентов, выявлять кластеры и корреляцию между наборами данных, а также разрабатывать прогностические модели с использованием методов интеллектуального анализа данных.

Список литературы

- [1] Телемедицина – современные технологии и пути развития. [Электронный ресурс]. – Электронные данные. – Режим доступа: <https://trueconf.ru/telemedicina.html>
- [2] Приложение для использования телемедицины. [Электронный ресурс]. – Электронные данные. – Режим доступа: <https://medical-club.net/prilozheniya-dlya-ispolzovaniya-telemeditsiny/>
- [3] Телемедицина или удаленный мониторинг. [Электронный ресурс]. – Электронные данные. – Режим доступа: <https://evercare.ru/news/telemedicina-ili-udalennyu-monitoring-chto-luchshe-dlya-pacienta>
- [4] Apple Watch detects irregular heart beat in large U.S. study. [Электронный ресурс]. – Электронные данные. – Режим доступа: <https://www.reuters.com/article/us-health-heart-apple/apple-watch-detects-irregular-heart-beat-in-large-u-s-study-idUSKCN1QX0EI>
- [5] Hypoglycemia Prevention After Exercise in Adolescent T1DM Patients Using a Control to Range System. [Электронный ресурс]. – Электронные данные. – Режим доступа: <https://ichgcp.net/ru/clinical-trials-registry/NCT01390259>
- [6] Преимущества дронов для удаленного здравоохранения и телемедицины. [Электронный ресурс]. – Электронные данные. – Режим доступа: <https://ts2.space/ru/>

BIG DATA AND ADVANCED ANALYTICS IN TELEMEDICINE

M. Y. Chyrkova

*Student of engineering and economics
at the BSUIR*

O.N. Shkor

*Senior Lecturer at the Department of Economics
at the BSUIR*

*Belarusian State University of Informatics and Radioelectronics
Minsk, Republic of Belarus
E-mail: marinchi03@gmail.com*

Abstract. This article discusses such concepts as Big Data and Advanced Analytics in the field of telemedicine. A list of modern methods and trends in the use of big data technologies is proposed.

Keywords: Big Data, Advanced Analytics, telemedicine, medicine.

УДК 621.396

ПРИМЕНЕНИЕ МЕТОДА ИЛЬИНА-МОРОЗОВА И В ЗАДАЧАХ ФОРМИРОВАНИЯ МУЛЬТИСФОКУСИРОВАННОГО ЭЛЕКТРОМАГНИТНОГО ИЗЛУЧЕНИЯ



Ю.Е. Седельников
Профессор кафедры
Радиофоники и
микроволновых технологий
КНИТУ-КАИ, доктор
технических наук, профессор
sedhome2013@yandex.ru



Д.А. Веденькин
Доцент кафедры
Радиофоники и
микроволновых технологий
КНИТУ-КАИ, кандидат
технических наук, доцент
denis_ved@mail.ru



А.Ф. Гильфанова
Аспирант кафедры
Радиофоники и микроволновых
технологий КНИТУ-КАИ
gilfanova-almira@mail.ru

Ю.Е. Седельников

Окончил Казанский авиационный институт. Область научных интересов связана с оценкой электромагнитной совместимости, разработкой и анализом антенн и СВЧ устройств и их технических приложений.

Д.А. Веденькин

Окончил Казанский национальный исследовательский технический университет им. А.Н. Туполева-КАИ. Область научных интересов связана с оценкой электромагнитной совместимости, разработкой и анализом антенн и СВЧ устройств, методами обработки информации и статистической теорией радиотехнических систем.

А.Ф. Гильфанова

Окончила Казанский национальный исследовательский технический университет им. А.Н. Туполева-КАИ. Область научных интересов связана с анализом антенн и СВЧ устройств, методами передачи и обработки информации в телекоммуникационных сетях.

Аннотация. Рассмотрены способы формирования мультисфокусированного излучения в антенных решетках, функционирующих в зоне ближнего излученного поля. Показано, что метод Ильина-Морозова по преобразованию одночастотного излучения в двухчастотное обладает рядом достоинств и является более предпочтительным. Представлены результаты моделирования мультисфокусированной антенной решетки в области ближнего излученного поля. Представленные результаты показывают, что возможность формирования нескольких максимумов поля позволит повысить эффективность использования сфокусированных антенных решеток в радиотехнических задачах.

Ключевые слова: мультифокусировка, сфокусированные антенные решетки, метод Ильина-Морозова, область фокусировки.

Введение.

Начиная со второй половины 40-х годов XX в. и по настоящее время ведутся научные исследования эффекта трехмерной фокусировки электромагнитного излучения в некоторой области пространства. Известна работа [1], в которой впервые были рассмотрены ограничения, накладываемые на сфокусированные в зоне ближнего излученного поля электромагнитные волны. В работе [2] рассмотрены свойства сфокусированных электромагнитных полей. Существенный вклад в развитие теории сфокусированных антенных решеток внес известный советский ученый Я.С. Шифрин, который в одной из своих работ [3] рассматривал сфокусированные поля в зоне ближнего излученного поля, сформированные случайными антенными решетками. В настоящее время активно ведутся исследования в области

сфокусированных антенных решеток и рассматриваются некоторые прикладные аспекты их применения на кафедре Радиопотоники и микроволновых технологий КНИТУ-КАИ под руководством проф. Ю.Е. Седельникова. Особо следует подчеркнуть монографию [4], в которой обобщены результаты исследований антенн и антенных решеток, сфокусированных в зоне ближнего излученного поля, а также их технических приложений.

Однако следует отметить, что несмотря на более чем семидесятилетнюю историю становления и развития теории сфокусированных антенн и антенных решеток вопросы формирования нескольких областей фокусировки одновременно. Рассмотрению данного аспекта и посвящена настоящая работа.

Принципы организации мультисфокусированного излучения.

В современных реалиях наблюдается тенденция к использованию радиотехнических устройств, функциональное назначение которых заключается в формировании распределения электромагнитного поля на расстояниях близких к геометрическим размерам апертуры. К ним можно отнести радиотехнические средства медицинской диагностики, промышленности и сельского хозяйства, радиолокации, телекоммуникационные технологии и многое другое. СВЧ-энергия, ставшая традиционным методом воздействия на материалы и вещества, позволяет расширить круг технологических задач, как, например, разогрев и сушка.

Одним из вариантов реализации мультисфокусированного излучения является способ при котором все излучатели антенной решетки делятся на группы по числу сфокусированных областей, формируя таким образом несколько локальных максимумов напряженностей электромагнитного поля. Иллюстрация подобной идеи представлена на рисунке 1.

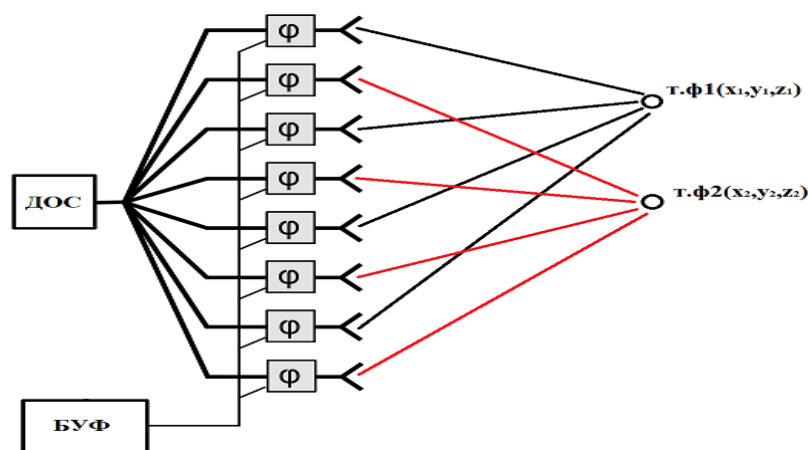


Рисунок 1. Принцип организации мультисфокусированного излучения (ДОС – диаграммообразующая схема, БУФ – блок управления фазами)

При кажущейся простоте подобному решению присущи ряд недостатков. К одному из них можно отнести фактическое разделение единой антенной решетки на ряд отдельных антенных «подрешеток» с меньшим числом излучателей в каждой. Уменьшение числа излучателей неизбежно приведет к снижению эффективности фокусировки и увеличению размеров сфокусированных областей. Иллюстрирующие графики представлены на рисунке 2.

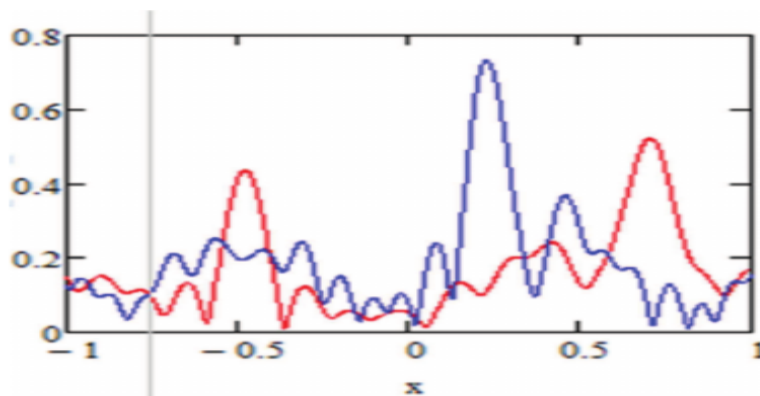


Рисунок 2. Напряженности ЭМ поля при реализации единственной (синяя линия) и двойной (красная линия) фокусировок

Как видно из графиков, представленных на рисунке 2, значения напряженностей электромагнитного поля мультисфокусированного излучения меньше, а поперечные размеры области фокусировки больше, что хорошо согласуется с результатами, отраженными в [4]. Ко второму недостатку относится сложность эффективной и относительно компактной реализации диаграммообразующей схемы особенно при решении задачи адаптивного управления значениями частот, излучаемых отдельными элементарными излучателями. К третьему ограничению можно отнести необходимость отдельной стабилизации значений частот зондирующего излучения, что приведет к усложнению оборудования.

Одним из вариантов преодоления указанных недостатков и ограничений является организация многочастотного излучения с использованием метода Ильина-Морозова [5]. В этом случае каждый из элементов антенной решетки излучает многочастотное излучение, параметры которого подобраны таким образом, чтобы сформировать две независимые друг от друга области фокусировки. Структурная схема подобной антенной решетки представлена на рисунке 3, блок преобразователей частоты по методу Ильина-Морозова находится в фазированном блоке и на рисунке не показан.

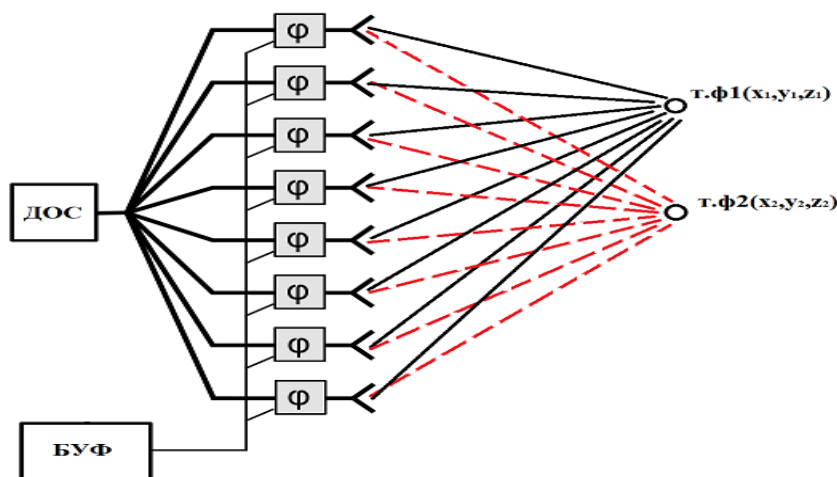


Рисунок 3. Мультисфокусированная антенная решетка с многочастотным излучением

Отметим, что метод Ильина-Морозова к настоящему моменту широко применяется в самых различных областях радиотехники. Несомненными достоинствами применения способа преобразования одночастотного излучения в двухчастотное являются постоянство амплитуд двухчастотного излучения и значения частот, симметричные относительно центральной частоты

с высокой степенью спектральной чистоты. Всё это позволяет значительно повысить эффективность генераторного и диаграммообразующего оборудования, устройств поддержания стабильности частот.

Математическое моделирование мультисфокусированной антенной решетки.

Способ формирования двухчастотного излучения по методу Ильина-Морозова хорошо известен и описан во многих научных трудах, например [6]. Рассмотрим модель линейной сфокусированной антенной решётки со следующими параметрами:

- Количество излучателей – 6 и 11
- Значения частот двухчастотного излучения - 0,95 ГГц и 1,05 ГГц.

Согласно «токовой» модели значения напряженностей электромагнитного поля могут быть оценены по формуле:

$$E(x, y, z) = \sum_{n=1}^{NN} \left[\frac{I_0 * e^{-j*k_1*(R_{\phi i} - R_i(x,y,z))}}{R(x, y, z)} \right],$$

где I_0 – ток, $R_{\phi i}$ – расстояние от i -го излучателя до точки фокусировки, $R_i(x, y, z)$ – расстояние от i -го излучателя до точки наблюдения.

Некоторые результаты моделирования представлены на рисунках 4 и 5.

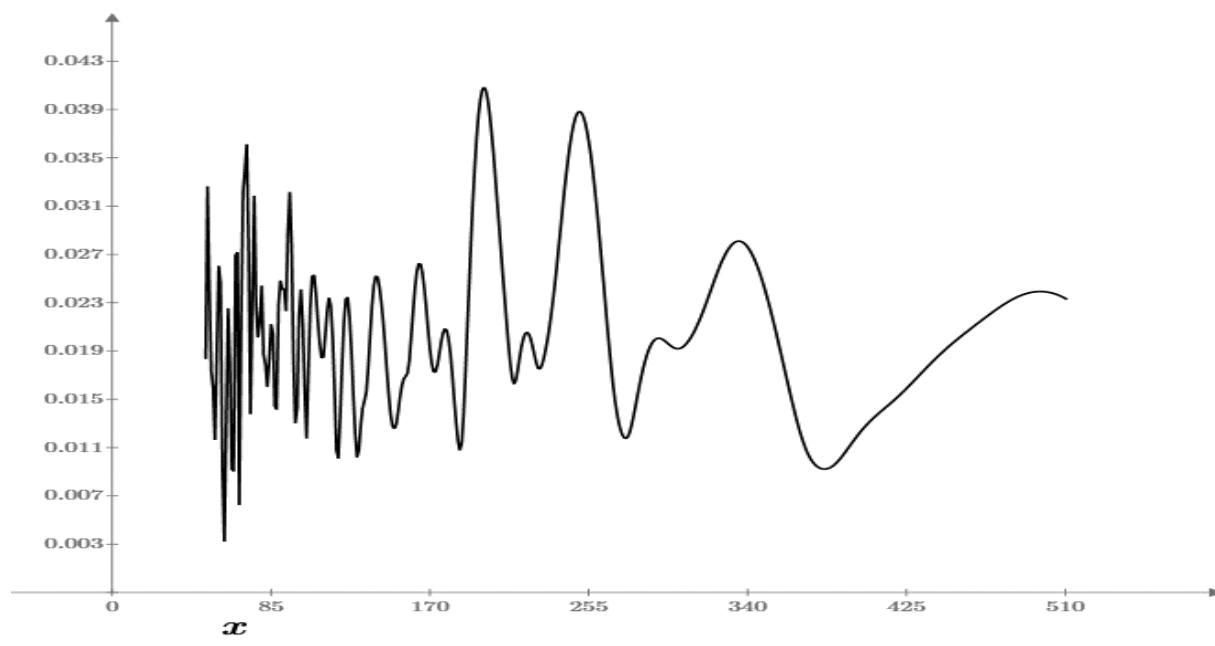


Рисунок 4. Эффект двойной фокусировки многочастотного излучения на дальностях 200 и 250 м. для антенной решетки из 6 излучателей.

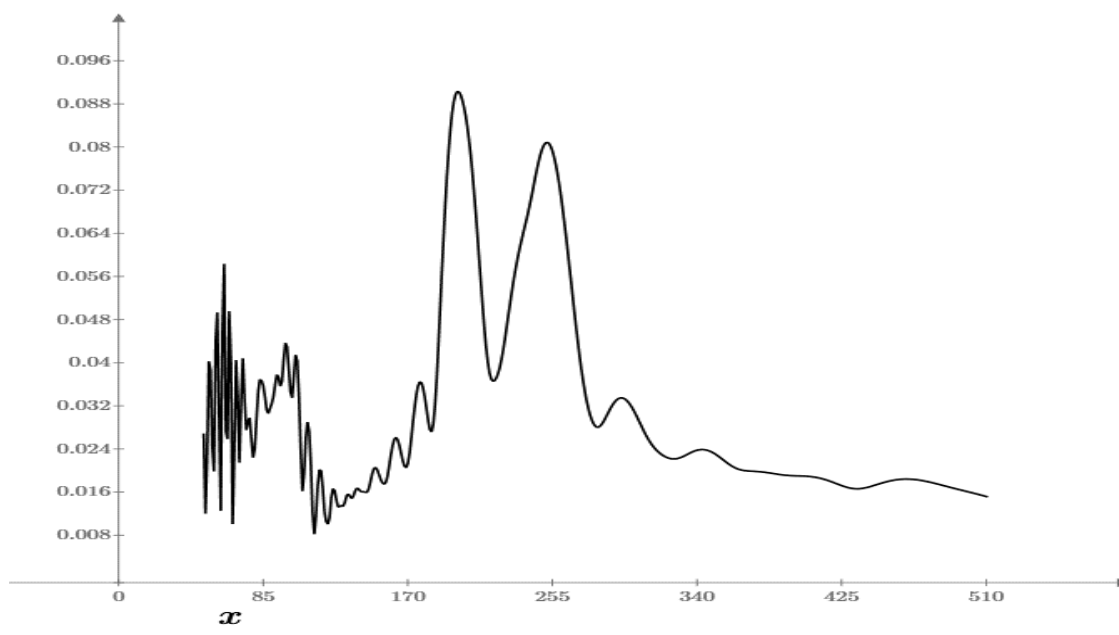


Рисунок 5. Эффект двойной фокусировки многочастотного излучения на дальностях 200 и 250 м. для антенной решетки из 11 излучателей

Как видно из графиков, представленных на рисунках 4 и 5 отдельные области фокусировки имеют четко выраженные обособленные границы, сравнимы по амплитуде и размерам.

Заключение.

Возможность формирования мультисфокусированного излучения открывает перед учеными и специалистами новые возможности по формированию пространственно-распределенных электромагнитных полей с адаптивным управлением числа, местоположения и размеров сфокусированных областей. Использование углубленной аналитики в задачах формирования амплитудно-фазовых распределений возбуждающих токов также позволит повысить эффективность сфокусированных мультисфокусированных апертур.

Список литературы

- [1] R.S. Wehner, Limitations of Focused Aperture Antennas, October 1949.
- [2] W. Sherman, Properties of Focused Apertures in the Fresnel Region, IEEE Trans. Antennas and Propagation, 1962. Т. 10, - №4. – С.399-408.
- [3] Шифрин Я.С. Поле случайных антенных решеток в зоне Френеля / Я.С. Шифрин, В.А. Назаренко // Радиотехника и электроника, 1991. Т.31. С.52-62
- [4] Антенны, сфокусированные в зоне ближнего излученного поля : монография / под общ. ред. Ю.Е. Седельникова, Н.А. Тестоедова ; Сиб. гос. аэрокосмич. ун-т. -Красноярск, 2015. –308 с.
- [5] Ильин Г.И., Морозов О.Г. Патент №: SU 1338647. «Способ преобразования одночастотного когерентного излучения в двухчастотное».
- [6] Нургазизов М.Р. Оптико-электронные системы измерения мгновенной частоты радиосигналов СВЧ-диапазона на основе амплитудно-фазового модуляционного преобразования оптической несущей /дисс. на соискание уч. степени канд. техн. наук // КНИТУ-КАИ, Казань, 2014 г.

APPLICATION OF THE ILYIN-MOROZOV METHOD IN PROBLEMS OF FORMATION OF MULTISFOCUSED ELECTROMAGNETIC RADIATION

Y.E. Sedelnikov

Professor of Radiophotonics and Microwave Technologies Department of KNRTU-KAI, Doctor of Technical Sciences, Professor

D.A. Vedenkin

Associate Professor of Radiophotonics and Microwave Technologies Department KNRTU-KAI, PhD, Associate Professor

A.F. Gilfanova

Postgraduate student of Radiophotonics and Microwave Technologies Department of KNRTU-KAI

*Department of Radiophotonics and Microwave Technologies
Institute of Radioelectronics, Photonics and Digital Technologies
Kazan National Research Technical University n.a. A.N.Tupolev-KAI, Russia
E-mail: denis_ved@mail.ru*

Abstract. Methods for the formation of multifocused radiation in antenna arrays operating in the zone of the near radiated field are considered. It is shown that the Ilyin-Morozov method for converting single-frequency radiation into two-frequency radiation has a number of advantages and is more preferable. The results of modeling a multifocused antenna array in the region of the near radiated field are presented. The presented results show that the possibility of forming several field maxima will improve the efficiency of using focused antenna arrays in radio engineering problems.

Keywords: multifocusing, focused antenna arrays, Ilyin-Morozov method, focus area.

УДК 004.8+78.02

ИСПОЛЬЗОВАНИЕ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА ДЛЯ ГЕНЕРАЦИИ МУЗЫКИ



Н.И. Потапенко

Старший преподаватель кафедры инженерной психологии и эргономики БГУИР



К.Ю. Назарук

Студент кафедры инженерной психологии и эргономики БГУИР



А.Н. Василькова

Ассистент кафедры инженерной психологии и эргономики БГУИР, магистр
a.vasilkova@bsuir.by

Н.И. Потапенко

Старший преподаватель кафедры инженерной психологии и эргономики БГУИР

К.Ю. Назарук

Студент кафедры инженерной психологии и эргономики БГУИР

А.Н. Василькова

Ассистент кафедры инженерной психологии и эргономики БГУИР, магистр

Аннотация. Музыка является неотъемлемой частью культуры, способом творческого самовыражения человека. Однако, помимо творческой составляющей, музыка также содержит строгие правила, шаблоны, последовательности нот, параметры высоты тона, скорости, темпа и так далее. Подобная информация при наличии большого количества аудиозаписей представляет из себя большие данные (Big Data), которые могут быть использованы для анализа существующей, а также генерации новой музыки при помощи искусственного интеллекта и нейронных сетей. Для синтеза музыки необходимо привести аудиофайлы в удобный для анализа MIDI-формат с помощью применения современных алгоритмов, также необходимо использовать алгоритмы генерации последовательностей и непосредственно звука для создания новых мелодий. Несмотря на инновационность подобных технологий, существуют различные проблемы, в том числе и этического характера, связанные с внедрением искусственного интеллекта и Big Data как в область музыкального, так и любого другого творчества.

Ключевые слова: генерация, Big Data, искусственный интеллект, нейронные сети, наборы данных, алгоритмы, MIDI, MAESTRO, Wave2Midi2Wave, Transformer, RNN, LSTM.

Введение.

Музыка – вид искусства, который на протяжении ни одной тысячи лет является частью культуры. С развитием технологий создание и потребление музыки претерпели значительные изменения. Одним из значительных стало использование больших данных (*Big Data*) и искусственного интеллекта (ИИ) для генерации музыки. *Big Data* и ИИ способны произвести революцию в создании музыки: от генерации простых мелодий и гармоний до аранжировок полноценных композиций.

Применение *Big Data* и ИИ в генерации музыки уже показало многообещающие результаты и есть немало примеров музыки, созданной при помощи ИИ, которые привлекли внимание за последние несколько лет. Однако предстоит ещё немало работы, прежде чем музыка, созданная при помощи ИИ, достигнет уровня качества и художественной выразительности музыки, созданной человеком.

Тем не менее технологии генерации музыки с использованием ИИ и *Big Data*, различные техники и модели, а также такие этические аспекты создания музыки при помощи ИИ, как влияние на музыкальную индустрию и роль творчества в создании музыки заслуживают внимания.

Основная часть.

Помимо очевидной творческой составляющей, музыкальные композиции подчиняются строгим правилам теории музыки, которые используют как для анализа, так и для написания музыки. Она содержит множество закономерностей, шаблонов и последовательностей нот, что можно структурировать, проанализировать, сформировать определенный набор данных, который можно использовать как источник для нейронных сетей, чтобы создать на их основе новые композиции.

Наиболее распространенный тип нейронных сетей плохо справляется с последовательными или временными данными, требует фиксированные размеры входных данных [1].

Рекуррентные нейронные сети решают эту проблему, за счет того, что последующие итерации передают данные от последней. Это означает, что информация передается через сеть каждый раз [1].

Принимая выходные данные одного прямого прохода и передавая их в следующий, можно генерировать совершенно новые последовательности данных. Это называется выборкой (*sampling*) [1]. Рекуррентные нейронные сети (*RNN*) имеют некоторые проблемы, такие как затухание или взрыв градиентов, которые возникают, когда сеть слишком глубока. Это решается с помощью сети долгой краткосрочной памяти (*Long Short-Term Memory* или *LSTM*), которая создает короткие пути в сети [1].

RNN достаточно хорошо справляется с задачей генерации музыки, однако для ее функционирования необходим определенный набор данных, такой как, например, *MAESTRO* (*MIDI and Audio Edited for Synchronous Tracks and Organization*), а также архитектуру *Wave2Midi2Wave*, которая сочетает в себе три современных алгоритма, которые обучает на наборе данных *MAESTRO*.

Набор данных *MAESTRO* содержит 172 часа аудио и *MIDI* транскрипций, что гораздо больше аналогов. Например, набор данных *MAPS* содержит только 17,9 часов аудио, *MusicNet* – всего 15,3 часа [2].

MIDI (*Musical Instrument Digital Interface*) – это технический стандарт, который включает в себя множество компьютерных протоколов для взаимодействия с различными типами аудиоустройств. Передаваемые данные содержат информацию о нотах, высоте, скорости и темпе [2].

Wave2Midi2Wave представляет собой комбинацию трех различных современных моделей, каждая из которых выполняет свою задачу. *Wave2Midi* используется для транскрибирования аудио в символическое представление (*MIDI*), что представлено на рисунке 1. Затем *Midi*-часть сети генерирует новый контент.

Все это синтезируется с помощью *Midi2Wave*, чтобы получить реалистично звучащую музыку [2].

Первая сеть в *Wave2Midi2Wave* использует современную архитектуру под названием *Onsets and Frames*, которая автоматически превращает аудиозаписи в заметки, представленные в *MIDI* – формате [2].

Для второй сети в *Wave2Midi2Wave* используется особый тип Трансформера для генерации совершенно новых последовательностей музыки с долгосрочной когерентностью. Выход этой сети гораздо более структурирован по сравнению с другими нейронными сетями [3].

В обычном Трансформере механизм *Self-Attention* используется для моделирования отношений между словами, потому что в предложениях значение слова основано не только на словах, которые были перед ним, но и на контексте всего предложения. Трансформеры агрегируют информацию из всех других частей сети и генерируют представление для каждого слова на основе полного контекста. Этот процесс повторяется для каждого слова для создания новых представлений [3].

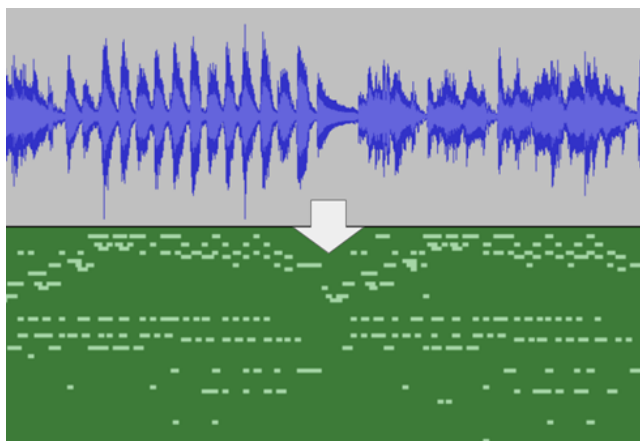


Рисунок 1. Транскрипция звука из аудиофайла в MIDI-представление

С помощью Трансформеров можно приписывать информацию различным фрагментам данных на основе контекста всей сети.

Одна из проблем со стандартным Трансформатором заключается в том, что он полагается на абсолютные позиции для *Self-Attention*. Применительно к музыке, Трансформеры борются с расстояниями, порядком и повторением. Используя вместо этого относительное внимание, музыкальная модель Трансформера может сосредоточиться на реляционных особенностях и генерировать последовательности, выходящие за рамки того, что было дано в примерах для его обучения [3].

Последняя часть сети использует модель *WaveNet* и обучает ее на наборе данных генерировать музыку, которая буквально звучит как запись. *WaveNet* – это модельная архитектура, которая основана на *PixelCNN* и специализируется на синтезе звука [4].

Ее архитектура использует сверточные слои (*convolutional layers*), представленные на рисунке 2.

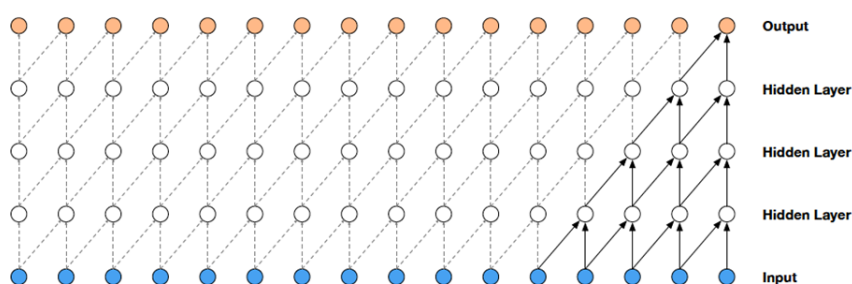


Рисунок 2. Диаграмма стопки сверточных слоев

Поскольку свертки не используют повторяющиеся соединения, подобные тем, что используются в *RNN*, это означает, что их обычно намного легче обучить, чем *RNN*. Но одна из проблем заключается в том, что для увеличения области восприятия (объема данных, которые может покрыть модель) требуется тонна слоев или сверхбольших фильтров, что увеличивает вычислительные затраты [4].

Чтобы обойти это, используются расширенные свертки (*dilated convolutions*), представленные на рисунке 3.

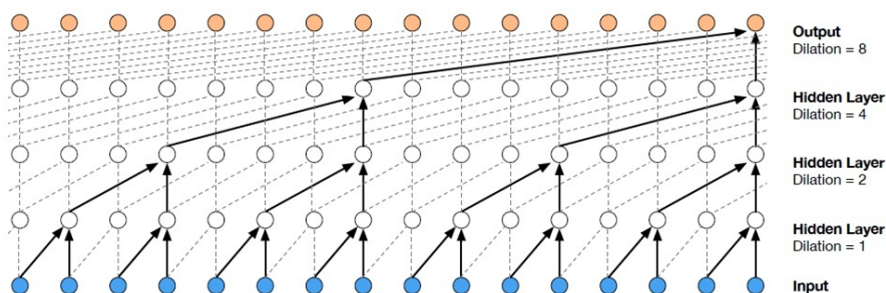


Рисунок 3. Диаграмма стопки расширенных сверточных слоев

Это означает, что фильтры могут быть применены на большей площади, если определенные входные значения пропущены. Получается почти тот же эффект, что и при использовании более крупного фильтра, если расширить его нулями, но расширенные свертки намного эффективнее. Обучение *WaveNet*, современной модели синтеза речи, на наборе данных *MAESTRO* дает довольно впечатляющие результаты [4]. Глядя на фактическую структуру песен, становится еще более ясно, насколько эффективен музыкальный Трансформер в создании новых произведений, которые имеют структурный смысл, что прослеживается на рисунке 4.

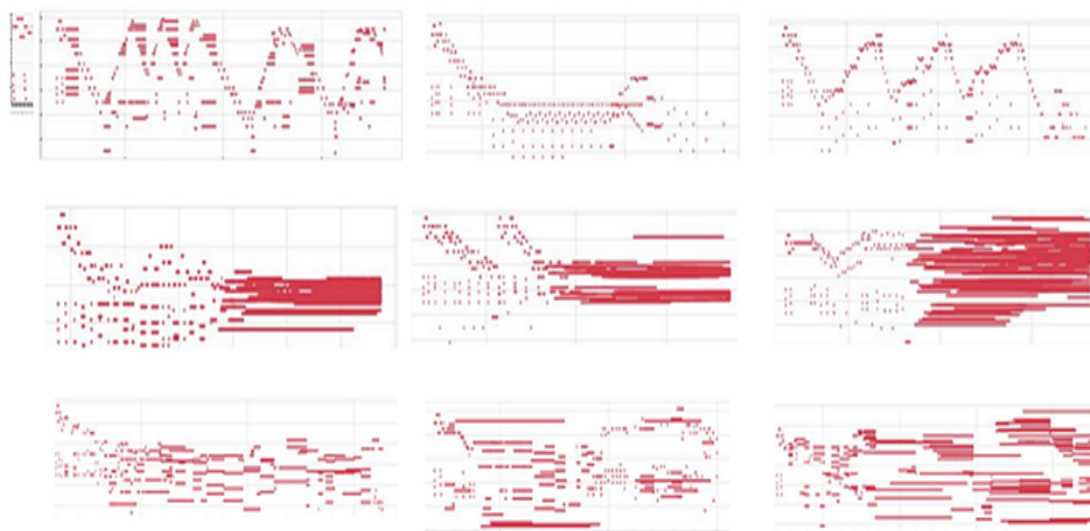


Рисунок 4. Сравнительные диаграммы музыкального Трансформера (сверху), Трансформера (в середине) и LSTM (снизу)

Верхний ряд четко указывает на долгосрочную структуру и повторяющиеся узоры в музыке. Второй ряд начинается первоначально с когерентных паттернов, но затем через несколько секунд становится нагромождением нот и аккордов без видимой структуры.

Последняя строка показывает еще меньше повторений, структуры и согласованности.

Заключение.

Использование искусственного интеллекта и *Big Data* при создании музыки является быстро развивающейся областью, которая имеет потенциал изменить способ создания и потребления музыки. С развитием алгоритмов машинного обучения и нейронных сетей, музыка, созданная искусственным интеллектом, может быть неразличима от музыки, созданной человеком. Однако, все еще существуют этические и творческие соображения, такие как

потенциальная возможность замены музыкантов и необходимость в том, чтобы музыка, созданная искусственным интеллектом, была действительно инновационной, а не просто повторением существующей музыки. Для генерации музыки при помощи искусственного интеллекта необходим значительный набор данных из аудиозаписей и *MIDI*-файлов, чтобы получить информацию о нотах, высоте тона, скорости и темпе. Также необходимо использовать рекуррентные нейронные сети как самые эффективные для выполнения таких целей и модели, применяющие современные алгоритмы транскрипции, генерации и синтеза реалистично звучащей музыки.

Список литературы

[1] Elham Rastegar-Mojarad. Recurrent Neural Networks for Short-Term Load Forecasting / Elham Rastegar-Mojarad, Ali Abbaszadeh. – Berlin: Springer, 2019. – 107 с.

[2] Tao Li. Music Data Analysis: Foundations and Applications / Tao Li, Mitsunori Ogihara, George Tzanetakis – Boca Raton: CRC Press, 2018. – 396 с.

[3] Henrique Malvar. Data Science for Musicians: Foundations, Techniques, and Applications / Henrique Malvar, Benoit Meudic, Emmanouil Benetos. – Berlin: Springer, 2020. – 253 с.

USING ARTIFICIAL INTELLIGENCE FOR MUSIC GENERATION

N.I. Potapenko

*Senior Lecturer, Department of
Engineering Psychology and
Ergonomics*

K.Y. Nazaruk

*Student of the Department of
Engineering Psychology and
Ergonomics.*

A.N. Vasilkova

*Assistant of the Department of
Engineering Psychology and
Ergonomics, master*

*Belarusian State University of Informatics and Radioelectronics, Minsk, Republic of Belarus
E-mail: a.vasilkova@bsuir.by*

Annotation. Music is an integral part of culture and a way for people to express themselves creatively. However, in addition to its creative aspect, music also contains strict rules, patterns, note sequences, and parameters such as pitch height, speed, tempo, and more. With a large number of audio recordings available, such information represents big data that can be used for analyzing existing music and generating new music using artificial intelligence and neural networks. To synthesize music, it is necessary to convert audio files into a convenient MIDI format for analysis using modern algorithms. Additionally, it is necessary to use algorithms for generating sequences and sound to create new melodies. Despite the innovation of such technologies, there are various issues, including ethical concerns, associated with the implementation of artificial intelligence and big data in the field of music and any other creative field.

Keywords: generation, Big Data, artificial intelligence, neural networks, datasets, algorithms, MIDI, MAESTRO, Wave2Midi2Wave, Transformer, RNN, LSTM.

УДК 004.62

ПРИМЕНЕНИЕ БОЛЬШИХ ДАННЫХ В РАБОТЕ МОРСКИХ ПОРТОВ И ТЕРМИНАЛОВ



Ю.И. Голубович

Студент 4 курса специальности
ИиТП, КСиС, БГУИР
gyuliya2001@gmail.com



С.Н. Нестеренков

Декан факультета
компьютерных систем и
сетей БГУИР, кандидат
технических наук, доцент
s.nesterenkov@bsuir.by



Д.В. Басак

Ассистент кафедры
электронных вычислительных
машин БГУИР, инженер-
программист ОСТ ЦИИР
d.basak@bsuir.by

Ю.И. Голубович

Студентка 4 курса специальности «Информатика и технологии программирования» факультета компьютерных систем и сетей БГУИР.

С.Н. Нестеренков

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

Д.В. Басак

Окончил БГУИР в 2021 году по специальности "Вычислительные машины, системы и сети", магистрант первого года обучения по специальности "Компьютерная инженерия" БГУИР.

Аннотация. В настоящее время сбор, хранение, анализ и передача большого количества данных – неотъемлемая часть портовой индустрии. Традиционные методы работы (ручные и полуавтоматизированные) постепенно перестают соответствовать темпам развития и объемам индустрии, поэтому с целью повышения эффективности, производительности и конкурентоспособности они дополняются возможностями новейших технологий. Морские порты и терминалы являются неотъемлемыми участниками постоянно растущей и развивающейся мировой экономики, поэтому использование технологий больших данных если не сейчас, то в обозримом будущем станет обязательным условием их функционирования. В данном докладе представлен обзор использования больших данных в работе морских портов и терминалов: описаны сферы применения и основные преимущества.

Ключевые слова: морские порты, терминалы, TOS, большие данные.

Введение.

Портовая индустрия – это сложная система, состоящая из множества людей, стран и организаций (таких как судовладельцы, власти, трейдеры, компании и предприятия, поставщики складских услуг, кранов, причалов и др.). Судходная отрасль весьма традиционна, поэтому внедрение новых технологий в ней происходит довольно медленного. Благодаря традиционным технологиям логистические компании получают такие данные, как время, скорость, количество и цены, и используют их в анализе, отчетности и визуализации для лучшего представления и оценки бизнес-моделей.

В современном мире наблюдается быстрый рост сложности, трафика и конкуренции в портовом, логистическом и морском секторах, что заставляет компании постоянно инвестировать в повышение эффективности и производительности, и снижение общих затрат, следовательно, растет спрос на высокотехнологические решения.

Большие данные в работе морских портов и терминалов.

Эффективное сочетание информационных технологий, статистики, распределенных систем и математических подходов – это жизненно важные в настоящее время методы для любой компании, которая хочет оставаться конкурентоспособной, анализировать общедоступные данные и преобразовывать их в ценную информацию для отдельных людей, общества или организаций, что может быть самым большим преимуществом для современных компаний [1].

Для регулирования работы терминалов в морских портах используется TOS. Операционная система терминала (TOS) – это система, предназначенная для помощи менеджерам в планировании, составлении графиков и контроле функционирования оборудования контейнерного терминала. TOS является сердцем терминальных операций, поэтому ее надежность и способность обеспечивать высокую производительность операций имеют первостепенное значение. Даже кратковременные сбои могут нанести существенный финансовый ущерб [2].

Чтобы извлечь полезную информацию из базы данных, составляются ежедневные отчеты до процедуры предварительной обработки. Исходные данные в Microsoft Excel генерируются из TOS, в которой каждый лист представляет один операционный день, а около 20 столбцов представляют каждую операцию, выполняемую оборудованием при перемещении контейнеров в пределах терминала. В них описывается время от момента получения команды на перемещение контейнера до завершения оператором требуемой операции. Поскольку каждая строка в базе данных представляет собой одно перемещение, отслеживание контейнера в терминале – непростой процесс и для получения такой информации требуется интенсивная обработка. В таблице описываются характеристики контейнера, такие как тип, размер, высота и IMO (идентификационный номер контейнера), а также номер T/C (Transfer Trane) или S/C (Straddle Carrier), GC (Gantry Crane), C/Y (Container Yard), внешнее шасси грузового автомобиля, TP (Transfer Point Transfer Point), компании и операторы, ответственные за выполнение заказа [3].

Использование больших данных можно определить как использование новых технологий для чтения и обработки огромных объемов различных типов данных (метрических, текстовых, аудио, видео и т.д.), которые позволяют действовать для достижения определенных целей [5].

С появлением методов управления большими данными в портовой индустрии стало возможным отслеживание таких факторов, как погодные условия, незапланированные ремонтные работы, задержки перевозок и транспортной логистик, а также большой объем данных, полученных от датчиков, Global Positioning System (GPS), Radio Frequency Identification (RFID) и других систем управления движением.

Однако каждый порт имеет свои особенности и потребности, к которым должна адаптироваться система, использующая большие данные. Например, для крупного логистического узла, который ориентируется на производительность, и для городского порта, которому необходимо уделять внимание на утилизацию отходов и другие экологические вопросы, понадобятся различные подходы.

Большие данные в анализе портовых данных делятся на три основные группы: управление судами, управление портами и грузами с использованием данных портовой администрации, анализ пространственных изображений по данным систем позиционного слежения [6].

При управлении судами используются данные, доступные в различных регистрах, параметры системы и статистика бункеровки, что позволяет качественнее обслуживать транспортные средства и лучше управлять экипажем.

Управление портами и грузами с использованием данных администрации порта относится к перевозчикам и торговым компаниям; включает эффективную обработку грузов, отслеживание товаров, оптимизацию портовых сооружений и др.

Анализ пространственных изображений происходит с использованием таких средств, как AIS (система автоматической идентификации) и LRIT (система дальней идентификации и

слежения за судами), изображения судов, береговых и космических радаров, оптических датчиков, которые обеспечивают эффективную маршрутизацию, слежение за флотом, анализ трафика, обнаружение аномалий и т.д.

Большие данные используются в работе морских портов при прогнозировании погодных условий. С помощью таких инструментов, как датчики IoT, расширенный интеллект и интеллектуальные данные о погоде, порт предоставляет точные данные о воде и погоде, что позволяет установить лучшее время для захода судов в порт [4].

Использование больших данных в морских портах позволяет анализировать информацию в режиме реального времени. Постоянный сбор и хранение данных, поступающих от технологических компаний и клиентов-перевозчиков, и передача данных различным заинтересованным сторонам порта по всему миру, позволяет экономить время и деньги и повышает производительность работы порта. Например, DataPorts (международный консорциум, софинансируемый Европейской комиссией), опирается на такие технологии, как большие данные и блокчейн, для формирования рынка данных, на который могут предоставлять информацию все заинтересованные стороны в цепочке портов. Простой, быстрый и безопасный доступ к этому рынку позволит осуществлять глобальный мониторинг товаров по всей цепочке поставок [5].

В современных морских портах для экономии времени при инспекции судов и порта используются специальные автономные аппараты, которые ведут подводную съемку и предоставляют информацию о судне и состоянии порта. Подводные дроны помогают увеличить скорость проверки, при этом анализируется большое количество данных.

Хранение данных об обработке, местоположении и передаче грузов, доступности и эффективности работы погрузочно-разгрузочных кранов и машин, и другого необходимого оборудования позволяет избегать задержек на любом этапе работы портовых терминалов.

Большие данные также могут быть полезны в предсказании движения пришвартованных кораблей. Siport21 вместе с Instituto Tecnológico de Informática (ITI) разработали алгоритм с использованием 3D-моделей, который прогнозирует шесть степеней свободы движения корабля во внешнем порту La Coruña, Punta Langosteira, Испания [5].

Заключение.

При разработке и внедрении новых алгоритмов, основанных на технологии больших данных, данные из различных источников собираются и интегрируются, что позволяет достичь следующих преимуществ:

- возможность разрабатывать и внедрять глобальное командное управление производительностью логистики и портовых операций, выявлять источники неэффективности и повышать эффективность принятия решений;

- возможность разрабатывать и реализовывать алгоритмы оптимизации системы управления движением, которые позволяют планировать работу кранов и оборудования, повышая общую производительность и производительность терминалов;

- возможность адаптировать и внедрить системы профилактического обслуживания, основанные на измерении, отслеживании и мониторинге датчиков оборудования, что позволит достичь минимизации непредвиденных простоев и поломок.

Создание умных портов помогает работать эффективнее и сокращать расходы. Объем данных, генерируемых интеллектуальными датчиками, – это бизнес-возможность для компаний, которые разрабатывают новые услуги на основе данных.

Список литературы

[1] Novaes Mathias T Shinoda T Hangga P Inutsuka H 2019 Big Data Approach to Identify the Waste Management of Container Terminal Resources Asian Transport Studies 5 (4) 653–678.

[2] Boer C A and Saanen Y A 2012 Improving container terminal efficiency through emulation Journal of Simulation 6, 267–278. doi:10.1057/jos.2012.10

[3] Tiago Novaes and Mathias Takeshi Shinoda 2022 Concepts of Big Data Analysis of Container Terminals in the Digital Era

[4] Хабибов, С.Х. Большие данные в наблюдении за океаном: возможности и проблемы / С.Х. Хабибов, А.Н. Марков, С.Н. Нестеренков // BIG DATA and Advanced Analytics = BIG DATA и анализ высокого уровня : сб. материалов VIII Междунар. науч.-практ. конф. (Республика Беларусь, Минск, 11-12 мая 2022 года). Редкол.: В. А. Богуш [и др.]. – Минск : Бестпринт, 2022. - С. 265-272.

[5] The use of Big Data at ports [Электронный ресурс]. – Режим доступа: <https://prosertek.com/blog/the-use-of-big-data-at-ports/>. – Дата доступа: 27.03.2023.

[6] Big Data Applications in the Port and Maritime Industry [Электронный ресурс]. – Режим доступа: <https://predikdata.com/big-data-applied-to-the-port-sector/>. – Дата доступа: 27.03.2023.

[7] Сидоркевич, Я.О. Использование Big Data в логистике / Я.О. Сидоркевич, С.Н. Нестеренков // BIG DATA and Advanced Analytics = BIG DATA и анализ высокого уровня : сб. материалов VIII Междунар. науч.-практ. конф. (Республика Беларусь, Минск, 11-12 мая 2022 года). Редкол.: В. А. Богуш [и др.]. – Минск : Бестпринт, 2022. - С. 219-222.

APPLICATION OF BIG DATA IN THE OPERATION OF SEA PORTS AND TERMINALS

Y.I. Golubovich

*Pregraduate student of the
BSUIR*

S.N. Nesterenkov

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

D.V. Basak

*Assistant of the Department of
Electronic Computing Machines of
BSUIR, Programming Engineer ITD
CIID*

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

E-mail: gyuliya2001@gmail.com, s.nesterenkov@bsuir.by, d.basak@bsuir.by

Abstract. Nowadays, collecting, storing, analyzing and transmitting large amounts of data is an integral part of the port industry. Traditional methods of work (manual and semi-automated) are gradually getting out of step with the pace and volume of industry development, so they are supplemented with the capabilities of the latest technologies in order to increase efficiency, productivity and competitiveness. Seaports and terminals are indispensable players in an ever-growing and evolving global economy, so the use of big data technologies will become a prerequisite for their operation in the foreseeable future. This report provides an overview of the use of big data in the work of seaports and terminals: it describes the areas of application and the main advantages.

Keywords: seaports, terminals, TOS, big data.

УДК 004.896

СИСТЕМА ИДЕНТИФИКАЦИИ РУКОПИСНЫХ ТЕКСТОВ КАК ОБЪЕКТ ЗАЩИТЫ И ОБРАБОТКИ МАССОВОГО ОБЪЕМА ДОКУМЕНТОВ



А.И. Калько

*Старший преподаватель кафедры ИТиФМД БарГУ, магистр информатики и вычислительной техники
lexa170594@gmail.com*

А.И. Калько

Окончил Барановичский государственный университет г.Барановичи. Область научных интересов связана с разработкой методов и алгоритмов построения информационно-компьютерных систем, нейросетевых архитектур, распознавание образов.

Аннотация. Выполнен анализ методов и алгоритмов идентификации рукописного текста. Описаны свойства фрагментов изображений рукописного текста, признаки классификации в подобластях цифровых изображений рукописного текста, признаки классификации для определения символов. Представлена разработанная система идентификации, как итог исследования.

Ключевые слова: Идентификация, нейронные сети, большие данные, признаки классификации, рукописный текст.

Введение.

Одной из основных причин, по которой следует разработать систему идентификации рукописного текста, является необходимость ускорения процесса обработки рукописных документов. Рукописные документы, такие как дневники, записные книжки, документы старых архивов и т.д., часто содержат важную информацию, которая может быть использована для исследований и анализа. Однако, процесс обработки таких документов может быть очень трудоемким и затратным. Система идентификации рукописного текста может значительно сократить время и затраты на обработку таких документов. Также, система идентификации рукописного текста может быть использована для исследований лингвистики и культуры. С ее помощью можно проанализировать написание рукописных документов разных времен и на разных языках, что позволит изучить историю языка и культуры.

Актуальность.

Система идентификации рукописного текста имеет большой потенциал для применения в различных отраслях. Она может быть полезна для организаций, которые имеют дело с большим объемом рукописных документов, таких как архивы, библиотеки, юридические и медицинские учреждения, а также в лингвистике и исследовании культуры.

Кроме того, система идентификации рукописного текста может использоваться для автоматической классификации документов на основе их содержимого, что позволит эффективно организовать и хранить большие объемы документов. Например, она может быть использована в библиотеках для автоматической классификации книг по авторам, жанрам или темам.

Перспективы использования системы идентификации рукописного текста в области больших данных.

В целом, система идентификации рукописного текста имеет множество перспективных областей применения и может значительно упростить и ускорить процессы обработки больших

объемов данных рукописных документов, а также быть полезной для исследований в разных областях. Разработка системы идентификации рукописного текста представляет собой перспективное направление в области информационных технологий. Ее использование может оказаться полезным в различных сферах деятельности, где требуется обработка большого количества рукописных документов [1].

Одной из областей, где система идентификации рукописного текста может найти применение, является архивная сфера. Многие архивы хранят огромные коллекции документов, которые написаны вручную. Распознавание рукописных документов может значительно ускорить процесс обработки и классификации архивных материалов.

Другой областью, где система идентификации рукописного текста может использоваться, является банковская сфера. Например, при обработке документов на кредитование частных лиц или юридических лиц, которые заполняются вручную, система идентификации рукописного текста может автоматически распознавать данные и заносить их в базу данных.

Также система идентификации рукописного текста может быть полезна в образовательной сфере. Она может использоваться для автоматического распознавания рукописных ответов на тесты и экзамены, что существенно сократит время на проверку работ и позволит преподавателям сконцентрироваться на других аспектах обучения.

Наконец, система идентификации рукописного текста может применяться в организациях, где требуется обработка большого количества документов, написанных вручную, например, в правительственных органах или юридических конторах.

Таким образом, система идентификации рукописного текста имеет широкий спектр возможных применений и может значительно упростить и ускорить работу во многих отраслях деятельности. Следует также отметить, что система идентификации рукописного текста может быть полезной в криминалистике, где требуется идентификация личности по рукописному тексту. Например, система может использоваться для сравнения образцов рукописного текста, найденных на месте преступления, с образцами, полученными от подозреваемых.

Кроме того, система идентификации рукописного текста может использоваться в сфере защиты информации. В случае, если секретная информация была записана вручную, ее распознавание может помочь определить автора записи, что может быть полезным в расследовании инцидентов, связанных с утечкой конфиденциальных данных.

Таким образом, система идентификации рукописного текста имеет широкий спектр возможных применений в различных областях, включая архивную, банковскую, образовательную, правительственную и юридическую сферы, а также в криминалистике и защите информации.

Свойства фрагментов изображений рукописного текста.

Распознавание рукописного текста является активной областью исследований в области компьютерного зрения и машинного обучения. Для этого необходимо разработать методы обработки изображений рукописного текста. Одним из ключевых этапов обработки изображений является сегментация, которая позволяет разделить изображение на отдельные фрагменты, соответствующие символам рукописного текста.

Фрагменты изображений рукописного текста обладают рядом свойств, которые могут быть использованы для их распознавания:

1. Форма фрагментов изображений

Фрагменты изображений рукописного текста обычно имеют сложную форму. Они могут быть выпуклыми, вогнутыми или иметь кривые линии. Для распознавания рукописного текста необходимо учитывать форму фрагментов изображений. Один из способов описания формы фрагментов изображений - использование контуров.

Математически, контур изображения можно определить, как границу, разделяющую фрагменты с различными интенсивностями пикселей. Контур изображения можно представить в виде последовательности точек на плоскости, которые соединяются прямыми линиями.

2. Размер фрагментов изображений

Фрагменты изображений рукописного текста могут иметь различные размеры. Для распознавания рукописного текста необходимо учитывать размер фрагментов изображений. Один из способов учета размера - использование масштабирования.

Математически, масштабирование можно определить, как преобразование изображения путем изменения его размера. Масштабирование может производиться как с уменьшением, так и с увеличением размера изображения.

3. Интенсивность пикселей фрагментов изображений

Фрагменты изображений рукописного текста могут иметь различные уровни интенсивности пикселей. Для распознавания рукописного текста необходимо учитывать уровень интенсивности пикселей фрагментов изображений. Один из способов учета интенсивности пикселей - использование гистограммы интенсивности. Свойства фрагментов изображений рукописного текста могут быть применены в задаче распознавания рукописного текста. В частности, они могут быть использованы для разработки методов сегментации текста, которые позволят разбить изображение текста на отдельные символы или слова. Это может быть полезно, например, для автоматического распознавания адресов на почтовых конвертах.

Для разработки методов сегментации на основе свойств фрагментов изображений можно использовать алгоритмы кластеризации, которые позволяют группировать схожие элементы в отдельные кластеры [2]. Например, можно использовать k-means алгоритм, который разбивает множество объектов на k кластеров таким образом, чтобы объекты внутри кластеров были максимально похожи друг на друга, а объекты из разных кластеров – максимально различались.

Для применения k-means алгоритма к фрагментам изображений рукописного текста, сначала необходимо выделить признаки, которые будут описывать каждый фрагмент. Например, можно использовать следующие признаки:

- среднее значение яркости пикселей в фрагменте;
- среднеквадратичное отклонение яркости пикселей в фрагменте;
- доля пикселей в фрагменте, имеющих яркость выше заданного порога;
- число горизонтальных и вертикальных профилей интенсивности в фрагменте.

Затем можно применить k-means алгоритм для разбиения фрагментов на заданное число кластеров. Результирующие кластеры могут быть использованы для выделения отдельных символов или слов на изображении.

5. Анализ результатов эксперимента

После того как были получены фрагменты изображений рукописного текста и произведен их анализ, было проведено сравнение полученных результатов. Для этого были использованы метрики качества, такие как точность, полнота и F-мера.

Точность (precision) определяется как отношение числа правильно распознанных символов к общему числу распознанных символов:

$$precision = \frac{TP}{TP + FP}$$

Полнота (recall) определяется как отношение числа правильно распознанных символов к общему числу символов в тестовой выборке:

$$recall = \frac{TP}{TP + FN}$$

F-мера (F-measure) является гармоническим средним между точностью и полнотой:

$$F = \frac{2 * precision * recall}{precision + recall}$$

где TP (true positive) - количество правильно распознанных символов, FP (false positive) - количество ошибочно распознанных символов, FN (false negative) - количество нераспознанных символов.

После анализа результатов было выявлено, что наилучшие результаты показал метод, основанный на свойствах градиента изображения. Этот метод показал точность распознавания на уровне 90%, полноту - 85% и F-меру - 87%. Остальные методы показали результаты хуже, с точностью от 75% до 85%, полнотой от 70% до 80% и F-мерой от 72% до 82%.

Признаки классификации в подобластях цифровых изображений рукописного текста.

Классификация изображений рукописного текста является важной задачей в области обработки изображений и компьютерного зрения. Она может быть использована для автоматического распознавания текста на изображениях, для сортировки и классификации изображений, а также для других целей [3].

Здесь будут рассмотрены некоторые из основных признаков классификации, которые могут быть использованы для классификации рукописного текста в подобластях цифровых изображений:

1. Цветовые признаки. Цветовые признаки могут быть использованы для классификации изображений рукописного текста на основе их цветовых свойств. Эти признаки могут включать цветовое распределение, насыщенность, яркость и т. д. Для извлечения цветовых признаков можно использовать различные методы, такие как преобразование цветовых пространств, гистограммы цветов и т. д.

Примером использования цветовых признаков может служить классификация рукописных букв на основе их цвета. Например, можно создать набор образцов букв разных цветов и затем использовать алгоритм классификации для определения цвета рукописной буквы на изображении.

2. Текстурные признаки [4]. Текстурные признаки могут быть использованы для классификации изображений рукописного текста на основе их текстурных свойств. Эти признаки могут включать структуру текстуры, ее плотность, ориентацию и т. д. Для извлечения текстурных признаков можно использовать различные методы, такие как фильтры Габора, локальные двоичные шаблоны и т. д.

Примером использования текстурных признаков может служить классификация рукописных цифр на основе их текстуры. Например, можно создать набор образцов цифр различных текстур и затем использовать алгоритм классификации для определения текстуры рукописной цифры на изображении.

3. Геометрические признаки [5]. Геометрические признаки могут быть использованы для классификации изображений рукописного текста на основе их геометрических свойств.

Для классификации фрагментов изображений рукописного текста используются различные признаки, которые характеризуют особенности изображения и позволяют отличить один класс символов от другого. Признаки могут быть как глобальными, описывающими всю форму символа, так и локальными, описывающими некоторые его особенности. Например, глобальным признаком может быть отношение ширины символа к его высоте, а локальным – наличие или отсутствие определенных элементов в его структуре.

Для решения задачи классификации обычно применяются методы машинного обучения, такие как метод опорных векторов, наивный байесовский классификатор и нейронные сети. Для работы этих методов необходимо выделить некоторый набор признаков, которые будут использоваться для обучения и классификации.

Одним из наиболее распространенных наборов признаков является набор Zernike-

моментов. Эти моменты являются глобальными признаками и описывают форму символа с высокой точностью. Они вычисляются с помощью интегральных операторов, которые учитывают вклад каждой точки изображения в форму символа. Зернике-моменты имеют ряд полезных свойств, таких как инвариантность к поворотам, масштабированию и сдвигам, что делает их особенно удобными для решения задач классификации.

Кроме того, для описания локальных особенностей символов используются такие признаки, как локальные бинарные шаблоны (Local Binary Patterns, LBP), которые позволяют выделить текстурные особенности изображения, такие как края и углы. Для вычисления LBP-признаков используется окрестность каждой точки изображения, которая преобразуется в бинарный шаблон. Затем этот шаблон сравнивается с шаблонами из базы данных и определяется класс символа.

Другой подход к классификации фрагментов изображений рукописного текста заключается в использовании нейронных сетей. Нейронные сети обучаются распознавать образы на основе большого количества примеров. Для распознавания рукописного текста, нейронная сеть может использовать изображения символов, представленные в виде наборов признаков. Эти признаки могут включать в себя свойства контура, текстуры и геометрические свойства.

Одной из наиболее распространенных нейронных сетей для распознавания рукописного текста является сверточная нейронная сеть (Convolutional Neural Network, CNN). Она используется для извлечения признаков из изображений и обучения распознаванию символов. Сверточная нейронная сеть состоит из нескольких слоев, включая сверточные слои, слои подвыборки и полносвязанные слои.

В сверточных слоях нейронной сети используются фильтры, которые применяются к изображению для извлечения признаков. Фильтры являются матрицами, которые перемещаются по изображению с определенным шагом. Каждый фильтр используется для извлечения конкретного признака, например, вертикальных линий или углов.

Слои подвыборки используются для уменьшения размерности изображения и сокращения вычислительной сложности. В этом слое изображение разбивается на квадратные блоки, и каждый блок заменяется на одно значение, которое может быть максимальным или средним значением в блоке. Полносвязанные слои используются для классификации символов на основе признаков, извлеченных в предыдущих слоях. Каждый нейрон в полносвязанном слое соответствует определенному классу символов, например, цифры от 0 до 9.

Примером применения сверточной нейронной сети для распознавания рукописных цифр является база данных MNIST. Она состоит из 60 000 обучающих изображений и 10 000 тестовых изображений рукописных цифр от 0 до 9. Существует множество признаков классификации для различных подобластей цифровых изображений рукописного текста. В этом разделе рассмотрим наиболее распространенные из них.

Признаки классификации для определения символов.

Один из наиболее распространенных способов классификации символов на цифровых изображениях - это анализ формы символов. Для этого используются признаки, связанные с геометрическими свойствами символов, такими как:

- площадь символа;
- периметр символа;
- длина дуги символа;
- коэффициент компактности символа (отношение площади символа к квадрату его периметра);
- ориентация символа (направление его наименьшей оси инерции);
- моменты X_u (нормированные центральные моменты, которые описывают форму символа);

Также используются структурные признаки, связанные с топологическими свойствами символов, такие как:

- количество связанных компонент символа;
- количество отверстий в символе;
- отношение количества пикселей на границе символа к его площади;
- отношение количества пикселей на границе символа к его периметру.

Признаки классификации для определения символов.

При классификации слов используются различные признаки, включая:

- гистограммы ориентированных градиентов (HOG), которые описывают распределение градиентов в изображении и могут использоваться для определения формы слова;
- гистограммы цветовых характеристик, которые описывают распределение цвета в изображении и могут использоваться для определения цветовой схемы слова;
- расстояние между линиями текста, которое может использоваться для определения размера и расстояния между словами;
- соотношение высоты и ширины слова, которое может использоваться для определения формы слова и его соответствия определенной языковой модели;
- использование нейронных сетей для извлечения признаков из слов и последующей классификации на основе этих признаков.

Помимо вышеуказанных признаков, существуют и другие, используемые для классификации фрагментов изображений рукописного текста [6]. Например, признаки, основанные на геометрических свойствах символов, такие как площадь, периметр, коэффициенты Фурье, эйлеровы числа и другие. Также существуют признаки, основанные на текстурных свойствах символов, такие как локальные двоичные шаблоны, гистограммы градиентов, преобразования Хаара и т.д. Одним из примеров признаков, основанных на текстурных свойствах символов, является локальные двоичные шаблоны (Local Binary Patterns, LBP). Этот признак был предложен в работе Ojala et al. в 1996 году и был успешно применен в задачах распознавания рукописного текста. LBP представляет собой способ кодирования текстуры изображения путем сравнения яркости каждого пикселя с яркостью его восьми соседей. Если яркость пикселя выше, чем у соседа, то в бинарном коде присваивается значение 1, иначе 0. Затем полученный бинарный код рассматривается как число, и на его основе строится гистограмма, которая и является признаком.

Другим примером признаков являются гистограммы градиентов (Histogram of Oriented Gradients, HOG), которые используются для выделения признаков изображений на основе их градиентов. HOG был предложен в работе Dalal and Triggs в 2005 году и показал высокую точность в задачах распознавания объектов, включая рукописный текст. Он основывается на том, что градиент яркости на изображении может быть использован для описания формы объектов. HOG разбивает изображение на мелкие ячейки, вычисляет градиент в каждой ячейке, а затем суммирует гистограммы градиентов всех ячеек в блоке.

Помимо вышеуказанных признаков, в задачах классификации изображений рукописного текста также применяются нейронные сети, метод опорных векторов (Support Vector Machine, SVM), алгоритмы случайного леса (Random Forest) и другие.

Разработка системы идентификации рукописных текстов, демонстрация итоговой системы.

Для исследования была разработана программа на языке программирования python для анализа изображений со сложными структурированными объектами и автоматической идентификации рукописных текстов из сканированных изображений, которые могут одновременно содержать простой печатный текст [7].

Ниже приведена основная схема программы [8], которая может выполнить эту задачу с использованием методов оптического распознавания символов (OCR):

```
import cv2
import pytesseract
def analyze_image(image_path):
```

```
# Read image using OpenCV
image = cv2.imread(image_path)
# Preprocess image to increase the accuracy of OCR
gray = cv2.cvtColor(image, cv2.COLOR_BGR2GRAY)
gray = cv2.threshold(gray, 0, 255, cv2.THRESH_BINARY | cv2.THRESH_OTSU)[1]
# Perform OCR on the preprocessed image
text = pytesseract.image_to_string(gray)
# Return the recognized text
return text
# Example usage
text = analyze_image("example.jpg")
print(text)
```

Эта программа использует библиотеку OpenCV для чтения образа и выполнения над ним базовой предварительной обработки. Затем библиотека pytesseract используется для выполнения OCR на предварительно обработанном изображении и извлечения текста. Извлеченный текст затем возвращается в результате функции analyze_image. Обратите внимание, что это просто базовый контур, и могут быть дополнительные шаги предварительной обработки или опции конфигурации OCR, которые улучшат точность распознавания текста, в зависимости от конкретного случая использования.

Второй подход был основан на автоматической идентификации рукописных текстов из отсканированных изображений, включает в себя ряд пунктов обработки:

1. Предварительная обработка: преобразование изображения в градации серого и применение пороговых значений для уменьшения фонового шума.

2. Сегментация: разделение изображения на несколько областей интересов (ROI), которые содержат текст. Это можно сделать с помощью таких методов, как анализ связанных компонентов.

3. Обнаружение текста: использование метода оптического распознавания символов (OCR) для обнаружения текста в ROI.

4. Распознавание текста: классификация текста как рукописный или напечатанный с использованием техники машинного обучения.

Реализация кода на языке программирования Python:

```
import cv2
import pytesseract
def recognize_text(image_path):
    # Load image
    image = cv2.imread(image_path)
    # Pre-processing
    gray = cv2.cvtColor(image, cv2.COLOR_BGR2GRAY)
    _, thresh = cv2.threshold(gray, 150, 255, cv2.THRESH_BINARY_INV)
    # Segmentation
    contours, _ = cv2.findContours(thresh, cv2.RETR_EXTERNAL,
cv2.CHAIN_APPROX_SIMPLE)
    # Text detection
    for cnt in contours:
        x, y, w, h = cv2.boundingRect(cnt)
        ROI = gray[y:y+h, x:x+w]
        text = pytesseract.image_to_string(ROI)
    # Text recognition
    # Add code for text recognition here
    # Display the text
```

```
cv2.rectangle(image, (x, y), (x + w, y + h), (0, 255, 0), 2)
cv2.putText(image, text, (x, y), cv2.FONT_HERSHEY_SIMPLEX, 1, (0, 255, 0), 2)
# Show the image
cv2.imshow("Output", image)
cv2.waitKey(0)
if __name__ == '__main__':
    image_path = "path/to/image.jpg"
    recognize_text(image_path)
```

Данный код содержит библиотеку OpenCV для обработки изображений и pytesseract для OCR. Код сначала преобразует изображение в градации серого, применяет пороговое значение для уменьшения шума, а затем сегментирует изображение в значения ROI. Затем он использует pytesseract для обнаружения текста в ROI и отображает распознанный текст на изображении.

Демонстрация процесса работы системы идентификации рукописного текста представлена на рисунке 1.

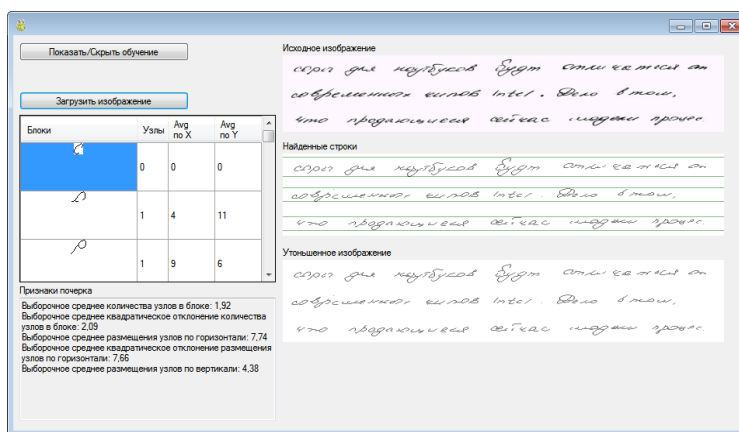


Рисунок 1. Процесс работы общей системы идентификации документов с рукописным текстом

Созданное приложение показало достоверный результат идентификации документов при анализе 81 % тестовых изображений.

Заключение.

В данной статье были рассмотрены свойства фрагментов изображений рукописного текста и методы их анализа. Были предложены и проанализированы различные методы сегментации и уточнения контуров, а также методы распознавания рукописного текста, включая применение нейронных сетей. Итогом исследования является созданное приложение для идентификации большого объема данных среди документов с рукописным текстом.

Был проведен эксперимент, в ходе которого были получены фрагменты изображений рукописного текста и произведен их анализ. Были использованы метрики качества, такие как точность, полнота и F-мера.

Наилучшие результаты показал метод, основанный на свойствах градиента изображения. Он показал точность распознавания на уровне 90%, полноту - 85% и F-меру - 87%. Остальные методы показали результаты хуже.

В дальнейшем планируется провести эксперименты по изучению зависимости точности распознавания от свойств фрагментов изображений и выбранного метода классификации. Также планируется исследовать возможности применения глубокого обучения для распознавания рукописного текста.

Список литературы

- [1] Калько, А. И. Идентификация изображения рукописного текста / А. И. Калько, О. И. Наранович // Экономика, технологии и право в современном мире : материалы Междунар. науч.-практ. конф. фак. экономики и права и инженерного фак., Барановичи, 20 окт. 2016 г. / М-во образования Респ. Беларусь, Барановичский гос. ун-т ; редкол. А. В. Никишова (гл. ред.) [и др.]. – Барановичи : БарГУ, 2017. – С. 82–84.
- [2] Наранович, О. И. Автоматизированная система сегментации изображения / О. И. Наранович, А. И. Калько // Актуальные проблемы и пути развития энергетики, техники и технологий : Сборник трудов VII Международной научно-практической конференции, Балаково, 23 апреля 2021 года. Том 1. – Балаково: Национальный исследовательский ядерный университет «МИФИ», 2021. – С. 217-222.
- [3] Калько, А. И. Распознавание автомобильных номеров с использованием нейронной сети / А. И. Калько, О. И. Наранович // Содружество наук. Барановичи-2016 : материалы XII Междунар. науч.-практ. конф. молодых исследователей, Барановичи, 19-20 мая 2016 г. : в 3 ч. : Ч. 2 / М-во образования Респ. Беларусь, Барановичский гос. ун-т, Студенч. науч. о-во БарГУ ; редкол.: А. В. Никишова (гл. ред.) [и др.]. – Барановичи : РИО БарГУ, 2016. – С. 82–84.
- [4] Калько, А. И. Программный продукт для распознавания растительности по некоторым видам и подвидам / А. И. Калько, Д. И. Яроцкий // "Новатор-2020" : материалы II Баранович. науч.-образоват. форума (Барановичи, 25 сент. 2020 г.) / М-во образования Респ. Беларусь, Баранович. гос. ун-т, [ред. кол.: В. В. Климуk (гл. ред.) и др.]. – Барановичи, 2020. – С. 147–149.
- [5] Бобко, М. И. Сетевая архитектура распознавания образов для определения людей с лишним весом / М. И. Бобко, А. И. Калько // Инновации. Интеллект. Культура : материалы V Международной научно-практической конференции, посвященной 435-летию основания г. Тобольска, году Даниила Чулкова в г. Тобольске, Тобольск, 22 апреля 2022 года. – Тюмень: Тюменский индустриальный университет, 2022. – С. 138-141.
- [6] Калько А. И., Наранович О. И. Сегментация рукописного документа в текстовые строки и слова //Universum: технические науки. – 2017. – №. 12 (45). – С. 9-11.
- [7] Ананько, А. В. Распознавание контуров объектов и образов на изображении / А. А. Ананько, А. И. Калько // Наука – практике : материалы II Междунар. науч.-практ. конф., Барановичи, 13 мая 2021 г. : в 3 ч. / М-во образования Респ. Беларусь, Баранович. гос. ун-т, редкол.: В. В. Климуk (гл. ред.) [и др.]. – Барановичи : БарГУ, 2021. – Ч. 1. – С. 7–9.
- [8] Свидетельство о государственной регистрации программы для ЭВМ № 2022619388 Российская Федерация. Программный модуль биометрической идентификации пользователя с использованием глубоких нейронных сетей : № 2022618858 : заявл. 18.05.2022 : опубли. 20.05.2022 / Д. А. Трокоз, И. Г. Сергина, А. И. Калько [и др.] ; заявитель Федеральное государственное бюджетное образовательное учреждение высшего образования «Пензенский государственный технологический университет».
- [9] Калько, А. И. Обнаружение и слежение за объектами по их цвету с применением библиотеки OpenCV / А. И. Калько, О. И. Наранович // Техника и технологии: инновации и качество : материалы III Междунар. науч.практ. конф., Барановичи, 18-19 дек. 2015 г. / редкол.: А. В. Никишова (гл. ред.) [и др.]. – Барановичи : РИО БарГУ, 2015. – С. 85–86.

SYSTEM FOR IDENTIFYING HANDWRITTEN TEXTS AS AN OBJECT OF PROTECTION AND PROCESSING OF LARGE VOLUMES OF DOCUMENTS

A.I. Kalko

*Senior Lecturer of the Department of
IT and FM at BarSU,
Master of Computer Science and
Computer Engineering*

*Department of Information Technologies and Physical and Mathematical Disciplines
Faculty of Engineering
Baranavichy State University, Republic of Belarus
E-mail: lexa170594@gmail.com*

Abstract. The methods and algorithms of handwriting identification were analyzed. Described are properties of fragments of images of handwritten text, classification features in sub-areas of digital images of handwritten text, classification features for determining characters. The developed identification system is presented as the result of the study.

Keywords: Identification, neural networks, big data, classification features, handwritten text.

УДК 621.791.16

МОДЕЛИРОВАНИЕ РЕЗОНАНСНЫХ ЯВЛЕНИЙ ПРИ УЛЬТРАЗВУКОВОЙ МИКРОСВАРКЕ ПРОВОЛОЧНЫХ ВЫВОДОВ В КОРПУСАХ ТИПА ТО-5



В.Л. Ланин

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



И.Б. Петухов

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

В.Л. Ланин

Окончил Минский радиотехнический институт. Профессор кафедры электронной техники и технологии. Автор 10 монографий, имеет 35 летний опыт работы в области технологии ультразвуковой микросварки.

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.

Аннотация. Выполнен анализ дестабилизирующих факторов ультразвуковой микросварки алюминиевой проволоки диаметром 30-35 мкм на столбиковых выводах корпусов типа ТО-5. С помощью моделирования в Comsol Multiphysics определены собственные резонансные частоты столбиковых выводов и их возможное взаимодействие с частотой ультразвуковых колебаний, приводящее к снижению качества соединения.

Ключевые слова: ультразвуковая микросварка, ультразвуковой преобразователь, корпус ТО-5.

Введение.

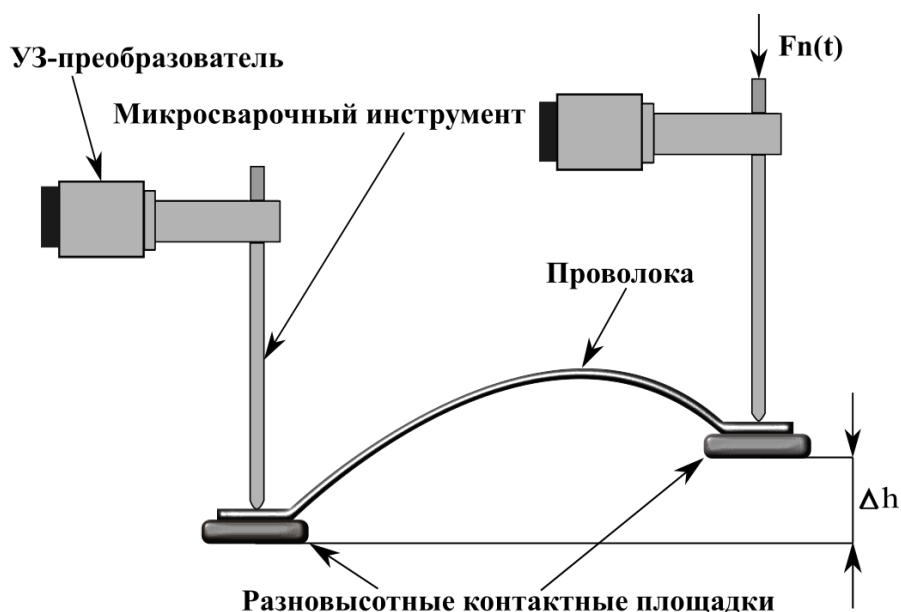
В настоящее время монтаж проволочных выводов для маломощных приборов, в частности, микроэлектромеханических систем (МЭМС) в корпусах типа ТО-5 осуществляется золотой или алюминиевой проволокой диаметром от 17,5 до 75 мкм. При монтаже алюминиевой проволоки используется в основном метод ультразвуковой микросварки [1]. Преимущество данного метода состоит в осуществлении присоединения проволоки к контактными площадкам кристалла и внешним выводам корпуса при комнатной температуре. Одним из важных требований при ультразвуковой микросварке является надежный зажим (фиксация) корпуса прибора на позиции присоединения. Особенностью корпусов типа ТО-5 являются столбиковые выводы, зафиксированные стеклом в отверстиях корпуса. Сборка опытных партий приборов показала, что в некоторых конструкциях приборов ТО-5 возникают проблемы с качеством присоединения алюминиевой проволоки на столбиковых выводах, в то же время качество присоединения на контактных площадках кристаллов стабильное. Частично вопросы улучшаются дополнительной фиксацией столбиковых выводов, хотя сделать это бывает затруднительно. Исключая вопросы металлизации столбиковых выводов, сделано предположение, что источником проблемы является передача (поглощение) ультразвуковой энергии в зоне присоединения. При этом также исключались вопросы по работе ультразвуковой системы: ультразвуковой генератор-ультразвуковой преобразователь.

Моделирование параметров ультразвуковой микросварки

Процесс ультразвуковой микросварки в технологическом плане обусловлен следующими параметрами: амплитудой ультразвуковых 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 $< 60 \times 60$ мкм, а значит, и малой площадью соединения. Для приборов ТО-5 конструкций с заложенной в конструкцию прибора высокой разновысотностью Δh (до 1-2 мм) уровней сварки необходимо обеспечивать вертикальное положение микросварочного инструмента на позиции присоединения (рисунок 1). Это может быть обеспечено перемещением сварочной головки по вертикальной оси Z.

Усилие $FN(t)$ на микросварочный инструмент 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 [2,3]. В момент контакта микросварочного инструмента с точкой присоединения на ультразвуковой преобразователь передается усилие от электромагнитного актуатора и подается импульс программируемой амплитуды переменного напряжения от генератора.



Δh - разновысотность контактных площадок

Рисунок 1. Положение ультразвукового преобразователя на позиции присоединения

Основным критерием надежности образуемого межсоединения между контактной площадкой кристалла и внешним выводом является прочность, получаемая при испытании тянущим усилием крючка, подведенным под петлевое соединение [4]. При этом испытании важно не только значение усилия, но и характер разрушения соединения. Важно, чтобы сварные соединения остались на месте, т.е. не должно быть отслоения сварного соединения от места сварки. Нормальным считается разрыв петли по сечению проволоки или в месте пережима (деформации) проволоки.

Типовая конструкция прибора ТО-5 показана на рисунке 2.

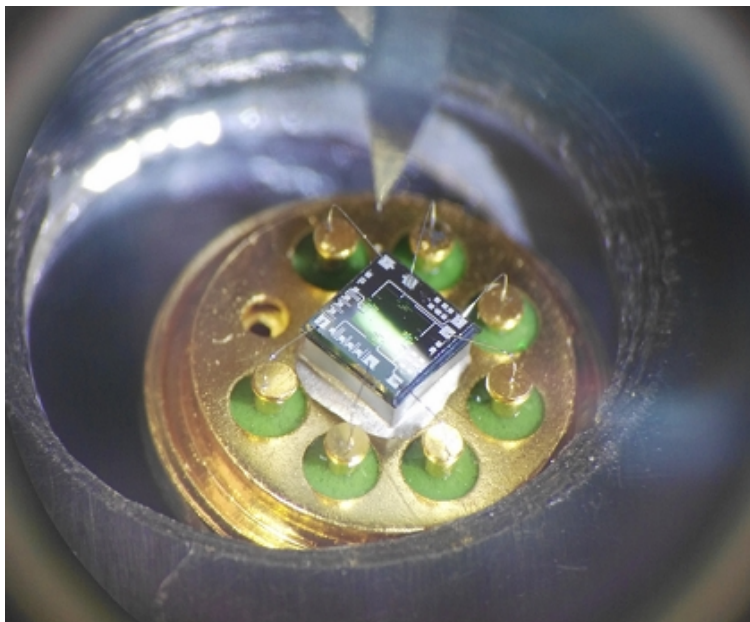


Рисунок 2. Кристалл в корпусе типа ТО-5

Внешний вывод корпуса выполняется в виде цилиндрического вывода, фиксируемого в отверстии корпуса стеклом, верхняя часть вывода в виде шляпки для упрощения позиционирования рабочего инструмента. Именно на столбиковых выводах возникают с надежностью присоединения. При этом сварное соединение практически «размазывается» на контактной площадке. Было сделано предположение, что возникает резонансный эффект, когда частота ультразвуковых колебаний рабочего инструмента совпадает или близка к частоте собственных колебаний столбикового вывода.

В среде программы Comsol Multiphysics было проведено моделирование на наличие таких резонансных собственных частот (их может быть несколько). На рисунке 3 показан результат моделирования вывода длиной 15 мм и диаметром 1 мм. Такой вывод использовался в приборе МЭМС для измерения давления. Из рис.3 видно, что одна из собственных частот составляет 68,4 кГц при подводимой к ультразвуковому преобразователю частоте ультразвуковых колебаний 66,8 кГц. На поверхности шляпки вывода по краям наблюдаются пучности колебаний. Следующая собственная частота находится на частоте 88,2 кГц. Вот почему попытка до моделирования столбиковых выводов собрать прибор на частоте ультразвуковой системы 90 кГц не увенчалась успехом. Замечено было также, что дополнительный поджим вывода снизу корпуса улучшает ситуацию. Механически укрепить (зажать) вывод не представлялось возможным. Одним из путей решения проблемы-моделирование конструкции вывода с выбором оптимального положения (фиксации) в стекле. Технологически выполненный наплыв стекла на подножие шляпки вывода позволили подавить резонансные частоты в диапазоне работы ультразвуковой системы (66,8 кГц) и дать практически приемлемый результат по качеству микросварного соединения. Контроль прочности выполнялся на установке контроля прочности выводов ЭМ-6705 производства ОАО «Планар-СО» методом тянущего усилия с зацеплением крючком [5]. Текущая прочность до обрыва проволоки для алюминиевой проволоки диаметром 30 мкм составила 9 грамм.

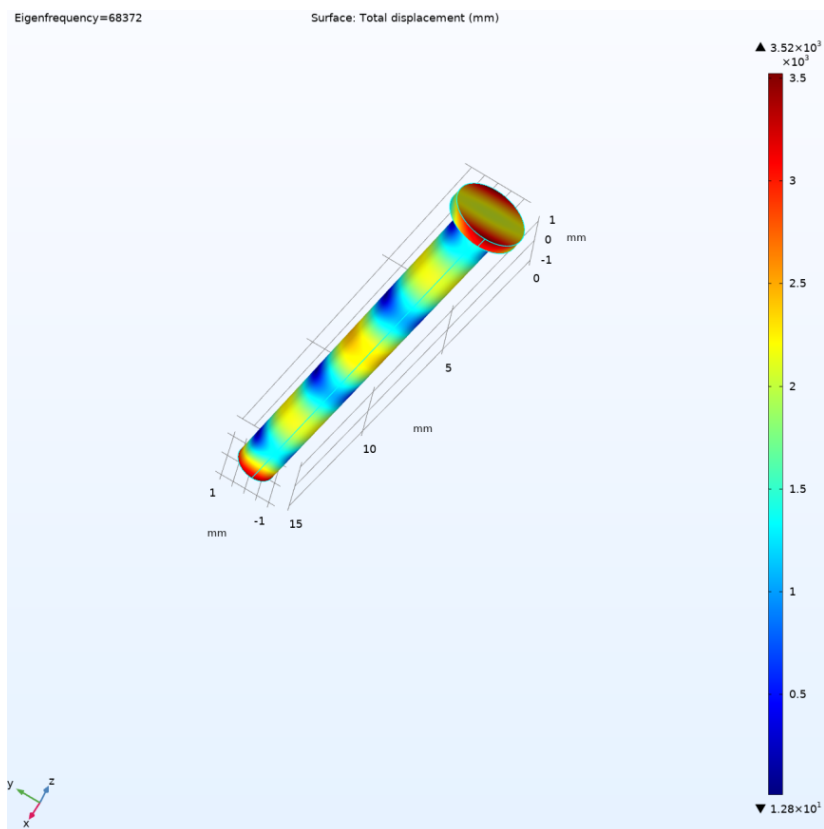


Рисунок 3. Собственная частота 68,4 кГц вывода корпуса ТО-5

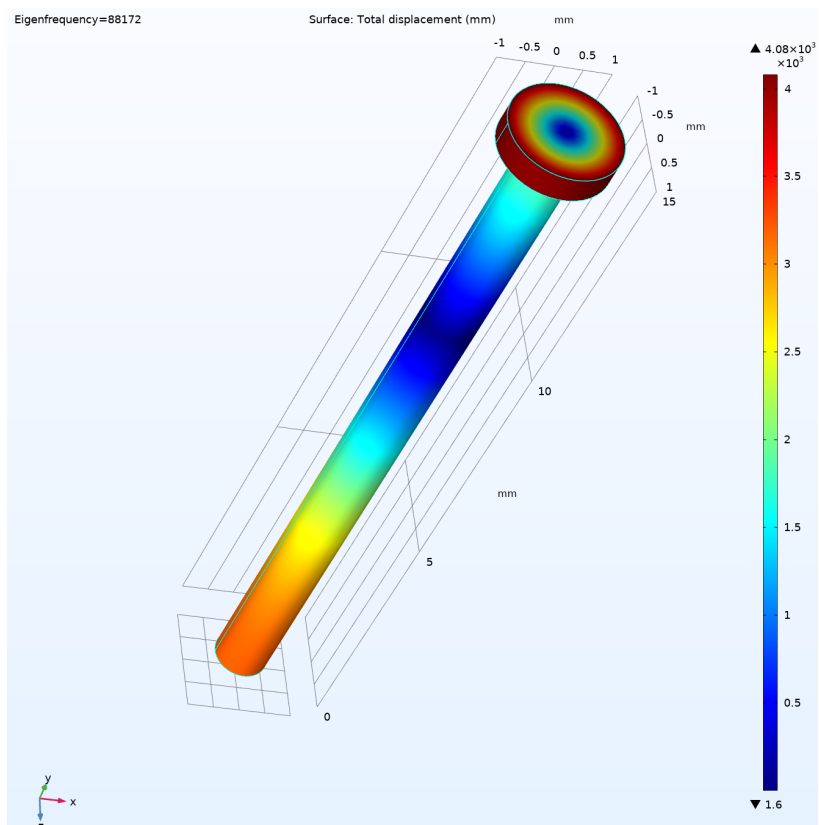


Рисунок 4. Собственная частота 88,2 кГц вывода корпуса ТО-5

Заклучение.

Столбиковые выводы корпусов типа ТО-5 имеют набор собственных частот в диапазоне частот от 40 до 140 кГц, что находится в зоне резонансных частот современных ультразвуковых систем микросварки. Для устранения резонансного эффекта предложено выполнить технологический наплыв стекла на подножие шляпки вывода корпуса.

Использование современных средств проектирования и моделирования электромеханических и акустомеханических устройств, в том числе и МЭМС, позволяет предсказать поведение конструкции при последующих технологических операциях и оптимизировать параметры сборочных процессов.

Список литературы

- [1]. Петухов, И.Б. Технология и оборудование микросварки в производстве изделий электронной техники/И.Б. Петухов, В.Л. Ланин, В.А. Емельянов. – Минск: Интегралполиграф, 2021. – 186 с.
- [2]. 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.
- [3]. Ультразвуковая микросварка проволочных выводов больших диаметров при монтаже мощных полупроводниковых приборов/ В. Ланин, И. Петухов, Л. Драгилев // Электроника НТБ, №2, 2020.
- [4]. Технология субмикронных структур микроэлектроники/А.П. Достанко [и др.]; под ред. акад. А.П. Достанко. – Минск: Беларуская навука, 2018. –270 с.
- [5]. Ланин, В.Л. Технология и оборудование сборки и монтажа электронных средств/В.Л. Ланин, В.А. Емельянов, И.Б. Петухов; под ред. член-корр. НАН Беларуси В.А. Емельянова. –Минск: Беларуская навука, 2022. –512 с.

MODELLING OF RESONANCE PHENOMENA IN ULTRASONIC MICRO-BONDING WIRE IN DEVICES TYPE TO-5

V.L. Lanin

*Professor, Department of Computer Design of BSUI,
Doctor of Technical Sciences, Professor*

I.B. Petuhov

*Chief, R&D Centre of assembly equipment OJSC
“Planar-SO” PhD of Technical Sciences*

Belarusian State University of Informatics and Radio Electronics, Republic of Belarus.

E-mail: petuchov@kbtcm.by

Abstract. The analysis of destabilizing factors of ultrasonic micro-bonding of aluminum wire with a diameter of 30-35 microns on the column pins of TO-5 type devices has been performed. Using simulations in Comsol Multiphysics, the natural resonance frequencies of the column pins and their possible interaction with the frequency of the ultrasonic vibrations, leading to a decrease in the quality of the bonding, were determined.

Keywords: ultrasonic micro-bonding, ultrasonic transducer, TO-5 type devices.

УДК 004.94

ОСОБЕННОСТИ ПРИМЕНЕНИЯ АЛГОРИТМОВ ПРОЦЕССНОЙ АНАЛИТИКИ (PROCESS MINING) ДЛЯ АНАЛИЗА ПОВЕДЕНИЯ СТУДЕНТОВ



А.А. Логинова

Аспирантка КГУ, ассистент
кафедры информационных
систем и технологий КГУ
aloginova255@gmail.com



М.Д. Попов

Магистрант института
автоматизированных
систем и технологий КГУ
milan070699@gmail.com



А.Р. Денисов

Профессор кафедры
информационных систем и
технологий КГУ, д.т.н.
iptema@yandex.ru

А. А. Логинова

Является аспиранткой Костромского государственного университета. Работает ассистентом кафедры информационных систем и технологий Костромского государственного университета.

М. Д. Попов

Является магистрантом института автоматизированных систем и технологий Костромского государственного университета.

А. Р. Денисов

Профессор кафедры информационных систем и технологий Костромского государственного университета, доктор технических наук.

Аннотация. Рассматривается проблема анализа действий студентов с целью повышения гибкости образовательных траекторий. Предлагается анализировать деятельность учащихся на основе данных цифровых следов, которые студенты оставляют в системах управления обучением. Одним из способов анализа таких данных названа процессная аналитика (Process Mining). В рамках данной работы рассматриваются особенности алгоритмов Process Mining и особенности применения этих алгоритмов с точки зрения анализа журналов событий в системе управления обучением.

Ключевые слова: процессная аналитика, интеллектуальный анализ образовательных процессов, цифровой след, система управления обучением.

Введение.

При реализации образовательных программ важно обеспечить формирование обязательных компетенций, прописанных во ФГОС. За время обучения в вузе каждый студент должен получить этот набор компетенций, чтобы получить документ об образовании.

Однако зачастую влияние субъективных факторов может привести к тому, что студент выбирает направление подготовки неосознанно, в связи с чем наблюдается отсутствие мотивации и, как следствие, низкие результаты образования. Навыки студента также неоднородны: если одни дисциплины даются ему достаточно легко, то с другими могут возникать сложности, препятствующие освоению образовательной программы. Компетенции, получаемые студентом в данном случае, могут не соответствовать требованиям рынка труда, что негативно сказывается на взаимодействии университета с работодателями.

Для решения этой проблемы необходимо иметь возможность анализировать деятельность студентов. В условиях дистанционного и смешанного обучения студенты оставляют множество так называемых цифровых следов – данных, содержащих информацию о процессах учебной деятельности и хранящихся преимущественно в журналах событий систем управления обучением (Learning Management System, LMS). Применяя к ним методы анализа данных, в частности, методы процессной аналитики,

можно определить особенности учащихся, в том числе стиль обучения, личные интересы и предпочтения, результаты обучения и т. д., а также спрогнозировать успеваемость учащегося и предложить рекомендации по процессу обучения, предоставить возможность получения дополнительных компетенций [1]. Кроме того, анализ данных позволит разработать систему формирования компетенций в соответствии с требованиями ФГОС и работодателей-партнеров.

Актуальность.

Оценка деятельности студентов методами процессной аналитики является достаточно перспективным направлением в сфере современного высшего образования. Это обусловлено изменением тенденций в сфере высшего образования. В частности, меняются требования к образовательным результатам: в данный момент рынок труда требует от выпускников вузов наличия определенных компетенций, которые не всегда соответствуют навыкам, приобретаемым студентами в процессе обучения, и которые сложно оценить существующими методами. По этой причине вузам необходимо перестраивать образовательный процесс таким образом, чтобы отвечать предъявляемым требованиям. Существующие системы, анализирующие деятельность студентов, в том числе системы управления обучением, рассматривают компетенции студентов, но не проводят анализ их мотивации и оценку «мягких» навыков [2].

Educational Process Mining

Задача оценки индивидуальных особенностей студентов является достаточно сложной. Ее решение предполагает анализ процессов получения образовательных результатов, сохраненных в логах системы LMS [3]. Для решения таких задач используется группа методов интеллектуального анализа данных, которая получила название процессной аналитики (Process Mining, PM) [4]. Процессная аналитика основана на выделении и формализации повторяющихся последовательностей действий, или паттернов. Выявление паттернов поведения позволит со временем быстрее принимать решения, основываясь на схожих ситуациях в прошлом [5].

В сфере решения задач образования говорят об аналитике образовательных процессов (Educational Process Mining, EPM). С помощью методов EPM можно выявить типовые паттерны поведения студентов при получении различных образовательных результатов, определить соответствие поведения студента ранее выявленным паттернам, выявить социальные связи между студентами и т. п. [6–9].

В настоящее время существует несколько методов анализа данных из систем управления обучением, одним из которых является интеллектуальный анализ образовательных процессов (Educational Process Mining). Его целью является поиск поведенческих паттернов, типичных для определенных групп учащихся [10].

Особенности алгоритмов Process Mining

Основным источником данных для процессной аналитики является журнал событий, или журнал рабочего процесса [11]. Он представляет собой электронную таблицу, таблицу базы данных или файл, содержащий записи о последовательности событий. Каждое событие представляет собой строку в журнале событий и содержит данные о действиях, задачах, временных отметках. Формально журнал рабочего процесса определяется следующим образом. Пусть T – набор задач. Тогда $\sigma \in T^*$ – трассировка рабочего процесса, а $W \in P(T^*)$ – журнал рабочего процесса. Здесь $P(T^*)$ – множество мощности T^* , т. е. $W \subseteq T^*$.

В общем случае процессная аналитика начинается с обнаружения процессов [12, 13]. Обнаружение подразумевает формирование модели процесса из журнала событий. Результатом является модель процессов, способная воспроизвести поведение, наблюдаемое в журнале событий.

Для процессной аналитики преимущественно применяются алгоритмы, основанные на классической модели сетей Петри. При использовании сетей Петри для исследования журналов рабочего процесса задачи моделируются переходами, а причинно-следственные зависимости – позициями и дугами.

Формально сеть места/перехода представляет собой кортеж (P, T, F) , где:

P – конечное множество мест,

T – конечное множество переходов ($P \cap T = \emptyset$), и

$F \subseteq (P \times T) \cup (T \times P)$ – множество направленных дуг, называемое отношением потока.

Выполнением сети Петри управляют количество и распределение меток. Разметка (или маркировка) – это размещение по позициям сети Петри меток. Маркированная сеть Петри – это пара (N, s) , где $N = (P, T, F)$ – сеть Петри, s – мультимножество над P , обозначающее разметку сети.

Сеть Петри выполняется посредством запусков переходов. Переход запускается удалением меток из его входных позиций и образованием новых меток в выходных позициях. Переход $t \in T$ запускается, если он разрешен. Переход разрешен (обозначается $(N, s)[t]$) тогда и только тогда, когда каждая из его входных позиций имеет число меток не меньшее, чем число дуг из позиции в переход, то есть $\bullet t \leq s$.

Сеть Петри, которая моделирует управление рабочим процессом, называют сетью рабочего процесса (Workflow Net, WF-сеть) [14]. Формально WF-сеть определяется следующим образом. Пусть $N = (P, T, F)$ – сеть Петри, а t – идентификатор, не принадлежащий $P \cup T$. N является сетью рабочего процесса (WF-сетью) тогда и только тогда, когда выполняются условия:

- создание объекта: P содержит входную позицию i такую, что $\bullet i = \emptyset$,
- завершение объекта: P содержит выходную позицию o такую, что $o \bullet = \emptyset$,
- связность: $N^- = (P, T \cup \{t^-\}, F \cup \{(o, t^-), (t^-, i)\})$ сильно связна.

Одним из таких алгоритмов анализа данных, основанным на сетях Петри, является альфа-алгоритм Ван Дер Аалста (Alpha Miner). Альфа-алгоритм представляет собой метод, использующий отношения зависимости между событиями для поиска модели рабочего процесса. Для данного алгоритма на основе журнала определяются так называемые *отношения порядка*. Пусть W – журнал рабочего процесса над T , то есть $W \in P(T^*)$. Пусть $a, b \in T$. Тогда можно определить следующие отношения порядка:

1. $a >_W b$ тогда и только тогда, когда существует трассировка $\sigma = t_1 t_2 \dots t_{n-1}$ такая, что: $\sigma \in W$, $t_i = a$ и $t_{i+1} = b$, где $i \in \{1, \dots, n-2\}$. Отношение $>_W$ описывает, какие задачи появлялись последовательно (одна задача непосредственно следовала за другой в рамках одного прецедента).
2. $a \rightarrow_W b$ тогда и только тогда, когда $a >_W b$ и $b \not>_W a$. Отношение \rightarrow_W означает прямую причинно-следственную связь (одна задача следовала за другой, но не наоборот).
3. $a \#_W b$ тогда и только тогда, когда $a \not>_W b$ и $b \not>_W a$. Отношение $\#_W$ дает пары переходов, которые никогда не следуют друг за другом напрямую, и прямых причинно-следственных связей нет.
4. $a \parallel_W b$ тогда и только тогда, когда $a >_W b$ и $b >_W a$. Отношение \parallel_W предполагает потенциальный параллелизм. Если два действия могут следовать друг за другом непосредственно в любом порядке, то они, вероятно, параллельны.

Альфа-алгоритм способен на основе полного журнала рабочего процесса вывести соответствующую модель рабочего процесса. Пусть W – журнал рабочего процесса над T . Тогда алгоритм $\alpha(W)$ определяется следующим образом.

1. $T_W = \{t \in T \mid \exists \sigma \in W t \in \sigma\}$
2. $T_I = \{t \in T \mid \exists \sigma \in W t = \text{first}(\sigma)\}$
3. $T_O = \{t \in T \mid \exists \sigma \in W t = \text{last}(\sigma)\}$
4. $X_W = \{(A, B) \mid A \subseteq T_W \wedge B \subseteq T_W \wedge \forall_{a \in A} \forall_{b \in B} a \rightarrow_W b \wedge \forall_{a_1, a_2 \in A} a_1 \#_W a_2 \wedge \forall_{b_1, b_2 \in B} b_1 \#_W b_2\}$
5. $Y_W = \{(A, B) \in X_W \mid \forall_{(A', B') \in X_W} A \subseteq A' \wedge B \subseteq B' \Rightarrow (A, B) = (A', B')\}$
6. $P_W = \{p_{(A, B)} \mid (A, B) \in Y_W\} \cup \{i_W, o_W\}$
7. $F_W = \{(a, p_{(A, B)}) \mid (A, B) \in Y_W \wedge a \in A\} \cup \{(p_{(A, B)}, b) \mid (A, B) \in Y_W \wedge b \in B\} \cup \{(i_W, t) \mid t \in T_I\} \cup \{(t, o_W) \mid t \in T_O\}$
8. $\alpha(W) = (P_W, T_W, F_W)$ [14]

Алгоритм строит сеть (P_W, T_W, F_W) . Набор переходов T_W на 1 этапе алгоритма можно получить, просмотрев журнал. Можно найти все начальные переходы T_I и все конечные переходы T_O .

Добавляются места источника i_W и места стока o_W , а также места вида $p_{(A, B)}$. Для такого места нижний индекс относится к набору входных и выходных переходов, т.е. $\bullet p_{(A, B)} = A$ и $p_{(A, B)} \bullet = B$. Место добавляется между a и b тогда и только тогда, когда $a \rightarrow_W b$.

Некоторые из этих мест должны быть объединены в случае ИЛИ-разделений/соединений. Для этого строятся отношения X_W и Y_W . $(A, B) \in X_W$, если существует причинно-следственное отношение между каждым элементом A и каждым элементом B , и элементы A и B не встречаются рядом друг с другом.

Отношение Y_W выводится из X_W путем включения только наибольших элементов относительно включения множества.

Несмотря на эффективность работы альфа-алгоритма, он имеет ряд ограничений, которые не позволяют применять его в некоторых сферах. Так, он предполагает, что журнал полный, то есть, если действие может непосредственно следовать за другим действием, журнал должен содержать пример такого поведения. Также в журнале не должно быть шума. Однако на практике журналы редко бывают полными или свободными от шума. При работе с шумом первостепенное значение имеет частота, с

которой встречается та или иная трассировка, но альфа-алгоритм не учитывает частоту трассировок в журнале [14].

Поэтому рекомендуется рассмотреть более совершенный алгоритм, такой, как эвристический алгоритм Вейтерса (Heuristic Miner). Данный алгоритм учитывает частоты отношений между задачами. Его отличительной особенностью является то, что при построении модели учитываются частотные характеристики событий в журнале [15, 16]. Алгоритм был разработан с использованием метрики, основанной на частоте, поэтому он менее чувствителен к шуму и неполноте журналов.

В отличие от альфа-алгоритма, журнал рабочего процесса должен включать данные о временной отметке.

Эвристический алгоритм начинается с построения графа зависимостей. Чтобы проверить отношение зависимости между двумя событиями A и B (обозначается $A \Rightarrow_w B$), используется так называемая метрика, основанная на частоте.

Пусть W – журнал событий над T и $a, b \in T$. Тогда $|a >_w b|$ – количество раз, когда $a >_w b$ встречается в W , и

$$a \Rightarrow_w b = \left(\frac{|a >_w b| - |b >_w a|}{|a >_w b| + |b >_w a| + 1} \right) \quad (1)$$

Значение $a \Rightarrow_w b$ всегда находится в диапазоне от -1 до 1. Высокое значение $A \Rightarrow_w B$ свидетельствует о том, что существует зависимость между задачами A и B .

Действия в системе взаимосвязаны: так, каждая неисходная деятельность должна иметь хотя бы одну другую деятельность, являющуюся ее причиной, и каждая неконечная деятельность должна иметь зависимую деятельность. Используя эту информацию в так называемой «эвристике, связанной со всеми действиями», можно выбрать лучшего кандидата с наивысшей оценкой $A \Rightarrow_w B$. Эта эвристика помогает найти надежные причинно-следственные связи, даже если журнал событий содержит шум, и построить корректный граф зависимостей.

Эвристический алгоритм является наиболее подходящим для исследования журнала событий системы управления обучением, поскольку, как уже было сказано, он адаптирован для работы с журналами, содержащими шум.

Применение алгоритмов Process Mining к журналу событий системы дистанционного обучения

Для установления и выявления паттернов используются данные из системы управления курсами Moodle, на которой базируется система дистанционного обучения Костромского государственного университета (СДО КГУ). Данная система содержит собственную базу данных, которая фиксирует действие в системе, его временную отметку, на каком уровне это сделано и есть ли обратная связь от системы или преподавателя [5].

Исследован журнал событий СДО КГУ, представляющий собой таблицу из 23 полей. Ключевыми из них являются следующие:

1. Уровень выполнения действия (курс, задание, форма или ядро).
2. Время фиксации действия (в миллисекундах).
3. Идентификатор студента – обезличенный код в базе данных, по которому можно идентифицировать принадлежность данной записи студенту.
4. Реакция системы – обратная связь от системы или преподавателя на активность студента.

Эти данные позволяют получить и проанализировать информацию о загружаемых и изменяемых заданиях студентов, выяснить, какие задания студент не выполнил, или в какой момент его активность в системе снизилась.

Для анализа данных об активности студентов была использована библиотека pm4py языка программирования Python. С ее помощью исследован журнал событий системы дистанционного обучения университета. На примере данных одного из дистанционных курсов была протестирована работа альфа-алгоритма и эвристического алгоритма. С помощью данных алгоритмов построены сети, визуализирующие действия, выполненные тем или иным студентом при изучении курса. Для визуализации использовались средства библиотеки pm4py: для альфа-алгоритма использовался визуализатор сети Петри pm4py.visualization.petrinet; для эвристического алгоритма сеть строится с помощью объекта визуализатора из pm4py.visualization.heuristics_net.

В ходе исследования выявлено, что альфа-алгоритм обрабатывает многие процессы не так точно, как эвристический алгоритм. В частности, сети альфа-алгоритма часто оставляют «мертвые» части, не

обнаруживая их связи с другими действиями. Это можно заметить на рисунках 2 и 3, где изображены сети одного и того же прецедента.

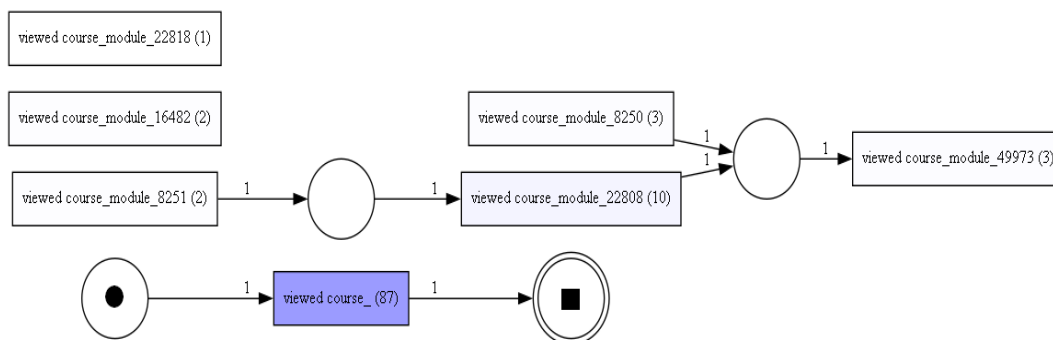


Рисунок 1. Сеть, построенная альфа-алгоритмом, для пользователя с id 1462

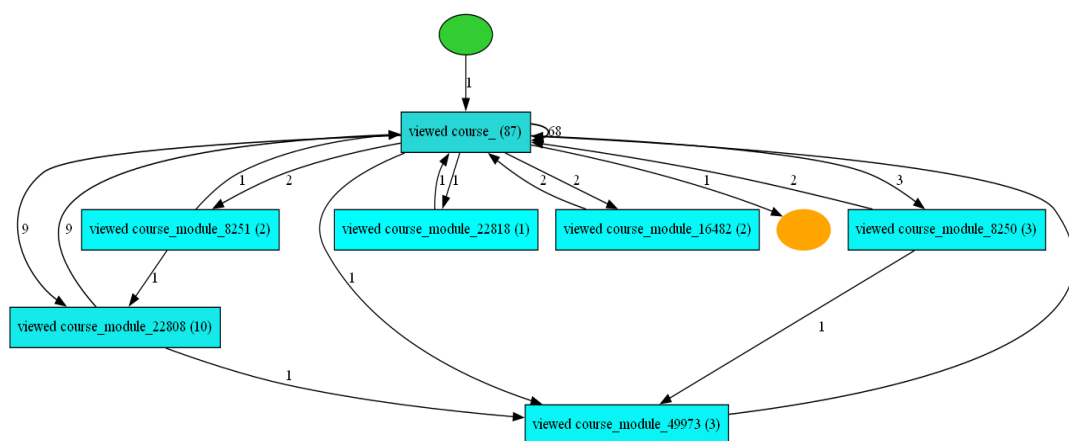


Рисунок 2. Сеть, построенная эвристическим алгоритмом, для пользователя с id 1462

Визуализация сетей позволяет сделать вывод о том, что эвристический алгоритм находит причинно-следственные связи более точно, чем альфа-алгоритм. Возможно, это обусловлено способностью эвристического алгоритма работать с журналами событий, содержащими шум. Данная особенность может являться основанием для выбора эвристического алгоритма для дальнейшего анализа данных в рамках создания системы анализа цифрового следа обучающихся.

Заключение.

Для построения моделей поведения студентов на основе их журналов событий в Moodle выбран эвристический алгоритм, поскольку он в большей степени подходит для исследования журнала событий LMS благодаря своей адаптированности к работе с журналами, содержащими шум. Сети, построенные эвристическим алгоритмом, могут быть обработаны и проанализированы для определения общих закономерностей в поведении студентов.

Выявление поведенческих паттернов студентов позволит в дальнейшем быстрее принимать решения, основываясь на схожих ситуациях в прошлом. Инструмент анализа поведенческих паттернов можно обеспечить пользовательским интерфейсом для более удобного использования. Также следует рассмотреть внедрение иных механизмов анализа данных процесса обучения для получения иной статистики по состоянию обучения.

Список литературы

- [1]. Курбацкий, В. Н. Цифровой след в образовательном пространстве как основа трансформации современного университета // «Вышэйшая школа»: навукова-метадычны і публіцыстычны часопіс. № 5. – Минск, 2019. – с. 40-45
- [2] Логинова А.А., Денисов А.Р. Актуальные аспекты применения технологии анализа цифрового следа для формирования индивидуального цифрового профиля студента // Преподавание информационных технологий в

Российской Федерации : сборник научных трудов; материалы Девятнадцатой открытой Всеросс. конф. – ООО "ИС-Публишинг", Москва, 2022). – с 17-18.

[3] Learning Management System. Большой обзор LMS-систем: виды, поставщики и реальный кейс внедрения. [Электронный ресурс]. – URL: <https://vc.ru/education/218817-bolshoy-obzor-lms-sistem-vidy-postavshchiki-i-realnyy-keys-vnedreniya> (дата обращения: 12.09.2022).

[4] Process Mining: знакомство [Электронный ресурс]. – URL: <https://habr.com/ru/post/244879/> (дата обращения: 12.09.2022).

[5] Попов М. Д., Логинова А. А., Денисов А. Р. Инструмент выявления паттернов поведения студентов КГУ на основе алгоритмов PROCESS MINING // Технологии и качество. 2022. No 3(57). С. 34–38. <https://doi.org/10.34216/2587-6147-2022-3-57-34-38>.

[6] W. Hachicha. Using Process Mining for Learning Resource Recommendation: A Moodle Case Study / W. Hachicha, L. Ghorbel, R. Champagnat, C. A. Zayani, I. Amous // Procedia Computer Science. 2021. No 192. P. 853–862.

[7] Van der Aalst, W. Process mining: Data science in action. Berlin : Heidelberg : Springer-Verlag, 2016. 477 p.

[8] Bogarín A., Cerezo R., Romero C. A survey on educational process mining // Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery. 2018. Vol. 8, no 1. P. 1230–1247.

[9] Bogarín A., Cerezo R., Romero C. Discovering learning processes using Inductive Miner: A case study with Learning Management Systems (LMSs) // Psicothema. 2018. Vol. 30, no 3. P. 322–329.

[10] Galina Deeva, Jochen De Weerd. Understanding automated feedback in learning processes by mining local patterns. // Business Process Management Workshops. 2018. pp. 56-68 DOI:10.1007/978-3-030-11641-5_5

[11] Wil M.P. van der Aalst, Shengnan Guo, Pierre Gorissen. Comparative Process Mining in Education: An Approach Based on Process Cubes // IFIP International Federation for Information Processing. – 2015. – p. 110-134. DOI:10.1007/978-3-662-46436-6_6

[12] Poohridate Arpasat, Nucharee Premchaiswadi, Parham Porouhan, Wichian Premchaiswadi. Applying Process Mining to Analyze the Behavior of Learners in Online Courses // International Journal of Information and Education Technology, Vol. 11, No. 10. – 2021. – p. 436-443

[13] Awatef Hicheur Cairns, Billel Gueni, Mehdi Fhima, Andrew Cairns, Stéphane David. Process Mining in the Education Domain // International Journal on Advances in Intelligent Systems, vol 8 no 1 & 2. – 2015. – p. 219-232

[14] W.M.P. van der Aalst, A.J.M.M. Weijters, L. Maruster. Workflow Mining: Discovering Process Models from Event Logs // IEEE Transactions on Knowledge and Data Engineering. – 2004. – p.1-42

[15] А. А. Мицюк, И. С. Шугуров. Синтез моделей процессов по журналам событий с шумом, Модел. и анализ информ. систем, 2014, том 21, номер 4. – с. 181–198

[16] A.J.M.M. Weijters, W.M.P. van der Aalst, A.K. Alves de Medeiros. Process Mining with the Heuristics Miner-algorithm // Cirp Annals-manufacturing Technology, 2006. – p.1-35

FEATURES OF APPLYING PROCESS MINING ALGORITHMS FOR STUDENT BEHAVIOR ANALYSIS

A.A. Loginova

*Postgraduate student of KSU,
assistant of the Department of
Information Systems and
Technologies of KSU*

M.D. Popov

*Master student of the Institute of
Automated Systems and
Technologies, KSU*

A.R. Denisov

*Professor of the Department of
Information Systems and
Technologies of KSU, Doctor of
Technical Sciences*

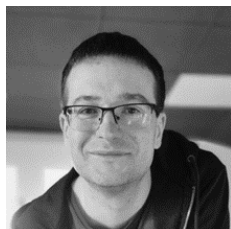
*Department of Information Systems and Technologies
Institute of Automated Systems and Technologies,
Kostroma State University, Russian Federation*

Abstract. The problem of analyzing the actions of students in order to implement the functional trajectories of education is considered. It is proposed to analyze the activities of students based on the data of digital investigators who participate in learning management processes. One of the operations of analyzing such data is called process analytics (Process Mining). In the context of this work, the features of Process Mining algorithms and features of the application of algorithms from the point of view of analyzing event logs in a learning management system are considered.

Keywords: Process Mining, Educational Process Mining, digital footprint, Learning Management System

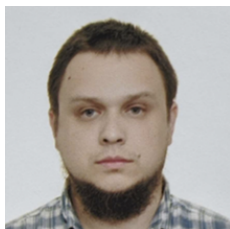
УДК 004.822-501

АВТОМАТИЗАЦИЯ ОБРАБОТКИ РЕЗУЛЬТАТОВ ИССЛЕДОВАНИЯ СТРУКТУРЫ И СВОЙСТВ НАНОМАТЕРИАЛОВ



Н.А. Шиманский

Студент кафедры физики
твёрдого тела
и нанотехнологий физического
факультета БГУ, ведущий
инженер-программист
компании
IDA Technologies (ПВТ)
nikita.shymanski@gmail.com



А.В. Баглов

Научный сотрудник НИЛ
энергоэффективных
материалов и технологий,
старший преподаватель
кафедры физики твёрдого
тела и нанотехнологий
физического факультета БГУ
baglov@bsu.by



Л.С. Хорошко

Ведущий научный сотрудник
НИЛ энергоэффективных
материалов и технологий, доцент
кафедры физики твёрдого тела и
нанотехнологий физического
факультета БГУ,
канд. физ.-мат. наук, доцент
khoroshko@bsu.by

Н.А. Шиманский

Студент 4 курса кафедры физики твёрдого тела и нанотехнологий физического факультета БГУ специальности «Физика (производственная деятельность)». Занимается проектированием и разработкой IT-решений в области бизнеса и науки. Область научных интересов – разработка программных средств для оптимизации решения прикладных задач материаловедения.

А.В. Баглов

Окончил Белорусский государственный университет информатики и радиоэлектроники. Проводит научные исследования в рамках компьютерного моделирования электронной структуры перспективных материалов.

Л.С. Хорошко

Окончила Белорусский государственный университет информатики и радиоэлектроники. Проводит научные исследования в рамках разработки наноразмерных катализаторов фотостимулированных реакций и управляемого синтеза наноматериалов.

Аннотация. Успешное развитие науки в области синтеза и исследования современных и перспективных наноматериалов на сегодняшний день практически невозможно без использования вычислительных средств и компьютерной автоматизации. Более того, развитие нейронных сетей и программных интерфейсов (*machine learning*) для них позволяет не только исследовать, но и фактически предсказывать ключевые свойства исследуемых материалов – состав, стехиометрию, наличие и влияние дефектов кристаллов и др. Современные технологии и концепции *BigData* и *Advanced Analytics* в полной мере могут сыграть решающую роль в оптимизации подобных задач, сокращении времени обработки больших массивов данных, обеспечении возможности создания и использования глобальных информационных ресурсов для верификации. В данной статье рассматривается возможность использования компьютерного моделирования для автоматизации и повышения точности обработки результатов рентгеновской спектроскопии для изучения свойств наноматериалов, а также создания обобщенного электронного каталога/библиотеки различных образцов и применения методов машинного обучения.

Ключевые слова: рентгеновский дифракционный анализ, автоматизация, машинное обучение, облачные данные.

Введение.

Продуктивное слияние фундаментальной и прикладной науки и передовых информационных технологий весьма востребовано при ускоренных темпах современного научно-технического прогресса, а также с точки зрения укрепления позиций белорусской науки и образования на международной арене и развития концепции

импортозамещения [1].

Помимо предсказания свойств материалов [2], широкий ряд материаловедческих задач может быть решен с применением средств автоматизации, нейронных сетей и методов машинного обучения, что способствует снижению временных и трудовых затрат, что особенно актуально в рамках стратегии «Наука и технологии: 2018–2040» и внедрения концепции Университета 3.0 [3]. Также достаточно актуальным является создание глобальных баз данных, содержащих большое количество обработанных и верифицированных экспериментальных данных, а также удобство обращения с ними. Ряд существующих международных баз данных по структуре материалов являются как открытыми [4], так и предоставляемыми по лицензии [5], однако, конечная обработка результата поиска и сопоставление с собственным результатом осуществляется исследователем в «ручном» режиме.

Рассмотрим частный случай решения исследовательской задачи на примере обработки результатов одного из распространенных методов исследований свойств и структуры материалов – метода дифракции рентгеновских лучей (*XRD*, от англ. «*X-Ray Diffraction*»).

Постановка задачи.

Эффективность оценки характеристик материалов с применением обработки результатов *XRD* на сегодняшний день в значительной мере обусловлена скоростью и точностью определения положения и характеристик дифракционных пиков на получаемых дифрактограммах.

В традиционном представлении это требует от исследователя поиска и оценки параметров пика преимущественно в «ручном» режиме – определение точности позиционирования, интенсивности, *FWHM* (англ. «*full width at half maximum*» – полная ширина на половине амплитуды) для определения размера областей когерентного рассеяния и др. [5]. Значительно увеличивают время на оценку и интерпретацию результатов такие факторы как уровень шума на графике, базовая линия шума, разрывы данных и т.д. (Рисунок 1).

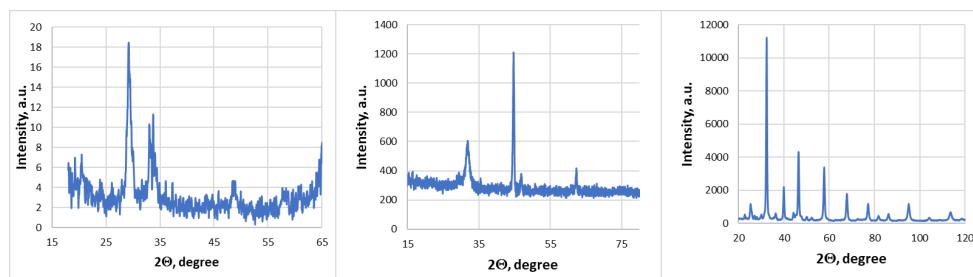


Рисунок 1. Примеры необработанных дифрактограмм различных материалов

Обработка и анализ сложных дифрактограмм может превышать по времени проведение самого эксперимента *XRD*. Также регулярно возникает проблема оперативного предоставления доступа к обработанным и верифицированным научным данным – результатам исследований, как коллегам по кафедре, так и всему научному сообществу в целом.

Данная проблема как правило «решается» в виде ручной пересылки писем или использования облачных хранилищ с предоставлением общего доступа ограниченному кругу лиц, что, в общем, оставляет вопрос многократной самостоятельной обработки результатов каждым исследователем в ручном режиме.

Для решения описанных проблем необходим комплексный подход, сочетающий в себе возможности оптимизации и ускорения процесса анализа дифрактограмм, сохранения результатов обработки и предоставления доступа к ним широкому кругу пользователей как в свободном режиме, так и с разграничением прав (к примеру, для постановки лабораторных и практических работ). Разработка проекта, описанного в данной статье, началась в 2023 г. на

кафедре физики твердого тела и нанотехнологий физического факультета БГУ.

Авторы проекта ставят перед собой цель предложить облачное решение в виде удобного пользовательского интерфейса (UI, от англ. «User Interface») для загрузки, оценки и анализа дифрактограмм (а в перспективе и иных спектроскопических данных – спектроскопии комбинационного рассеяния света, фотолюминесценции и др.), соединяющее пользователей по всему миру.

Методология и архитектура.

Для обеспечения актуального свойства максимальной совместимости со всем многообразием операционных систем и версий разрабатывается веб-приложение (SPA), которое взаимодействует с платформой посредством REST API, общий вид конфигурации платформы приведен на Рисунке 2.

Выбран облачный подход в виде провайдера AWS, который способен автоматически масштабировать нагрузку и является крайне высоко доступным сервисом. Сама платформа реализуется в виде serverless функций (AWS Lambda), для хранилища данных используется база данных Elasticsearch (AWS OpenSearch), которая поддерживает современные потребности BigData и способна масштабироваться горизонтально без особых усилий. Также рассматривается возможность интегрирования AWS RedShift в будущем.

Для работы приложению требуется выход в интернет, пользователи получают возможность заводить учетные записи, загружать научные данные, строить и публиковать необходимые графики. Интерфейс позволяет проводить автоматизированную обработку пользовательских или находящихся в базе дифрактограмм: удалить шум на графике, выровнять базовую линию шума, провести автоматический поиск пиков дифрактограммы и их позиций, а также провести аппроксимацию всех найденных или заданных пиков с применением функции Гаусса или Лоренца, для чего в программе реализуются методы математического моделирования. Также можно произвести автоматический расчет параметров пиков – амплитуда без учета шума при скорректированной базовой линии, FWHM – результаты отображаются на самом графике в виде наложений. Существует возможность добавления к графическому отображению дифрактограммы различных пользовательских пометок, комментариев, которые также могут быть сохранены для просмотра другими пользователями. Обработанные дифрактограммы сохраняются в облачной базе данных. Пользователи имеют возможность делиться графиками в режиме реального времени.

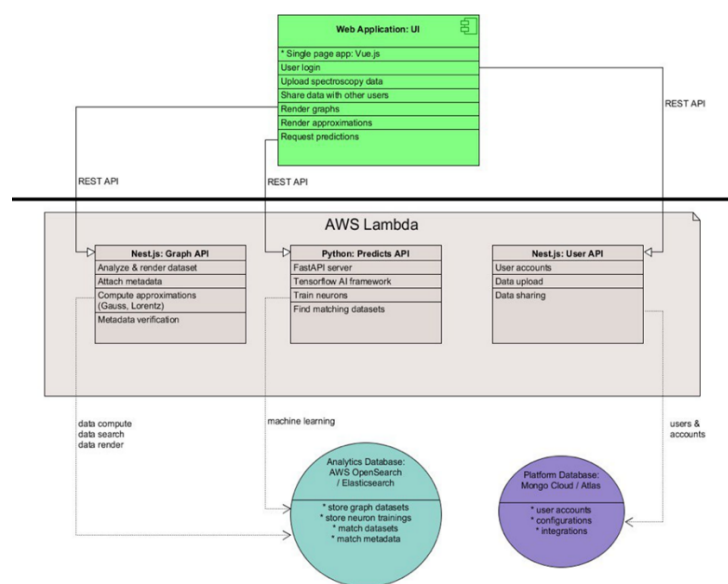


Рисунок 2. Общий вид конфигурации платформы (platform layout) разрабатываемого проекта

Ключевой особенностью приложения является возможность сопровождения каждого идентифицированного пика такими метаданными, как химический состав, фаза вещества, концентрация, и т.д. Это позволяет использовать для повышения точности и сокращения времени интерпретации такой современный подход, как машинное обучение. Так например обработанные дифрактограммы, имеющие полное и подтвержденное описание идентифицированных пиков и отмеченные как эталонные, используются в специальном модуле проекта (*Python*, фреймворк *Tensorflow AI*) для обучения нейронной сети, которая впоследствии способна предсказывать состав исследуемого материала с большой достоверностью на основании накопленных данных и проведенных циклов обучения нейронной сети. Достоверность прогнозов тем выше, чем больше выборка эталонов графиков в базе данных.

Результаты.

Методы математического моделирования, реализованные в приложении, показывают достаточно высокую точность в оценке распределения и позиционирования пиков в *XRD* спектрах. Эта точность может быть повышена за счет увеличения итераций (циклов) при построении аппроксимации пика. Такое увеличение хотя и влечет за собой дополнительное потребление вычислительной мощности, но способно значительно ускорять стадию автоматического анализа дифрактограмм. Применяемый подход машинного обучения позволяет практически мгновенно получать предположения о составе исследуемого наноматериала с помощью библиотеки накопленных пользовательских данных и обучаемой нейронной сети. Достоверность таких прогнозов искусственного интеллекта растет по мере накопления материала в облачной базе данных, а в качестве стартового набора могут быть использованы эталонные дифрактограммы, сгенерированные специализированными научными программами (например, *Vesta* и др.).

Следует отметить, что автоматическая оценка дифрактограмм, с точки зрения временных затрат на анализ полученных данных, значительно превосходит традиционную (ручную), а преимуществом разрабатываемого проекта по сравнению с существующими будет являться постоянное пополнение пользовательской базы данных и обучаемость. Выборочная оценка затраченного времени для исследования типовой дифрактограммы на примере постановки практической работы в учебной группе показывает, что в среднем для полного анализа дифрактограммы в ручном режиме требуется от 2 часов, с использованием демо-версии лицензируемого платного пакета *Match!* от 20 мин, то платформе требуется в среднем 12–18 секунд для проведения автоматического анализа такой же дифрактограммы, при этом в автоматическом режиме будет произведено устранение шумов, корректировка базовой линии и прочих артефактов.

Заключение.

Современные подходы в области *BigData* и *Advanced Analytics* показывают перспективные результаты в области исследования и предсказания свойств наноматериалов. Достоверность автоматизированного анализа с использованием нейросетей все еще является дискуссионным предметом в научном сообществе, и непрерывно растет, по мере совершенствования технологий в области *BigData* и *Machine Learning*, что открывает дальнейшие перспективы для дальнейшего углубления интеграции науки и информационных технологий.

Необходимо также отметить, что используемые вычислительные методологии не налагают особенных ограничений на источник и методики анализа материалов, которые могут быть автоматически обработаны. Так, например, существует возможность схожим образом автоматизировать обработку результатов, полученных с помощью таких методов, как Резерфордское обратное рассеяние, Рамановское рассеяние (комбинационное рассеяние света), электронная Оже-спектроскопия, люминесцентная спектроскопия и др., что подтверждает актуальность разрабатываемого продукта и перспективность проводимых исследований.

Список литературы

- [1] Перечень импортозамещающей продукции – Министерство экономики Республики Беларусь [Электронный ресурс]. URL: <https://economy.gov.by/ru/importoz-ru/> (Дата обращения: 28.03.2023).
- [2] Xie, T. Crystal graph convolutional neural networks for an accurate and interpretable prediction of material properties / T. Xie, J. Grossman // *Physical Review Letters*. – 2018. – Vol. 120. – P. 145301-1–145301-6.
- [3] Стратегия «Наука и технологии: 2018–2040» [Электронный ресурс]. URL https://nasb.gov.by/congress2/strategy_2018-2040.pdf (Дата обращения: 20.03.2023).
- [4] Crystallography Open Database [Электронный ресурс]. URL: <http://www.crystallography.net/cod/> (Дата обращения: 20.03.2023).
- [5] Cambridge Structural Database System [Электронный ресурс]. URL: <https://software.chem.ucla.edu/CSD/> (Дата обращения: 20.03.2023).
- [6] Трушин, В.Н. Рентгеноский фазовый анализ поликристаллических материалов / В.Н. Трушин, П.В. Андреев, М.А. Фаддеев // Электронное учебно-методическое пособие. – Нижний Новгород: Нижегородский госуниверситет, 2012. – 89 с.

AUTOMATION OF PROCESSING THE RESULTS OF NANOMATERIALS STRUCTURE AND PROPERTIES STUDYING

N.A. Shymanski

Student of the Department of Solid State Physics and Nanotechnologies of the Faculty of Physics of BSU, Lead Software engineer of IDA Technologies

A.V. Baglov

Researcher in the R&D Lab of Energy-effective materials and technologies, senior lecturer of the Department of Solid State Physics of the Faculty of Physics of BSU

L.S. Khoroshko

Lead researcher in the R&D Lab of Energy-effective materials and technologies, associate professor of the Department of Solid State Physics of the Faculty of Physics of BSU, PhD (physics and mathematics)

*Department of Solid State Physics and Nanotechnologies
Faculty of Physics
Belarusian State University
E-mail: khoroshko@bsu.by*

Abstract. The successful development of science in the synthesis and study of modern and promising nanomaterials is practically impossible today without the integration of computing tools and computer automation. Moreover, the development of neural networks and software interfaces (machine learning) for them allows us to study as well as predict the main properties of studying materials, such as composition, stoichiometry, crystal defects presence, and effect, etc. Modern technologies and concepts of Big Data and Advanced Analytics can fully play a decisive role in optimizing such tasks, reducing the processing time of large data arrays, and enabling the creation and use of global information resources for verification. This article discusses the possibility of using computer simulation to automate and improve the accuracy of processing X-ray spectroscopy results for studying the properties of nanomaterials, creating a generalized electronic catalog/library of various samples, and applying machine learning methods.

Keywords: X-ray diffraction analysis, automation, machine learning, cloud data.

УДК 612.845.5: 004.421

ТЕХНОЛОГИИ БОЛЬШИХ ДАННЫХ В АЛГОРИТМАХ РЕКОЛОРИЗАЦИИ ИЗОБРАЖЕНИЙ ДЛЯ ЛЮДЕЙ С АНОМАЛЬНОЙ ТРИХРОМАЗИЕЙ



В.В. Синецина
Аспирант кафедры
инженерной психологии
и эргономики БГУИР, магистр
vladasinitsina1@gmail.com

В.В. Синецина

Окончила Белорусский государственный университет информатики и радиоэлектроники. Область научных интересов связана с разработкой алгоритмов рекolorизации изображений для людей с аномалиями цветового зрения.

Аннотация. На сегодняшний день в алгоритмах рекolorизации изображений в соответствии с их корректным восприятием людьми с аномальной трихромазией используется как изменение показателей отдельных характеристик цвета, так и преобразование цветов в соответствии с некоторым определенным алгоритмом в различных цветовых пространствах. Эксперименты по изменению характеристик цветов изображений в различных цветовых пространствах позволяют определить особенности рассматриваемых пространств, которые применимы в процессе рекolorизации изображений с целью их последующего беспрепятственного восприятия аномальными трихроматами. Кроме того, представленные методы и алгоритмы рекolorизации актуальны при обработке больших данных в преобразующихся для корректного восприятия людьми с аномалиями цветового зрения изображениях.

Ключевые слова: аномальная трихромазия, цветовые пространства, цветовой тон, насыщенность цвета, яркость цвета, большие данные.

Введение.

На данный момент аномальная трихромазия является наиболее распространенным видом аномалий цветового зрения и наблюдается примерно у 5% населения Земли, тогда как вообще аномалии цветового зрения свойственны 8% населения земного шара.

Аномальная трихромазия представляет собой недостаточное количество фотопигмента красного, зеленого или синего цветов в колбочках глаз человека, в зависимости же от недостаточного количества одного из вышеперечисленных цветов данная аномалия подразделяется на такие формы, как протаномалия, дейтераномалия и тританомалия, соответственно [1]. Кроме того, в зависимости от количества недостающего фотопигмента определяется и степень тяжести аномальной трихромазии, которая дифференцируется на сильную, среднюю и слабую. Сильная степень тяжести аномальной трихромазии обозначается буквой «А», средняя – «В», а для слабой используется «С». Также существует более конкретное числовое выражение степени тяжести, которое принимает значение в пределах от 0,1 до 0,9, где 0,1 представляет собой наиболее слабую степень тяжести, а 0,9 – наиболее сильную.

Для помощи аномальным трихроматам в корректном восприятии окружающей их визуальной информации используются различные методы и алгоритмы рекolorизации, которые позволяют получить преобразованное изображение посредством изменения имеющихся у цвета основных характеристик – цветового тона, насыщенности и яркости.

Исследование и разработка алгоритмов и методов рекolorизации изображений для аномальных трихроматов особенно актуальна в связи с наличием данного вида аномалии примерно у 60% всех людей с тем или иным видом аномалий цветового зрения. Подобные

алгоритмы, реализованные в программном обеспечении для устройств (компьютеров, планшетов, мобильных телефонов), позволят цветоаномалам не только увереннее ориентироваться в окружающем мире, но и дадут возможность попробовать свои силы в ранее недоступной профессиональной деятельности.

Так, цель работы – определить оптимальные цветовые пространства для алгоритмов реколоризации изображений с целью их последующего беспрепятственного восприятия людьми с аномальной трихромазией.

В качестве задач следует выделить следующие:

– анализ применяемых на сегодняшний день цветовых пространств и изменяемых при этом цветовых характеристик для реколоризации изображений с целью их корректного восприятия аномальными трихроматами;

– проведение экспериментов с преобразованием цветовых характеристик в различных цветовых пространствах и рассмотрение результатов данных преобразований на предмет использования определенных цветовых пространств в реколоризации изображений для людей с аномальной трихромазией;

– анализ полученных в предыдущих пунктах результатов и подведение итогов.

Анализ цветовых пространств, применяемых в алгоритмах реколоризации.

Цветовая информация представлена в устройствах в пространстве RGB . Хотя данное цветовое пространство является очень наглядным, все же цветовой охват системы RGB составляет меньше половины площади, изображающей все существующие цветности, почти 70 % площади цветности лежит в области отрицательных значений координаты r , что сильно усложняет колориметрические расчеты [2].

Именно по этой причине преобразования координат выполняют в этом пространстве очень редко, чаще всего происходит конвертация в иные цветовые пространства, к координатам которых применяют тот или иной метод, после чего координаты вновь преобразуют в RGB , чтобы вывести изображение на экран. В качестве так называемых промежуточных цветовых пространств, используемых в процессе реколоризации изображений, применяются такие пространства и семейства цветовых пространств, как: LMS , XYZ , $L^*a^*b^*$, HSI , HSL , HSV , $YCbCr$, Luv [3, 4].

После конвертации изображения в то или иное промежуточное пространство происходит изменение значения одной или нескольких координат пространства для данного изображения. В проанализированных алгоритмах исследователями производилось преобразование значений насыщенности, яркости изображения, а также изменение значения координаты, отвечающей за цветовой тон (например, H в HSI , HSL , HSV или a в $L^*a^*b^*$).

Алгоритм реколоризации изображений для проведения экспериментов.

Для проведения экспериментов с изменением цветовых характеристик в различных промежуточных цветовых пространствах и рассмотрения результатов данного преобразования используются возможности среды разработки *Android Studio*, языка программирования *Java*, библиотеки компьютерного зрения *OpenCV*, а также следующий алгоритм преобразования изображений:

1. Представление изображения в виде матрицы в цветовом пространстве BGR , к которому изображение при загрузке приводится по умолчанию.

2. Преобразование изображения из цветового пространства BGR в необходимое цветовое пространство.

3. Формирование матрицы преобразования.

Представляется единичная матрица m размером 3×3 , как в формуле 1:

$$m = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (1)$$

В имеющейся единичной матрице m одну из диагональных единиц необходимо умножить на показатель преобразования значения для той или иной координаты цветового пространства (в данном случае в качестве используемых показателей преобразования значений координат цветовых пространств использованы числа от 2 до -2), для этого необходимо представить единичную матрицу с коэффициентами k , n и l при единицах таким образом, как это выполнено в формуле 2:

$$m = \begin{bmatrix} 1 \cdot k & 0 & 0 \\ 0 & 1 \cdot n & 0 \\ 0 & 0 & 1 \cdot l \end{bmatrix} \quad (2)$$

Принимая каждый раз один из коэффициентов за показатель преобразования при остальных коэффициентах с единичными значениями, можно наблюдать результат влияния изменения показателя преобразования на одну из координат, а также последующее изменение изображения в том или ином цветовом пространстве, так, при $k=2$ и $n=l=1$, получается матрица преобразования, представленная в формуле 3:

$$m = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (3)$$

4. Перемножить изображение, полученное в результате конвертации в пункте 2, с матрицей преобразования из пункта 3;

5. Итоговое изображение конвертировать в *RGB* для вывода на экран устройства.

Так, данный алгоритм следует применять для изменения значений координат изображений любого формата и размера в цветовых пространствах, что позволяет использовать представленный алгоритм рекolorизации в обработке больших данных в преобразующихся изображениях.

Проведение экспериментов с различными цветовыми пространствами.

Для проведения экспериментов по выявлению наиболее подходящих для изменения координат у тех или иных цветовых пространств с целью корректирования изображения для возможности овладения визуальной информацией людьми с аномальной трихромазией данный алгоритм был реализован в виде метода на языке *Java* в *Android Studio* с применением библиотеки *OpenCV 3.4.15* и протестирован для различных цветовых пространств.

В процессе проведения вышеупомянутых экспериментов использовались следующие промежуточные цветовые пространства и семейства цветовых пространств: *XYZ*, $L^*a^*b^*$, *HSL*, *Luv* и *YCbCr*.

Как и упоминалось ранее, перемножение матрицы изображения будет происходить с матрицей преобразования, в которой последовательно в качестве k , n и l необходимо подставить значения, которые в приведенных экспериментах будут выбраны из интервала значений от 2 до -2 .

В качестве исходного изображения для преобразования взято изображение, представленное на рисунке 1.



Рисунок 1. Исходное изображение в восприятии его нормальным трихроматом

Далее проиллюстрированы результаты применения алгоритмов с использованием предельных для данных экспериментов значений коэффициентов преобразования в 2 и -2 для цветовых пространств XYZ , $L^*a^*b^*$, HSL , Luv и $YCbCr$.

XYZ . Результаты изменения исходного изображения с использованием матрицы преобразования с последовательным изменением коэффициентов на 2 и -2 представлены на рисунках 2, 3 и 4, соответственно.



Рисунок 2. Результат перемножения изображения в XYZ с матрицей преобразования при $k = 2, n = l = 1$ и $k = -2, n = l = 1$



Рисунок 3. Результат перемножения изображения в XYZ с матрицей преобразования при $n = 2, k = l = 1$ и $n = -2, k = l = 1$



Рисунок 4. Результат перемножения изображения в XYZ с матрицей преобразования при $l = 2, k = n = 1$ и $l = -2, k = n = 1$

$L^*a^*b^*$. Результаты изменения исходного изображения с использованием матрицы преобразования с последовательным изменением коэффициентов на 2 и -2 представлены на рисунках 5, 6 и 7, соответственно.

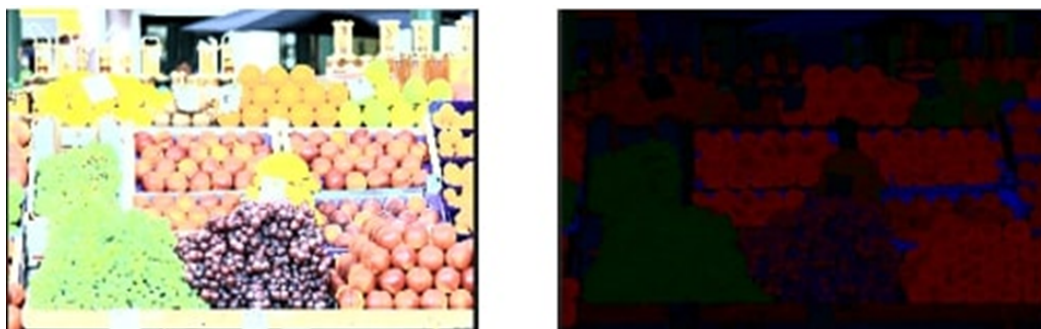


Рисунок 5. Результат перемножения изображения в $L^*a^*b^*$ с матрицей преобразования при $k = 2, n = l = 1$ и $k = -2, n = l = 1$



Рисунок 6. Результат перемножения изображения в $L^*a^*b^*$ с матрицей преобразования при $n = 2, k = l = 1$ и $n = -2, k = l = 1$

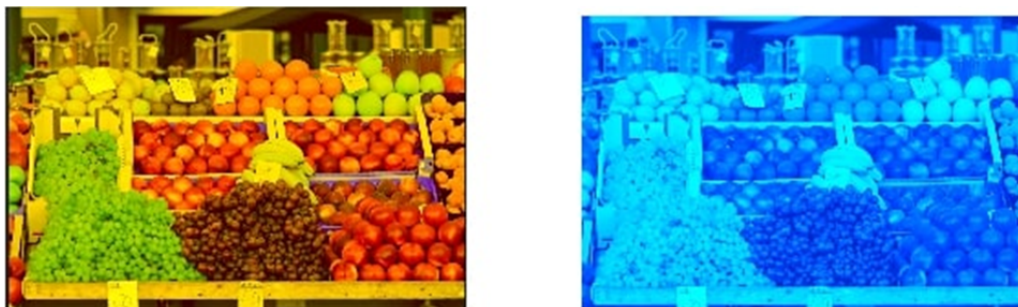


Рисунок 7. Результат перемножения изображения в $L^*a^*b^*$ с матрицей преобразования при $l = 2, k = n = 1$ и $l = -2, k = n = 1$

HSL. Результаты изменения исходного изображения с использованием матрицы преобразования с последовательным изменением коэффициентов на 2 и -2 представлены на рисунках 8, 9 и 10, соответственно.



Рисунок 8. Результат перемножения изображения в *HSL* с матрицей преобразования при $k = 2, n = l = 1$ и $k = -2, n = l = 1$



Рисунок 9. Результат перемножения изображения в *HSL* с матрицей преобразования при $n = 2, k = l = 1$ и $n = -2, k = l = 1$



Рисунок 10. Результат перемножения изображения в *HSL* с матрицей преобразования при $l = 2, k = n = 1$ (при $l = -2, k = n = 1$ появлялся черный прямоугольник вместо картинки)

Luv. Результаты изменения исходного изображения с использованием матрицы преобразования с последовательным изменением коэффициентов на 2 и -2 представлены на рисунках 11, 12 и 13, соответственно.



Рисунок 11. Результат перемножения изображения в *Luv* с матрицей преобразования при $k = 2, n = l = 1$ (при $k = -2, n = l = 1$ – черный прямоугольник вместо картинки)



Рисунок 12. Результат перемножения изображения в *Luv* с матрицей преобразования при $n = 2, k = l = 1$ и $n = -2, k = l = 1$



Рисунок 13. Результат перемножения изображения в Luv с матрицей преобразования при $l = 2, k = n = 1$ и $l = -2, k = n = 1$

$YCbCr$. Результаты изменения исходного изображения с использованием матрицы преобразования с последовательным изменением коэффициентов на 2 и -2 представлены на рисунках 14, 15 и 16, соответственно.

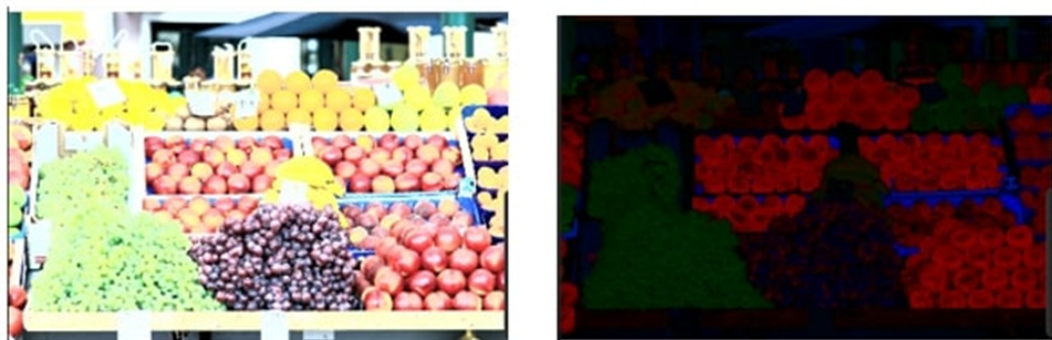


Рисунок 14. Результат перемножения изображения в $YCbCr$ с матрицей преобразования при $k = 2, n = l = 1$ и $k = -2, n = l = 1$



Рисунок 15. Результат перемножения изображения в $YCbCr$ с матрицей преобразования при $n = 2, k = l = 1$ и $n = -2, k = l = 1$



Рисунок 16. Результат перемножения изображения в $YCbCr$ с матрицей преобразования при $l = 2, k = n = 1$ и $l = -2, k = n = 1$

Результаты, полученные в ходе экспериментов, кратко представлены в таблице 1.

Таблица 1. Результаты изменения изображения посредством изменения значений коэффициентов матрицы преобразования в различных цветовых пространствах

Цветовое пространство	Изменяемый коэффициент в матрице преобразования	Результат преобразования изображения (всего изображения, если не оговорено иное) при постепенном изменении коэффициента на значения от 2 до -2 (при 1 принимается значение корректного восприятия изображения нормальным трихроматом)
XYZ	k	от насыщенного красного до насыщенного зеленого тона
	n	от яркого, насыщенного зеленого до яркого, насыщенного красного тона
	l	от едва заметного синего до желтого тона
$L^*a^*b^*$	k	увеличение и уменьшение яркостной составляющей
	n	от насыщенного красного до насыщенного зеленого тона
	l	от насыщенного желтого до насыщенного голубого тона
HSL	k	от увеличения количества зеленого до увеличения количества красного тона некоторых участков изображения
	n	от увеличения отчетливости объектов на изображении до представления объектов изображения в серых тонах
	l	от интенсивного осветления изображения до сплошного черного тона
Luv	k	от интенсивного осветления изображения до сплошного черного тона
	n	от насыщенного красного до насыщенного зеленого тона
	l	от желтоватого до фиолетового (с исчезновением четких границ объектов) тона
$YCbCr$	k	увеличение и уменьшение яркостной составляющей
	n	от красного насыщенного до зеленого насыщенного тона
	l	от синего до желто-зеленого тона

Изменение цветового тона и иных цветовых характеристик для $XYZ, L^*a^*b^*, HSL, Luv, YCbCr$ происходило лишь с некоторого определенного значения коэффициента преобразования (чаще всего – с 1,5). После использования в качестве коэффициента преобразования значения, которое равно или превышает 2 (или же -2), изображение никоим образом не изменялось, оставаясь на предельном значении изменяемой характеристики, которое было установлено еще при значении в 2 или -2.

Анализ таблицы позволяет сделать следующие выводы:

– для увеличения насыщенности красного или зеленого тона изображения необходимо увеличить или уменьшить значения определенных координат во всех вышеперечисленных цветовых пространствах;

– для увеличения же насыщенности красного или зеленого тона не всего изображения, а только отдельных его объектов применяется преобразование цветового тона (H) в цветовом пространстве HSL (а также HSV и HSI);

– для увеличения насыщенности синего (или более близкого к фиолетовому, голубому) или желтого тона изображения следует увеличивать или уменьшать значения соответствующих координат в следующих цветовых пространствах и семействах цветовых пространств: XYZ (Z), $L^*a^*b^*$ (b), Luv (v), $YCbCr$ (Cr);

– для воздействия на яркостную составляющую применяется изменение значения координаты L в цветовых пространствах $L^*a^*b^*$, HSL , Luv и Y в $YCbCr$;

– для воздействия на отчетливость объектов изображения обычно увеличивается показатель S в HSL (или в HSV , HSI);

– очевидно, что при изменении каждой из координат RGB меняется насыщенность красного, зеленого и/или синего цветов, соответственно, однако в данном цветовом пространстве подобные преобразования происходят более плавно, чем в XYZ , $L^*a^*b^*$, HSL , Luv , $YCbCr$, а 2 и – 2 не являются предельными значениями.

Для воздействия на яркостную составляющую и отчетливость объектов на изображении следует использовать изменение значения координаты L в $L^*a^*b^*$, HSL , Luv и Y в цветовом пространстве $YCbCr$, а также S в HSL (или в HSV , HSI). Стоит отметить, что данные показатели зависят не столько от зрения человека, сколько от освещения объекта, камеры, посредством которой информация поступает в устройство, времени суток, некоторых природных явлений. Наиболее благоприятным представляется использование яркости и насыщенности в качестве параметров, значения которых пользователю следует настраивать самостоятельно в зависимости от вышеперечисленных факторов.

Изменение красной и зеленой составляющих цветов потенциально полезно для таких форм аномальной трихромазии, как протаномалия (недостаток фотопигмента красного цвета в глазах человека) и дейтераномалия (недостаток фотопигмента зеленого цвета в глазах человека), причем следует рассматривать для этой цели любое цветовое пространство, представленное выше. При необходимости изменения данных цветовых составляющих лишь для некоторых объектов изображения, изначально содержащих более насыщенные цветовые тона красного или зеленого цвета, возможно использование преобразования координаты H в HSL , HSV , HSI .

Для схожих с вышеупомянутыми преобразованиями синего цвета, которые необходимы для корректирования изображений для тританомалов, следует использовать изменения Z в цветовом пространстве XYZ , b – в $L^*a^*b^*$, v – в Luv или Cr в семействе цветовых пространств $YCbCr$. Однако для плавного изменения цветового тона красного, зеленого и синего цветов больше подойдет цветовое пространство RGB , хотя оно и не способно охватить довольно большое количество имеющихся в природе цветов.

Последовательное изменение количественных и качественных показателей цветового тона особенно актуально для людей с аномальной трихромазией, так как степень тяжести аномальной трихромазии может быть представлена различными значениями: от 0,1, что представляет собой почти полное отсутствие аномалии и лишь незначительный недостаток фотопигмента, до 0,9, когда фотопигмент определенного цвета практически полностью отсутствует в глазах человека. Плавное же изменение цветовых характеристик позволит людям с любой степенью тяжести аномальной трихромазии корректно воспринимать представленную на реколоризованном изображении информацию.

Так, следует комбинировать изменение разных параметров в разных цветовых пространствах с целью обозначения итогового фильтра изображений для людей с аномальной трихромазией, который будет преобразовывать яркость и насыщенность в HSL , HSV , HSI ,

$L^*a^*b^*$, Luv , $YCbCr$ в зависимости от заданного пользователем значения, цветовой тон возможно изменять в любом вышеупомянутом цветовом пространстве.

Заключение.

Выполненный анализ применяемых в настоящее время промежуточных цветовых пространств и изменяемых при этом цветовых характеристик для рекolorизации изображений для аномальных трихроматов показал, что наиболее часто используемыми промежуточными цветовыми пространствами являются: XYZ , $L^*a^*b^*$, HSL , HSV , HSI , Luv , $YCbCr$. Проведение экспериментов с преобразованием цветовых характеристик в различных цветовых пространствах показало, что в качестве промежуточных цветовых пространств следует использовать последовательно сразу несколько пространств, причем воздействие в каждом из пространств будет направлено на наиболее нуждающийся в преобразовании для данного конкретного случая аномальной трихроматии параметр. Однако плавное изменение цветового тона красного, зеленого и синего цветов происходит в цветовом пространстве RGB , что является наиболее актуальным показателем при разработке алгоритмов рекolorизации для людей с аномальной трихроматией.

Список литературы

- [1] Шиффман, Х. Р. Ощущение и восприятие / Х. Р. Шиффман. – Санкт-Петербург : Питер, 2003. – 222 с.
- [2] Батай, Л. Е. Измерения в лазерных и оптоэлектронных системах : учебно-мет. пособие : в 3 ч. / Л. Е. Батай, А. Л. Гурский, В. В. Мирончик. – Минск : БГУИР, 2015. – Ч. 1. Фотометрические и колориметрические измерения. – 66 с.
- [3] Zhu, Z. Image Recoloring for Color Vision Deficiency Compensation: a Survey / Z. Zhu, X. Mao // The Visual Computer. – 2021. – Vol. 37. – P. 2999–3018.
- [4] Ribeiro, M. Recoloring Algorithms for Colorblind People: a Survey / M. Ribeiro, A. J. P. Gomes // ACM Comput. Surv. – 2019. – Vol. 52. – P. 1–37.

BIG DATA TECHNOLOGIES IN THE IMAGE RECOLORIZATION ALGORITHMS FOR PEOPLE WITH ANOMALOUS TRICHRMACY

V.V. Sinitsyna

*Postgraduate student of the Department
of Engineering Psychology and Ergonomics BSUIR,
Master of engineering*

*Department of Engineering Psychology and Ergonomics
Faculty of Computer Engineering
Belarusian State University of Informatics and Radioelectronics, Republic of Belarus
E-mail: vladasinitsina1@gmail.com*

Abstract. Image recolorization algorithms in accordance with their correct perception by people with anomalous trichromacy use both a change in the indicators of individual color characteristics and color conversion in accordance with some specific algorithm in different color spaces. Experiments on changing the color characteristics of images in different color spaces make it possible to determine the features of the spaces under consideration, which are applicable in the process of image recolorization for the purpose of their subsequent unhindered perception by anomalous trichromats. In addition, the presented recolorization methods and algorithms are relevant for processing big data in images that are transformed for correct perception by people with color vision anomalies.

Keywords: anomalous trichromacy, color spaces, hue, color saturation, color brightness, big data.

УДК 338.4

MARTECH КАК ТРЕНД ЦИФРОВОЙ ТРАНСФОРМАЦИИ БИЗНЕСА



Л.И. Архипова

Доцент кафедры экономики Белорусского государственного университета информатики и радиоэлектроники, кандидат экономических наук, доцент
l.arkhipova@gmail.com



Л.Ф. Медведева

Доцент кафедры управления экономическими системами Академии управления при Президенте Республики Беларусь, кандидат экономических наук, доцент
medvedevaL15@mail.ru

Л.Ф. Медведева

Окончила Белорусский государственный университет. Имеет практический опыт работы в радиоэлектронной отрасли экономики. Сфера научных интересов: синергетический подход в управлении, менеджмент устойчивого развития общества в условиях цифровой трансформации бизнеса.

Л.И. Архипова

Окончила Белорусский государственный университет и Академию управления при Президенте Республики Беларусь. Имеет многолетний опыт работы в реальном секторе экономики (НПО «Интеграл»). Работает в Белорусском государственном университете информатики и радиоэлектроники. Проводит научные исследования в области инновационных технологий и цифровой трансформации бизнеса.

Аннотация. В статье рассматривается MarTech (Marketing Technologies) как концепция современного бизнеса и дается характеристика ее компонентов. MarTech представляется, как сочетание и взаимопроникновение маркетинга, цифровых технологий, и менеджмента. Показано, что пересечение этих трех категорий входят в бизнес, как неопровержимый тренд, определяющий векторы развития маркетинга, продаж, и нового типа мышления. В исследовании отмечается, что концепция MarTech формирует сквозное взаимодействие информационных ресурсов, как актива, способного перевести маркетинг из категории центра затрат в центр прибыли.

Ключевые слова: цифровой бизнес, цифровые технологии, MarTech (маркетинговые технологии), автоматизация маркетинга, эффективность бизнеса.

Введение.

Термин «MarTech» был введен и популяризирован американским программистом и предпринимателем Скоттом Бринкером. Он описал MarTech, как технологии, которые определяют инновации в маркетинге и обусловлены изменением поведения потребителей, новыми каналами выхода на рынок, более фрагментированными базами потенциальных клиентов, нормативно-правовыми актами и технологическими инновациями. Объединив в единую концепцию «Marketing, Technology, Management», он описал новые возможности для маркетинга, которые не ограничиваются текущим технологическим развитием, а определяют будущее [1, 2]. В статье «2023 will be a chaotic year for Martech, yet the start of a massive wave of growth» он спрогнозировал начало мощной волны роста по типу S-образных кривых новых технологий за счет: искусственного интеллекта (AI); дополненной и виртуальной реальности (AR/VR); компонуемости бизнеса (Composable business). Некоторые специалисты дополняют этот список такими категориями, как «данные» и «персонализация» [1, 3-5].

Концепция MarTech подчеркивает эволюцию технологий и поддерживает высокую степень инновационности, обусловленную изменением ожиданий клиентов и, соответственно, развитием технологий под эти ожидания, что постоянно предоставляет бизнесу новые возможности (рисунок 1).

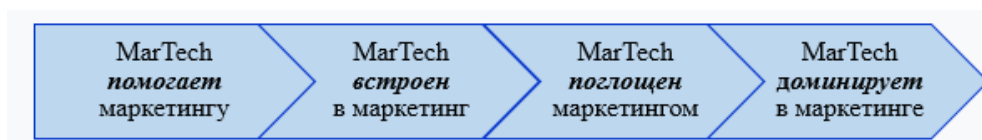


Рисунок 1. Эволюция MarTech и маркетинга (адаптировано) [2-4]

MarTech, как сочетание «маркетинга» и «технологии», фактически, относится к онлайн-инструментам и программному обеспечению, которые используются для охвата, привлечения, вовлечения и конверсионного преобразования клиентов. Эти инструменты образуют набор (stack) маркетинговых технологий, который помогает маркетологам более эффективно планировать, управлять и анализировать эффективность маркетинговых кампаний, оптимизируя процессы, которые ранее требовали значительного времени или усилий [6, 7].

Основная часть.

Инструменты маркетинговых технологий MarTech используются на всех этапах взаимодействия с клиентами, т.е. на протяжении всего *жизненного цикла клиента* и позволяют автоматизировать процессы *конверсионного пути клиента* от первой точки контакта с брендом или компанией до его логического завершения – доставки и послепродажного обслуживания.

Набор инструментов MarTech, который используют в маркетинговых процессах различные компании, может варьироваться от нескольких простых инструментов до сложного сочетания в зависимости от размера компании, отрасли и местоположения.

Для достижения целей маркетинга используются инструменты, обслуживающие несколько уровней управления [8]:

- корпоративные базы данных – данные о клиентах, продуктах и продажах, бизнес-аналитика;
- технологические платформы, CRM-системы – технологии автоматизации маркетинга, взаимодействия с социальными сетями, управления контентом;
- традиционные каналы маркетинговой системы – электронная почта, мобильная связь, телевидение, интернет-сайты, др.

Что вынуждает бизнес переключаться на новые технологии и инструменты? В первую очередь – это работа с данными, на основе которых должны приниматься стратегические и тактические решения. Необходимо отметить, что недостатка в данных сегодня нет, а дефицит, в большей степени, определяется инструментами, которые анализируют информацию, прогнозируют результаты и автоматизируют процессы. Именно это создает условия для разработки и внедрения в маркетинговую деятельность цифровых технологий и инструментов, поддерживая парадигму цифрового бизнеса (**Digital business**).

Цифровая трансформация бизнеса реализуется многоэтапным процессом, который включает в себя наращивание потенциала (решений и компетенций) по автоматизации основных процессов организации с учетом их приоритетности в решении стратегических задач: автоматизация, цифровизация, трансформация, цифровой бизнес (рисунок 2).



Рисунок 2. Цифровой бизнес как единая система, объединяющая процессы и структуры организации

Входы и выходы всех бизнес-процессов организации поддерживаются цифровыми технологиями и инструментами, которые синхронизируют внутренние и внешние взаимодействия организации, используя методологию «одного информационного окна» для принятия управленческих решений.

Следует отметить растущее значение данных в таких активностях маркетинга, как: реклама, продвижение, управляемый контент, клиентский опыт (CX), пользовательский опыт (UX), маркетинг взаимоотношений (relationship marketing), e-commerce, SMM, продажи, данные и аналитика, менеджмент.

В соответствии с принципами концепции MarTech, организация должна совмещать компетенции в области построения брендов, технологий и менеджмента. Однако концепция MarTech включает в себя не только управление командами или процессами, но и обмен данными между такими наборами компетенций, как: *автоматизация маркетинга; онлайн/офлайн продажи; программы лояльности; медиа-планирование; логистика; партнерские программы и др.*

Наилучшим образом этапы жизненного цикла клиента в омниканальной цифровой среде описывается моделью RACE (рисунок 3) [9].

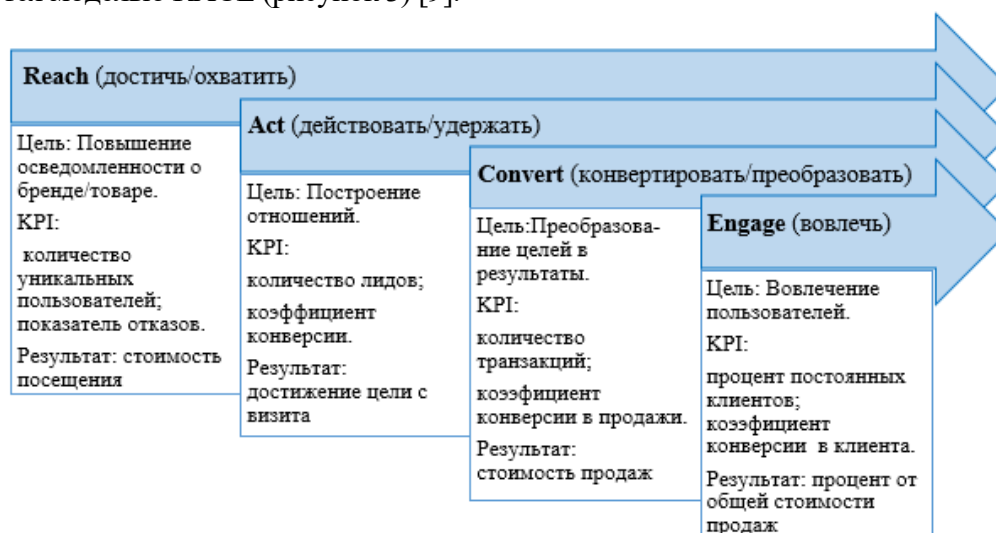


Рисунок 3. Модель RACE: структурированный подход к цифровому маркетингу

Модель структурирует маркетинговую деятельность на этапах жизненного клиента, ориентирует на действия с фокусированием на клиента (клиентоцентричность) и достижение КРІ. Фактически, это конверсионная модель, с точки зрения поведения клиента, которая предписывает использование определенных цифровых технологий и аналитических инструментов, оценивающих результат на каждом из этапов жизненного цикла клиента [9,10].

Сегодня MarTech максимально встраиваются в такие направления маркетинговой деятельности, как:

- Автоматизация процессов маркетинга (Marketing automation).
- Менеджмент взаимоотношений (CRM).
- Электронная коммерция (e-commerce Web & Aps).
- Социальные медиа (Social Media).
- Формирование клиентских баз данных (Customer Data).
- Данные и аналитика (Customer Data & Analytics).
- Проектный менеджмент (Project Collaboration Tools).

Набор технологий инструментов MarTech может быть использован практически во всех маркетинговых активностях (таблица 1), перекрывая этапы жизненного цикла клиента [7-10].

Таблица 1. Структуризация технологий и инструментов, используемых в маркетинге на этапах жизненного цикла клиента (по модели AARRR)

AARRR	Технологии и инструменты маркетинга на этапах жизненного цикла клиента				
Acquisition Привлечение	Social Tools Соц. медиа	Web-site, Blog Tools Инструменты управления web-сайтом и блоггингом	Advertising Tools Рекламные инструменты	Marketing Automation Автоматизация маркетинга	Analytical Platform Аналитическая платформа
Activation Активация	E-mail Marketing e-mail маркетинг				
Retention Удержание	Customer Engagement Tools Инструменты вовлечения клиентов	Data Management Tools Инструменты дата-менеджмента			
Referral Реферальность	Customer Relationship management (CRM) Маркетинг взаимоотношений				
Revenue Доход	Referral Marketing Реферальный маркетинг	Project Management & Communication. Проектный менеджмент и коммуникации			

Автоматизация маркетинга рассматривается, в первую очередь, как программная платформа, которая оптимизирует, автоматизирует и измеряет то, как происходит взаимодействие компании/бренда с потенциальными и действующими клиентами в условиях омниканальности. Набор технологий и инструментов MarTech позволяет:

1. Автоматизировать повторяющиеся задачи, такие как: электронная почта и рекламные кампании, распространение контента, обработка контактных форм, и планирование в социальных сетях. В результате маркетинговые операции становятся более оптимизированными, а конверсия потенциальных клиентов в реальных покупателей увеличивается.

2. Обеспечивает доступ к аналитике. Данные и маркетинг, в этом контексте, можно рассматривать как синонимы (Data Driven Marketing – маркетинг, управляемый данными). Актуальная информация о поведении клиентов дает больше возможностей для персонализации сообщений, точной настройки рекламных кампаний и подтверждения окупаемости инвестиций в маркетинг.

3. Поддерживает формирование системы Total Experience (TX) за счет интеграции UX, CX, EX, MX. При интеграции с CRM единая платформа гармонизирует подходы отделов продаж и маркетинга к клиентам.

Применение инструментов автоматизации (как правило, приложений) повышает эффективность мероприятий по увеличению продаж, создавая условия для: развития отношений с потенциальными клиентами, которые имеют низкую степень готовности к покупкам (не готовы покупать); своевременного информирования клиентов о новинках; сбора и анализа ключевых данных по рентабельности инвестиций в маркетинг (ROMI).

Основными компонентами автоматизации маркетинга являются:

1. *Маркетинговая база данных*: должна содержать подробные данные о клиентах (Empathy Map), их взаимодействиях, онлайн/офлайн поведении (Customer Journey Map) и др.

2. *Механизм интеграции маркетинга*: бизнес-среда, наполненная цифровыми технологиями и инструментами, для создания, управления и автоматизации маркетинговых процессов в онлайн/офлайн каналах.

3. *Механизм аналитики – диагностика и прогнозирования*: технологии и инструменты для тестирования, измерения и оптимизации бизнес-процессов, которые используются маркетингом

и продажами, с целью повышения рентабельности маркетинговых инвестиций и формирования потоков дохода.

4. *Набор приложений*: разработка и приобретение эффективных, интегрируемых и масштабируемых маркетинговых приложений, которые необходимы для достижения целей бизнеса.

В контексте сказанного, автоматизации подлежат [10,11]:

- 1) SEO-оптимизация web-сайта;
- 2) сбор информации и кластеризация пользователей;
- 3) CJM (путь клиента);
- 4) системы управления контентом (web-сайт, SMM и landing page);
- 5) рекламные кампании – создание, отслеживание и оценка эффективности (в том числе, используя programmatic-платформы);
- 6) реферальные программы и программы лояльности;
- 7) e-mail рассылки;
- 8) взаимодействие и обслуживание клиентов через систему чат-ботов;
- 9) сбор информации и исследование UX, CX, EX и MX – создание банков тотального опыта (Total experience) и др.

Отдел маркетинга с оптимально подобранным набором MarTech, как правило, оптимизирует работу с наиболее успешными маркетинговыми каналами и перенаправляет туда больше инвестиционных и других ресурсов.

Важно использовать описанные компоненты, как значимые механизмы взаимодействия с покупателями (внешние коммуникации) и обеспечивать согласованность подразделений организации, включая маркетинг и продажи (внутренние коммуникации).

В развитие структурного подхода к MarTech все инструменты можно разделить на шесть категорий: *реклама и продвижение; контент и опыт; социальные сети и отношения; коммерция и продажи; данные; управление.*

Конкретно каждая категория MarTech представлена технологиями и инструментами, которые комплексно используются ключевыми бизнес-процессами и активностями маркетинга [3-5]:

E-mail Marketing and Remarketing/E-mail маркетинг и ремаркетинг: несмотря на множество и популярность современных платформ цифрового маркетинга, электронная почта и ремаркетинг остаются важными аспектами в маркетинговой стратегии.

Social Media Management/Управление социальными сетями: планирование и анализ кампаний в социальных сетях включает несколько инструментов, каждый из которых имеет решающее значение для работы с клиентами.

Marketing Automation/Автоматизация маркетинга: маркетологам требуется настраивать и автоматизировать общение с существующими и потенциальными клиентами, поэтому автоматизация маркетинга должна быть включена в набор маркетинговых технологий.

Web Analytics Tools/Инструменты веб-аналитики: аналитика позволяет лучше понимать поведение посетителей веб-сайта, обеспечивать мониторинг трафика веб-сайта и оценивать эффективность рекламных кампаний.

Customer Relationship Management (CRM)/Управление взаимоотношениями с клиентами: взаимодействие с клиентами является критически важным элементом для маркетинга и продаж, поэтому инструменты управления взаимоотношениями с клиентами должны входить в набор MarTech.

Content Delivery Network (CDN)/Сеть доставки контента: гибкость в управлении контентом веб-сайта еще один важный элемент, обеспечивающий создание и распространение релевантного контента, а также оценка фактора виральности (контент является виральным, если он распространяется в сети самостоятельно – силами пользователей).

Преимущества использования MarTech для маркетинга можно сгруппировать по нескольким направлениям (рисунок 4) [12]:

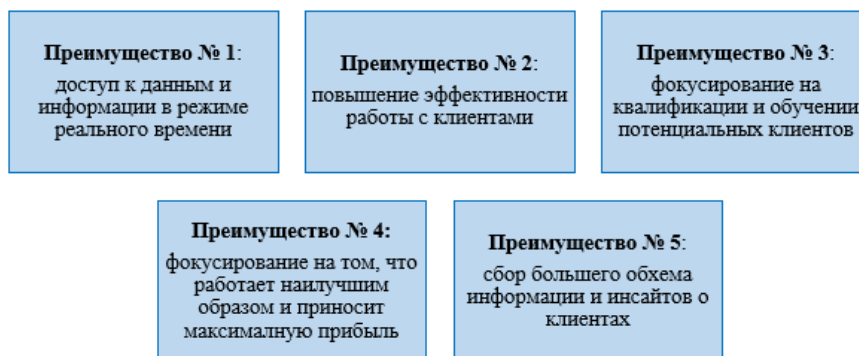


Рисунок 4. Преимущества применения MarTech в маркетинге

Быстрое развитие технологий «провоцирует» новые ожидания клиентов. В данных условиях **маркетинговые технологии** приобретают ключевое значение для организаций, которые фокусируются именно на ожиданиях клиентов и пытаются превзойти их, помогая организациям (в первую очередь, маркетингу) исключить «цифровые пробелы», связанные с изменением поведения покупателей и их предпочтений.

Использование подхода, сформулированного в концепции MarTech, предполагает [3,12]:

1. *Доступ к данным и аналитике в режиме реального времени.* Данные становятся ресурсом в принятии решений – повышение рентабельности инвестиций и эффективности маркетинговых стратегий (например, оптимизация маркетинга в социальных сетях, SEO, PPC, персонализация электронных рассылок и др.).

2. *Автоматизацию маркетинга с целью повышения эффективности бизнес-процессов.* Автоматизация рыночных процессов позволяет маркетологам взаимодействовать с потребителями более персонализированным способом, который предлагает реальную ценность и способствует сохранению «сильного присутствия» в умах потребителей. Автоматизация маркетинга также повышает производительность и увеличивает присутствие бренда в цифровой среде, что способствует масштабированию бизнеса.

3. *Повышение вовлеченности клиентов, партнеров и персонала с помощью цифровых инструментов управления.* Маркетинговые технологии соответствуют ожиданиям клиентов, автоматизируют и оптимизируют взаимодействие клиентов с брендами через платформы социальных сетей. Инструменты управления социальными сетями помогают планировать и публиковать контент с высокой вовлеченностью и охватом.

4. *Улучшение качества обслуживания.* Технологии Big Data и AI (программные решения) помогают маркетингу анализировать и обрабатывать обширные данные о клиентах, чтобы затем предоставить клиентам персонализированный опыт (CX, UX, TX).

5. *Повышение рентабельности инвестиций и производительности.* Маркетологи создают цифровой контент (изображения, видео, аудио и др.), нацеленный на определенный этап конверсионного пути покупателя (CJM). В данном случае важно создать легко доступную и бесшовную среду для омниканального взаимодействия с клиентами. С этой целью рекомендуется использовать:

- DAM (Digital Asset Management) – программное обеспечение для управления цифровыми активами, которое автоматизирует управление изображениями и другими нетекстовыми материалами;

- CDP (Customer Data Platform) – платформа, которая интегрирует данные о клиентах из различных источников, формируя единый взгляд на клиента (Single View of the Customer);

- Programmatic Advertising – программные рекламные платформы, которые работают по модели RTB (Real Time Biting) – самая популярная модель программных закупок, которая позволяет маркетологам запускать автоматизированные онлайн-кампании с predetermined рекламными атрибутами целевой аудитории: демография, интересы, или покупательские ожидания и др. С увеличением объема данных, доступных рекламодателям, бренды могут лучше понять свою целевую аудиторию и предоставлять более персонализированную и релевантную рекламу, точно контролируя рекламные кампании, данные и затраты.

6. *Более широкое использование технологий искусственного интеллекта (AI) и машинного обучения (ML).* Использование искусственного интеллекта и машинного обучения позволяет реализовать возможности таргетинга, оптимизации и прогнозирования. Алгоритмы AI используются также в автоматизации и оптимизации процесса покупки, позволяя брендам обращаться к потенциальным или действующим клиентам в соответствующее время и с релевантным сообщением.

Заключение.

Результаты исследования подтверждают, что маркетинговые технологии увеличивают эффективность и результативность маркетинга, а также позволяют устанавливать более тесную связь с целевой аудиторией, реализуя стратегии маркетинга взаимоотношений (CRM). Опираясь на MarTech, компании ставят аналитику в центр процесса принятия решений – они становятся более ориентированными на данные и результаты.

Набор технологий и инструментов MarTech улучшает бизнес-процессы работы с контентом, упрощают создание рекламных кампаний за счет автоматического сбора и анализа данных, а также обеспечивают оперативное отслеживание, персонализацию и автоматизацию конкретных задач маркетинговой деятельности, координируя усилия по совместному использованию ресурсов.

Выбор и внедрение подхода, основанного на концепции MarTech, рекомендуется реализовывать по следующему сценарию:

- Определить и описать «проблемные зоны» маркетинга и бизнеса. Здесь необходимо провести аудит, определить цифровую зрелость организации, а также ответить на ряд вопросов – какие технологии необходимо внедрить и как они смогут изменить бизнес.
- Определить и описать целевую аудиторию, используя инструменты Empathy Map.
- Сегментировать списки электронной почты по принципам технологии JTBD (Job To Be Done).
- Разработать Customer Journey Map для каждого целевого сегмента – путь клиента с учетом особенностей принятия решения о покупке.
- Выбрать программную платформу, которая согласуется и интегрируется с CRM-системой организации и приложениями.

Оптимальный набор технологий и инструментов MarTech, естественным образом способствует максимальной автоматизации процессов и более эффективному анализу данных для принятия бизнес-решений, извлекая выгоду из взаимодействий с потенциальными клиентами. Эффективность MarTech в данном случае определяется тем, что в программных продуктах активно используются принципы машинного обучения и аналитики больших данных. Таким образом, технологии и инструменты MarTech служат стратегическим целям организации и направлены на то, чтобы максимизировать эффективность и окупаемость инвестиций в маркетинговую деятельность и развитие бизнеса.

Успешно проделанная работа в направлении внедрения концепции MarTech, должна сместить акценты с маркетинга, как центра затрат, к созданию маркетинговых команд, работающих, как центры прибыли. Идея управления маркетингом, как центром прибыли

становится все более реальной с увеличением цифрового преобразования и технологической поддержки маркетинга. Это требует изменения мышления, новых компетенций и овладения цифровыми инструментами.

Список литературы

- [1] Marketing Technology Landscape Supergraphic (2019): Martech 5000 (actually 7,040) [Электронный ресурс]. – Режим доступа: [https://bethebean.com/blog/what-is-martech/#:~:text=\(Source%3A%20Marketing%20Technology%20Landscape%20by,%2C%20and%20location\(s\).](https://bethebean.com/blog/what-is-martech/#:~:text=(Source%3A%20Marketing%20Technology%20Landscape%20by,%2C%20and%20location(s).) – Дата доступа: 12.02.2023
- [2] Martech is marketing a wide spectrum of possibilities [Электронный ресурс]. – Режим доступа: <https://chiefmartec.com/2019/07/martech-marketing/> – Дата доступа: 12.02.2023
- [3] The state of martech in 2023 [Электронный ресурс]. – Режим доступа: <https://martech.org/the-state-of-martech-in-2023/> – Дата доступа: 12.02.2023
- [4] What is MarTech and what can it mean for your business? [Электронный ресурс]. – Режим доступа: <https://www.contentgrip.com/martech-meaning-how-it-benefit-business/> – Дата доступа: 10.01.2023
- [5] What is Marketing Technology and Martech? [Электронный ресурс]. – Режим доступа: <https://www.marketingevolution.com/marketing-essentials/marketing-technology-martech>. – Дата доступа: 12.12.2022
- [6] MarTech технологии маркетинга, которые придут к нам завтра [Электронный ресурс]. – Режим доступа: https://www.comagic.ru/martech/experience/marketing_technologies/
- [7] What is martech and marketing technology? <https://martech.org/what-is-martech/>. – Дата доступа: 18.02.2023
- [8] Комплексные решения сложных маркетинговых задач [Электронный ресурс]. – Режим доступа: <https://avada-media.ua/services/martech/> – Дата доступа: 22.02.2023
- [9] Модель RACE современное планирование цифровой маркетинговой деятельности [Электронный ресурс]. – Режим доступа: <https://dzen.ru/media/id/5c9df317f7498800b3d2b61b/model-race-sovremennoe-planirovanie-cifrovoi-marketingovoi-deiatelnosti-60584e2596537e64081663c7>. – Дата доступа: 16.01.2023
- [10] What is MarTech? [Электронный ресурс]. – Режим доступа: [https://bethebean.com/blog/what-is-martech/#:~:text=\(Source%3A%20Marketing%20Technology%20Landscape%20by,%2C%20and%20location\(s\).](https://bethebean.com/blog/what-is-martech/#:~:text=(Source%3A%20Marketing%20Technology%20Landscape%20by,%2C%20and%20location(s).) – Дата доступа: 16.02.2023
- [11] Что такое MarTech. Тренды цифрового маркетинга [Электронный ресурс]. – Режим доступа: <https://markway.ru/blog/cto-takoe-martech-trendy-cifrovogo-marketinga/> – Дата доступа: 24.02.2023
- [12] Top 5 Benefits of Implementing Marketing Technology in Your Business Strategy [Электронный ресурс]. – Режим доступа: <https://www.rohitprabhakar.com/2021/02/14/top-5-benefits-of-implementing-marketing-technology-in-your-business-strategy/> – Дата доступа: 10.01.2023

MARTECH AS A TREND OF BUSINESS DIGITAL TRANSFORMATION

L.F. MEDVEDEVA

*Associate professor, PhD, Academy of Management
under the President of the Republic of Belarus*

L.I. ARKHIPOVA

*Associate professor, PhD, Belarussian State
University of Informatics and Radioelectronics*

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

E-mail: l.arkhipova@gmail.com

Academy of Management under the President of the Republic of Belarus,

E-mail: medvedevaL15@mail.ru

Abstract. In this article, MarTech (Marketing Technologies) as a modern business concept is considering and main characteristics of its components have submitted. Stack of MarTech is defined as a combination and interpenetration of marketing, digital tools and management. It is shown that the intersection of this three category enter to the business as an irrefutable trend that will determine the direction of marketing, sales and new way of the thinking. In this Research confirmed that MarTech concept is forming the interaction of all IT tools and resources through the company as assets that will transform marketing from the cost generated to the profit-generated center.

Keywords: digital business,, digital technology, MarTech (marketing technologies), marketing automation, маркетинга, business efficiency.

УДК 338.24

ИНСТРУМЕНТАРИЙ АЛГОРИТМИЗАЦИИ ОТКРЫТЫХ И BIG DATA – ДАННЫХ В ЦЕЛЯХ РАЗВИТИЯ МЕЖДУНАРОДНОГО БИЗНЕСА КОМПАНИЙ



С.А. Затонский

*Аспирант ЮФУ, руководитель
отдела маркетинга BODIUM,
stanislav.zatonsky@gmail.com*

С.А. Затонский

Окончил Южный Федеральный Университет (ЮФУ). Аспирант ЮФУ, факультет Менеджмента. Работает в инновационном пространстве BODIUM в должности руководителя отдела маркетинга. Проводит научные исследования на тему цифровой трансформации международного бизнеса организации.

Аннотация. Благодаря интернету и распространению компьютерных технологий, международный бизнес может анализировать открытые и большие данные в целях наиболее эффективной адаптации полученной информации, влияющей на получение коммерческой прибыли. В данном исследовании описаны преимущества использования данных в целях развития бизнеса. Проведен пошаговый анализ алгоритма работы с данными. Определены ключевые этапы работы по исследованию данных: поиск и сбор информации, основная очистка и обработка данных, приведение данных к одному виду кодировки. Определены наиболее популярные инструменты, позволяющие работать с данными на всех представленных этапах.

Ключевые слова: международный бизнес, большие данные, массивы данных, исследование данных, алгоритмизация данных

Введение.

Сегодня исследование данных является востребованным направлением по всему миру. Международный бизнес может использовать большие данные в повседневной работе для множества целей, таких как прогнозирование тенденций и рыночной конъюнктуры, оптимизация процессов производства и логистики, управление рисками, а также улучшения обслуживания клиентов и модернизации маркетинговых стратегий.

Аналитика больших данных имеет потенциал преобразовать способ взаимодействия компаний с международным бизнесом, позволяя им лучше понимать и реагировать на сложную и динамичную глобальную среду [1]. По всему миру большие данные все чаще используются в исследованиях международного бизнеса для изучения широкого спектра тем, включая глобальные цепочки добавленной стоимости, интернационализацию и кросс-культурный менеджмент.

Одной из главных польз использования Big Data для международного бизнеса является возможность получения более точной и своевременной информации о клиентах и рынке [2]. Например, анализ данных социальных сетей может помочь компаниям лучше понимать потребности и предпочтения своей ЦА, а анализ данных покупок и поведения клиентов может помочь компаниям оптимизировать свои продукты и услуги. Еще одним преимуществом использования Big Data для международного бизнеса является возможность оптимизации бизнес-процессов и сокращения расходов [3]. Например, анализ данных о производственных процессах может помочь компаниям выявить узкие места и улучшить производительность, а анализ данных о расходах может помочь компаниям сократить затраты на определенные процессы (табл. 1).

Таблица 1. Соотношение задач МБ и типов больших данных

Задачи	Типы больших данных
Прогнозирование спроса	Исторические данные о продажах, данные о клиентах, данные о производственных циклах
Улучшение маркетинговых кампаний	Данные о клиентах, социальные медиа-данные, данные о транзакциях
Оптимизация логистики	Данные о транспортировке, данные о трафике, данные GPS
Анализ рынка	Данные об экономике, данные о конкурентной среде, данные о клиентах
Улучшение качества продукции	Данные о производственных процессах, данные о качестве продукции, данные о сбоях и проблемах
Управление рисками	Данные о рисках, данные о финансовых рынках, данные о клиентах
Исследование новых рынков	Данные о рынке, данные о конкурентной среде, данные о клиентах
Развитие инноваций	Данные о рынке, данные о клиентах, данные о технологиях и инновациях

Источник: составлено автором по материалам собственного исследования

Компании малого и среднего производственного бизнеса могут встречать проблемы в работе с большими данными, так как у них может не быть достаточных ресурсов для разработки и реализации программ и технологий, способных обрабатывать и анализировать большие объемы данных.

Вот некоторые из задач, которые могут возникнуть в этой области и которые менеджмент может решить с помощью больших данных:

1. Определение целей и преимуществ использования больших данных. Менеджеры могут помочь определить, какие данные могут быть наиболее полезными для компании и как они могут использоваться для увеличения эффективности производственных процессов, оптимизации затрат и улучшения качества продукции.

2. Разработка уникальных инструментов и методов для работы с данными в целях компании. Например, использование статистических методов, машинного обучения, искусственного интеллекта и других технологий [4].

3. Оценка стоимости внедрения систем обработки данных: в рамках развития международного бизнеса могут быть внедрены системы обработки данных, а также оценена потенциальная коммерческая выгода от использования этих систем.

4. Анализ данных и выявление тенденций: топ-менеджмент корпораций, используя методы работы с большими данными, может проанализировать информации, собранную компанией, и выявить определенные тенденции и закономерности, которые могут помочь принимать более эффективные решения по управлению бизнесом.

5. Разработка стратегии безопасности данных. Крупный международный бизнес может быть заинтересован в разработке собственной стратегии «безопасности данных», которая будет обеспечивать сохранность и конфиденциальность данных компании, включая защиту от внешних угроз и внутренних нарушений безопасности.

Тем не менее, реализация потенциала использования больших данных в международном бизнесе влечет за собой ряд технических, организационных и даже иногда этических вызовов для компаний по всему миру.

Следует отметить, что поскольку суть использования информации, основанной на больших и открытых данных, заключается в познании и поиске, то изначально зафиксировать единый статичный план первоначальной идеи исследования, необходимой для бизнеса, как правило, невозможно. Однако, какой бы ни была сумбурной динамика работы с большими и открытыми данными, существуют общепринятые алгоритмы работы, подходящие для различных бизнес-целей. Так, благодаря команде дата-исследователей из ‘The Guardian’, основавшей значимый для развития исследования данных ‘Datablog’, на данный момент установлен общий план выполнения работ по исследованию данных (рисунок 1).



Рисунок 1. Визуализированный производственный процесс Guardian Datablog, 2019 [5]

Вне зависимости от целей и сектора международного бизнеса, первым этапом плана по работе с данными всегда является их поиск и сбор. Как правило, источники дифференцированы. Ими могут являться научные, медицинские или государственные базы данных, а также поисковых сетей, новостные сводки, социологические опросы, открытая информация и разнообразные исследования. Так, на данном этапе необходимо создать основную базу данных

с актуальной для бизнеса информацией, напрямую связанный с определенной актуальной бизнес-задачей.

Следующий шаг корректной работы с данными – анализ полученной базы на предмет ее корректности, целостности и упорядочиваемости. На этом этапе, как правило, проводится первичная очистка характеризующаяся отсортировкой и удалением «битых» файлов с данными. В последствии необходимо провести анализ базы на предмет информационной заполненности. Результат будет влияет на последующее базы в последствии и другими массивами данных. Вдбоавок к логическому анализу, исследователю необходимо также определить потенциально возможные взаимосвязи и корреляции данных в рамках как уже собранной базы данных, так и при ее расширении в дальнейшем.

Далее необходимо перейти к этапу работы с данными, который характеризуется очисткой основных данных и их обработкой. На данном этапе необходимо выявить и удалить лишние элементы таблиц, таких как дубликаты, ненужные столбцы, объединенные ячейки, посторонние символы и т.д. Данная работа должна выполняться постепенно в момент построения базы. Если информация в базе данных является зашифрованной, то исследователю также потребуется полная дешифровка и чистка данных. В случае необходимости, если необходимо отдельно восстанавливать и расшифровывать значения каждой ячейки в таблице базы, на данном этапе может быть привлечен отдельный дешифровщик-специалист.

При составлении общей таблицы из исследуемых источников, для объединения данных и продолжения дальнейшей работы, необходимо кодировать данные к одному виду. Так, следует провести работу и с исходными форматами данных – в случае, если данные не являются машиночитаемыми (допустим, изображениями или PDF-форматам), то необходимо использовать специализированное программное обеспечение (ПО). Как правило, подходящее под данную задачу ПО не рассчитано на работу с данными в формате таблицы – как результат, это может привести к появлению недопустимых ошибок. Следовательно, вероятно, что полученные данные в таком случае придется переносить вручную или с помощью сторонних программ, использующих технологии искусственного интеллекта.

Таким образом, алгоритм работы с данными состоит из четко структурированных пошаговых этапов. Менеджмент компании, анализируя big data – данные, обязан следовать данному алгоритму, т.к., в противном случае, любая полученная информация может быть ошибочный и привести компанию к потерям и убыткам. Несмотря на близость работы с данными к деятельности программистов, в международном бизнесе навыки исследователя-данных играют первостепенную роль, ведь в центре задачи всегда стоит определенная бизнес-цель, а не просто ряд цифр.

Сбор и первичная обработка данных.

Наиболее важным и первым этапом в работе над дата-исследованием является поиск и сбор данных. Среди работающих с данными журналистами принято считать, что в зависимости от уровня технической подготовки и профессиональных навыков, исследователь имеет возможность выбрать наиболее удобный инструмент для сбора данных [6].

Традиционно первичный поиск осуществляется через наиболее популярные системы поиска, такие как Yandex или Google. Такие современные гибкие системы поиска позволяют дата-исследователю, используя поиск по ключевым фразам, цитатам, датам, ограничивая поиск по определенным ресурсам или форматам данных, получить возможность найти не только необходимые базы данных, но и другую важную информацию по изучаемой теме.

Сегодня многие государства поддержали концепцию открытых данных и свободы распространения официальных данных тех или иных структур. К сожалению, правительственные и научные структуры, в силу отсутствия дополнительных человеческих ресурсов или постановления других приоритетных задач, не всегда имеют возможность контролировать качества загружаемых в открытые базы данных, а также степень адаптации старых, плохо обработанных баз данных, под современное программное обеспечение

компьютеров. Таким образом, исследователь, работающий с базами данных с открытых источников такого рода, часто может столкнуться с нечитабельным материалом, непригодным для дальнейшего компьютерного анализа.

Помимо официальных государственных сервисов, таких как вышеупомянутые порталы открытых данных разных стран как `data.gov.ru`, `data.gov` или, допустим, `data.gov.uk`, исследователь может также использовать и агрегированные международные системы, занимающиеся составлением тематических баз данных. Ими могут являться следующие порталы: Open Data Network, Google Public Data Explorer, The Data Hub, Datamarket, The World Factbook, BuzzData, UNData и множество других. Данные системы открывают доступ к мировой коллекции данных из многообразия разных интернет-источников. Следует учитывать, что данные, попадающие в эти системы, могут поступать не только из официальных источников, но также и от интернет-пользователей: при работе с такими данным исследователю необходимо с особым вниманием проверять загруженные базы данных.

Не стоит забывать и о наиболее классическом и распространенном методе получения информации, котором часто пользуются исследователи: официальном запросе или обращении к экспертам [7]. Многие организации и ведомства не загружают все имеющиеся данные в собственные базы данных, однако по запросу могут предоставить необходимую для исследователя информацию. В таком случае, при обращении к организациям и ведомствам, исследователь данных обязан определить заранее какую именно информацию он ставит цель получить, так как подобная специализированная информация от «экспертов» может быть ограничена в распространении законом. Распространено мнение, что в случаях, если информация может представлять собой государственную или коммерческую тайну, исследователю надо до публикации проконсультироваться с юристом.

В дополнение к вышеописанным классическим способам получения необходимой дата-информации, исследователь может привести поиск по темам и запросам на узкоспециализированных форумах. Сегодня интернет объединяет большое количество исследователь данных, которые, являясь активными пользователями сети, создают специализированные форумы и порталы. Здесь диджитал-исследователи могут делиться опытом, получать рекомендации по поиску и обработке данных, а также запрашивать необходимую информацию у своих коллег. Примерами могут являться специализированные порталы Get The Data, Data Mos, AZSecure-daa, посвященные журналистики данных сабредита портала Reddit и менее узкоспециализированный сервис Quora. При желании, в целях получение постоянной информации о новых доступных базах данных, дата-исследователь может стать членом интернет-сообществ, связанных с открытыми и большими данными, а также AI-тенденциями в исследовании и обработки таких данных.

В случае, если исследователь не может использовать описанные выше способы, то для получения необходимой информации существуют и иные эффективные методы. Один из таких потенциально возможных, но более сложных методов, является получение API-доступа к данным [8].

Интерфейс прикладного программирования (application programming interface) является набором методов, процедур и функций, предоставляемым сервисом для использования профессиональными разработчиками и сторонним ПО. Так, API классифицируются по видам операционных и аутентификационных систем, а также звуковых или графических интерфейсов.

Возможно использование интерфейсов прикладного программирования, таких как: OpenAI, GDI, Motif, Pam Qt, Amiga ROM, Kernel, Cocoa, OS/2 API, SDL, X11 и Zune. Используя интерфейсы прикладного программирования, исследователи могут взаимодействовать с другими программами и получать доступ к непубличным данным социальных медиа, библиотек и сервисов. Безусловно, обработка данных, полученных таким образом, запрещена, так как изначально их стороннее использование третьими лицами нарушает законодательство.

Метод получения данных напрямую с интернет-страницы может быть использован, если для доступа к закрытым данным на сайте не предусмотрен API. Как правило, использование данного метода требует дополнительных технических навыков: более глубокий знаний в программировании, а также базовых знаний CSS и HTML.

Данный способ используется в случае, если информация встроена в страницу и ее невозможно скопировать. Безусловно, исследователь может попробовать провести провести аналогичный сбор информации: тем не менее, если информация распределена по различным страницам или даже отдельным сайтам, то данный сбор информации требует, как правило, большого количества времени. Таким образом, специалисты с навыками программирования используют готовые утилиты или создают собственные программы, автоматизирующие процесс сбора данных.

Так, сегодня в распоряжении исследователя-данных могут находиться инструменты, облегчающие работу по извлечению больших объемов данных с веб-страниц. Например, ими могут являться браузерные расширения, профессиональные веб-сервисы и отдельное программное обеспечение. Например, для поиска в браузере исследователь может использовать «Альтернативный поиск Google», «Контекстный поиск», «Chrome Scraper extension» для Chrome, DownThemAll для Firefox и расширение FireBug доступное сегодня для большинства браузеров: Firefox, Chrome, Internet Explorer и Safari. Интернет-сервис QuickCod являющийся примером сервиса для коллективной разработки программного, необходим для анализа и извлечения публичных данных. Благодаря QuickCod, исследователю открывается доступ к работе с кодами программ, написанных на языках программирования, таких как Ruby, Python и PHP.

Процесс поиска и извлечения данных, как и любые другие практические способы и методы работы, имеет свои ограничения. Ими могут являться: неправильно структурированный код веб-страниц, нарушающий алгоритмы поиска, аутентификационные барьеры, запрещающие автоматический доступ, а также блокировка широкого доступа со стороны администрации сервера.

Выше было упомянуто, что исследователю, помимо технических ограничений, никогда не стоит забывать и об ограничениях правового рода. Сегодня извлечение данных из веб-страниц-первоисточников не является официальным способом получения информации. Таким образом, данные, полученные в результате таких методов, имеют ограничения как на использование, так и на публикацию. В случае дальнейшего распространения данных, извлечение которых даже могло быть и не запрещено официально, могут возникнуть юридические конфликты.

Так, независимо от используемого метода получения данных, основной бизнес-целью является получение эффективной машиночитаемой базы. Основная задача, которую необходимо выполнить – собрать данные и затем подготовить их к дальнейшей компьютерной обработке. Среди исследователей данных наиболее удобными и популярными форматами для этого являются общепринятый Excel, а также XML, CSV и JSON [9].

Вторичная очистка, обработка и формирование базы данных исследователя.

Следующий этап обработки является упорядоченной последовательностью действий исследователя, направленных на выявление и исправление некорректных элементов баз данных. Цель данного этапа состоит в том, чтобы привести информацию к структурированному и упорядоченному состоянию, что является важным шагом в проведении последующего анализа. Сегодня исследователи выделяют три предварительных этапа обработки данных: получение входных данных, непосредственный процесс обработки и анализ результатов. Для базовой очистки данных автор может использовать бесплатные инструменты, широко представленные в сети Интернет. Ими являются: Potter's Wheel ABC, Wranglerm, OpenRefine и другие. Однако наиболее популярным и удобным на сегодняшний день инструментом является именно специализированная программа OpenRefine, являющаяся особым автономным настольным приложением. Программа OpenRefine имеет так называемый «открытый исходный код», что позволяет исследователю очищать и преобразовать данные в другие форматы, а также работать

с таблицами (например, программа может совершать автоматическое исправление ячеек и данных с ошибками), анализировать распределение значений по всему набору данных т.д. (рисунок 2) Исследователь имеет возможность определять критерии фильтрации. Особенным отличием от прочих электронных таблиц является возможность выполнения большинства операций над всеми видимыми строками.

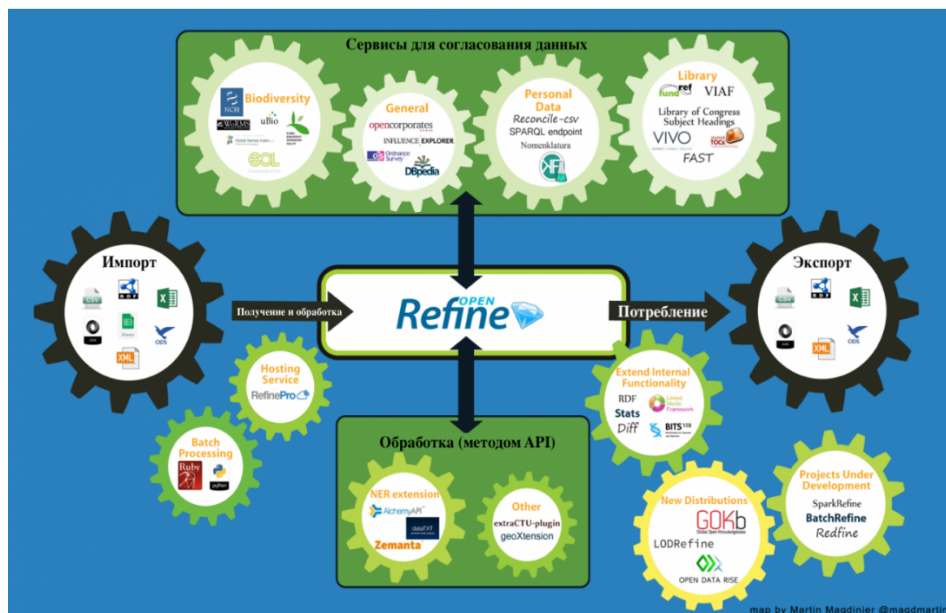


Рисунок 2. Процесс работы с данными в программе OpenRefine, 2019 г. [10]

Для работы с этим приложением исследователю хватит классических навыков работы с таблицами. Тем не менее, следует отметить, что программа OpenRefine позволяет выполнять и профессиональные задачи - данный инструмент повсеместно используют в организациях, чья работа связана с открытыми данными официальных государственных порталов.

Имея определённые навыки программирования (например, владение языками программирования Python и R), исследователь, как и в случае с процессом извлечения данных, его обработки и очистки, может автоматизировать процесс работы и в OpenRefine. Тем не менее, несмотря на то, что данный вариант работы является наиболее сложным, он позволяет использовать более широкий спектр возможностей для быстрой автоматической очистки.

Несмотря на системную необходимость обработки данных, следует исходить и из особенностей организаций международного бизнеса: требования и форматы оформления и структурирования данных могут кардинально отличаться друг от друга. Тем не менее, общепринятых требований и официальных стандартов к оформлению открытых данных на данный момент не существует. Следовательно, каждая организация формирует их на свое усмотрение в соответствии с нуждами бизнеса и его целями. Однако, это может приводить к возникновению неточностей и нестыковок в базах данных, что затрудняет выявление в таблице закономерностей между полученными данными.

В целях решения этой задачи, исследователь данных может изучить полученный материал и, устранив очевидные ошибки, провести сортировку данные. При работе с базой, созданной путем слияния других баз, потребуется процедура форматирования, позволяющая привести данные к одному формату. Кроме того, иногда при работе с данными могут быть использованы различные методы кодирования, затрудняющие обработку данных. В данном случае задачу решает создание специализированных словарей с перечнем ключевых значений, используемых

в таблицах. В таких словарях описаны все виды кодировок, позволяющих, как правило, избегать вольного трактования в работе с той или иной базой данных.

Таким образом, при глубокой обработке, очистке и формировании рабочей базы данных, менеджмент должен учитывать все перечисленные выше факторы, использовать специализированные программы и постоянно анализировать полученные результаты под цели компании. Именно на этапе формирования логической структуры и происходит формирование базы, а, следовательно, неверная интерпретация данных может привести к ложным выводам и доказательствам, что полностью подрывает ценность использования больших данных в международном бизнесе.

Результаты.

В целях развития международного бизнеса компаний, менеджмент может использовать большие данные как для оптимизации своей деятельности, так и для получения новой уникальной информации, позволяющей принимать решения на основе данных и доказательств.

Основными сервисами и инструментами для работы с данными являются:

- На этапе сбора и первичной обработки: государственные порталы открытых данных data.gov., порталы The Data Hub, Open Data Network, Google Public Data Explorer, Datamarket, а также метод получения доступа к данным через API через такие сервисы как miga ROM, Kemel, Cooa, OS/2 API, OpenAI, GDI, Motif, Pam Qt, SDL и др.

- На этапе глубокой обработки, отчистки и формирования рабочей базы исследователь данных может прибегнуть к использованию общедоступных в сети Интернет инструментов, таких как Potter's Wheel ABC, Wranglerm, OpenRefine или, имея определённые навыки программирования, модернизировать данные инструменты методом автоматизации процессов и путем определения критериев фильтрации данных.

Заключение.

Аналитика больших данных может предоставить компаниям ценные инсайты в международные рынки, позволяя им выявлять паттерны, прогнозировать спрос и предсказывать тенденции. Однако использование больших данных в международном бизнесе также создает значительные вызовы, связанные с навыками работы с данными, а также их конфиденциальностью, безопасностью и соответствием регулятивным требованиям. Предложенный инструментарий алгоритмизации открытых и big data – данных позволяет международному бизнесу наиболее эффективно получить новую информацию, в конечном счете влияющую на управление компаний на всех этапах: от анализа ЦА для получения уникального конкурентного преимущества и до управления рисками, логистикой, качества продукции и интернационализацией бизнеса.

Список литературы

- [1] A Systematic Review of Big Data: Research Approaches and Future Prospects / C. Cobanoglu, A. Terrah, M.-J. Hsu, V.D. Corte, G.D. Gaudio // Journal of Smart Tourism, 2(1), 2022, p. 21–31.
- [2] The Role of Big Data in International Business Strategy / M.H. Jensen // Twenty-Seventh European Conference on Information Systems, Stockholm-Uppsala, Sweden, 2019, p. 5-11.
- [3] Аксенова О. Н. Журналистика данных: проблемы и перспективы // Научный вестник Воронежского государственного архитектурно-строительного университета. Серия: Социально-гуманитарные науки No.3, 2015 г, с. 41-44.
- [4] Using machine learning to create and capture value in the business models of small and medium-sized enterprises / R.-C. Climent, D. Hafnor, M. W. Staniewsk // International Journal of Information Management, Swansea, United Kingdom, 2023, p.5-19.
- [5] Big Data (2023) // The Guardian URL: (<https://www.theguardian.com/data>) Дата обращения 12.03.2023
- [6] Арбатская Е.О. Открытые данные как ресурс региональной журналистики // Вестник Челябинского государственного университета, No: 5 (360) - Челябинск - 2015, г с. 52-58
- [7] Василика М.А. Основы теории коммуникации: учебник // М.: Гардарики, 2003 г., с. 610-615
- [8] Видовский Л.А., Янаева М.В., Мурлин А.Г, Мурлина В.А., Гвозденко А.А. Анализ возможности использования технологии обработки больших данных в системах для территориально-распределенных комплексов // Научный журнал КубГАУ - Scientific Journal of KubSAU. 2017. No132, с.1-11

[9] Нил К., Шатт Р. Data Science. Инсайдерская информация для новичков. Включая язык R // Издательский дом «Питер», Санкт-Петербург, 2018, с. 273-290

[10] Mapping OpenRefine Ecosystem / M. Magdinier // Open Refine Blog URL: (<https://openrefine.org/blog/2015/01/26/Mapping-OpenRefine-ecosystem>), 2015. Дата обращения: 15.03.2023.

ALGORITHMIZATION OF OPEN AND BIG DATA FOR INTERNATIONAL BUSINESS DEVELOPMENT

S.A. Zatonkii

*Postgraduate student at South Federal University, Head of
BODIUM Marketing Department*

*BODIUM, Marcel LLC, Rostov-on-Don, Russia
Southern Federal University, Rostov-on-Don, Russia
E-mail: stanislav.zatonsky@gmail.com*

Abstract. International business can analyze open and big data in order to effectively adapt the information that affects commercial profits. This study describes the benefits of using data for business development. A step-by-step analysis of the algorithm for working with data is conducted. The key stages of data research process are defined: searching and collecting information, basic data cleaning and processing, as well as reducing data to one type of coding. The most popular tools to work with data at all presented stages are identified.

Keywords: international business, big data, data arrays, data research, data algorithmization

УДК 004.021:004.75

BIG DATA И ADVANCED ANALYTICS В МЕДИЦИНЕ



А. А. Бакурина

Студент инженерно-экономического
факультета БГУИР
hannabakuryna@gmail.com



О.Н. Шкор

Старший преподаватель кафедры экономики
БГУИР
shkor@bsuir.by

А. А. Бакурина

Родилась в 2002 году в Жлобине. В 2020 году закончила ГУО «Средняя школа №12 в г. Жлобине». В этом же году поступила в УО «БГУИР», была зачислена на бюджетную форму обучения по специальности «Электронный маркетинг» инженерно-экономического факультета.

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. Концепция Big Data популярна в различных областях. Целью этой работы было суммировать особенности, применение, подходы к анализу Больших данных в здравоохранении. В статье рассматриваются основные сферы внедрения, применения и аналитики Big Data в медицине с опорой на последние достижения науки в этих сферах.

Преимущества Big Data в медицине рассмотрены со стороны как организаций здравоохранения и пациентов. С точки зрения пациента, применение анализа Больших данных может привести к улучшению лечения и снижению затрат. Кроме пациентов, правительство, больницы и исследовательские учреждения также могут получить преимущества от Больших данных.

Ключевые слова: Big Data в медицине, персонализированная медицина, общественное здравоохранение, аналитика данных в здравоохранении.

Введение

Современная медицина с каждым годом становится все более зависимой от данных. В связи с этим, применение BIG DATA и Advanced Analytics в медицине становится все более популярным. BIG DATA представляет собой большие объемы данных, которые могут быть использованы для определения путей улучшения здравоохранения и получения новых знаний о заболеваниях и лечении. Advanced analytics, в свою очередь, позволяет извлекать информацию из этих данных и использовать ее для принятия более точных и своевременных решений в медицинских вопросах.

Актуальность.

Сегодня технологии BIG DATA и Advanced Analytics стали широко применяться в медицине, где большие объемы данных могут дать более точные диагнозы, прогнозировать развитие заболеваний и улучшать качество лечения.

Использование больших данных и расширенной аналитики в медицине может иметь огромный потенциал для снижения затрат на лечение, увеличения эффективности медицинского ухода, улучшения точности диагностики и прогнозирования результатов лечения. Сейчас медицинские учреждения, исследовательские центры и производители медицинских изделий все

чаще используют технологии больших данных и расширенной аналитики для повышения эффективности работы и улучшения результатов в медицинской практике.

Сферы применения BIG DATA и Advanced Analytics в медицине. Одним из наиболее обещающих направлений использования BIG DATA и advanced analytics в медицине является **персонализированная медицина**. С помощью анализа больших объемов данных, таких как генетические данные, данные о здоровье пациента и данные о лекарствах, медицинские специалисты могут определить наиболее эффективное лечение для каждого конкретного пациента. Это позволяет улучшить результаты лечения и снизить риски для пациента [1].

В персонализированной медицине большие данные используются для создания профиля здоровья пациента, который включает данные о генетической предрасположенности, медицинской истории, результаты тестов, информацию о образе жизни и многие другие факторы. Этот профиль здоровья может быть использован для принятия индивидуальных решений по лечению, которые учитывают особенности конкретного пациента.

Например, компания IBM Watson Health разработала систему Watson for Oncology [2], которая использует анализ данных для поддержки онкологов в принятии решений о лечении рака. Система базируется на анализе большого количества данных о заболевании, лечении и результатах, и предлагает онкологам индивидуальные планы лечения для каждого пациента. Исследования показали, что система Watson for Oncology улучшает качество лечения и увеличивает выживаемость пациентов. Одним из примеров использования системы Watson for Oncology является работа с клиникой Manipal Hospitals в Индии. В этой клинике система используется для анализа медицинских данных пациентов и рекомендации лечения онкологических заболеваний. Watson for Oncology использует нейронные сети и алгоритмы машинного обучения для анализа больших объемов медицинских данных, таких как медицинские истории, результаты тестов и изображений, чтобы предоставить персонализированные рекомендации по лечению.

Big Data в общественном здравоохранении фокусируется на физиологических данных пользователей, которые часто собираются с за счет портативного оборудования, такого как электрокардиограмма, ежедневно носимые девайсы, устройства записи о здоровье, спорте и питании. К таким устройствам можно отнести, например, смартфоны с приложениями от третьих сторон (HealthKit от Apple, Google Fit от Google и S Health от Samsung), Android-часы и Google Glasses), умные фитнес-браслеты от Xiaomi. Только в магазине приложений Apple iTunes доступно более 40 000 приложений в области здравоохранения [1].

Жизненно важные показатели, собираемые устройствами, включают температуру, пульс, частоту дыхания и кровяное давление. Эти показатели являются четырьмя наиболее важными показателями функций тела. Носимое устройство в общественном здравоохранении относится к оборудованию, которое записывает детали о образе жизни и жизненно важных показателях людей, благодаря чему врачи могут получить помощь в лечении и диагностике у пациентов. Поскольку люди стали более озабочены своим здоровьем на повседневной основе, записи о ежедневном здоровье и поведении, показателях и симптомах пациентов стали играть важную роль в сборе данных. Кроме того, данные о спорте и питании людей также вносят значительный вклад в Big Data в общественном здравоохранении и поведении.

Big Data в медицинских исследованиях и экспериментах. Эта часть Big Data сфокусирована в основном на молекулярной биологии, наборе данных о человеческом организме, клинических испытаниях, биологических образцах, последовательностях генов, а также на клинических и медицинских исследованиях лабораторных тестов.

Молекулярная биология, важная часть как биологических, так и медицинских экспериментов, фокусируется на взаимодействии и регуляции биологических процессов внутри клеток, таких как взаимодействия между ДНК, РНК и белками, а также биосинтез. Основными техниками молекулярной биологии являются молекулярный клонирование, полимеразная цепная реакция (ПЦР), блотирование и пробирование макромолекул, микрочипы и т.д [3].

Наборы данных о человеческом организме включают образцы клеток, тканей и органов в человеческом организме, а также снимки человеческого тела в рамках проекта видимого человека, который используется для визуализации анатомии человеческого тела в поддержку медицинских деятельностей. Аналогично наборам данных о человеческом организме, биологические образцы лаборатории также получают при взятии образцов из человеческого организма и хранятся в биобанках. В случае создания нового типа лекарств, новых вакцин или новых медицинских устройств, перед их использованием должны быть проведены клинические исследования. Собранные данные и возможность проанализировать их помогает выявить и подтвердить взаимосвязи в ходе проведения разработки и тестирования новых продуктов. Например, при разработке новых лекарств для лечения рака, компания Roche использовала данные об экспрессии генов, полученные из биопсий опухолей пациентов, чтобы разработать индивидуализированные решения для лечения рака груди. Эти данные использовались для определения того, какие лекарства могут быть наиболее эффективными в конкретном случае, и для предсказания побочных эффектов этих лекарств.

Аналитика больших данных в генной инженерии. Big data является незаменимым инструментом в области генной инженерии и обеспечивает значительный прогресс в науке и медицине. При использовании больших данных находят новые возможности для идентификации генетических заболеваний. Генетические заболевания крови доказали свою эффективность для практики генной терапии, так как в аутологичных гемопоэтических стволовых клетках (ГСК) действие генной терапии может изменить причинный ген и внести необходимые изменения в гемопоэтическую систему. Например, болезнь Стоунта возникает из-за мутации глутамина в валин в нижней доле бета-глобина гемоглобина, что приводит к аномальному производству гемоглобина в организме. Универсальным методом лечения всех нарушений бета-глобина является повторное экспрессирование генов гамма-глобина. Такого результата позволяет достичь CRISPR / Cas9 технология для редактирования генома, что исправляет генетические мутации, связанные с наследственными заболеваниями. Так, CRISPR Therapeutics ведет клинические исследования по лечению бета-талассемии и других наследственных заболеваний крови с помощью трансплантации генно-редактированных гемопоэтических стволовых клеток (CTX001) [3].

Аналитика больших данных в профилактической медицине. Большие данные и здравоохранение имеют важное значение для пациентов с хроническими заболеваниями, которые не могут позволить себе расходы на госпитализацию. Optum Labs, американский центр исследований, собирает электронные медицинские записи 30 миллионов пациентов, чтобы создать базу данных для инструментов предиктивной аналитики, которые улучшат медицинскую помощь. Цель использования больших данных в профилактической медицине - помочь врачам принимать решения на основе данных за считанные секунды и улучшить помощь пациентам. Здравоохранительные учреждения могут обеспечить надежную профилактическую помощь и, в конечном итоге, снизить количество госпитализаций, понимая такие методы, как тип лекарств, симптомы и частота посещений пациентов. Это особенно полезно в случаях, когда у пациентов сложные медицинские истории и несколько заболеваний [4].

Это не только снизит риски, что приведет к меньшим затратам на стационарную медицинскую помощь, но также обеспечит наличие мест и ресурсов для тех, кто больше всего нуждается в них. Это является ясным примером того, как анализ больших данных в здравоохранении может улучшить и спасти жизни людей. Медицинские эксперты могут распознавать потенциальные преимущества и недостатки в исследованиях или процессах, используя комбинацию исторических, текущих и прогностических инструментов в качестве понятного сочетания методов визуализации данных.

Заключение

Большие данные обслуживают множество сфер в здравоохранении, включая управление рисками и болезнями, предотвращение самоповреждения, улучшение лечения рака,

телемедицину, улучшение стратегического планирования, разработку новых терапий, прогностический анализ, повышение вовлеченности пациентов и многие другие. Большие данные являются запуском для преобразования универсального здравоохранения на многих уровнях. Изменения в медицине, инфраструктуре и поддержке, которые обещает анализ данных в здравоохранении, обеспечивают решения, которые максимизируют клиническую помощь и увеличивают ценность медицинских профессионалов.

Список литературы

- [1] New Paradigm in Healthcare Industry Using Big Data Analytics [Электронный ресурс]. – Электронные данные. – Режим доступа: <https://iopscience.iop.org/article/10.1088/1757-899X/1099/1/012054/pdf>
- [2] A meta-analysis of Watson for Oncology in clinical application [Электронный ресурс]. – Электронные данные. – Режим доступа: <https://www.nature.com/articles/s41598-021-84973-5>
- [3] Big Data in Health Care: Applications and Challenges [Электронный ресурс]. – Электронные данные. – Режим доступа: <https://iopscience.iop.org/article/10.1088/1757-899X/1099/1/012054/pdf>
- [4] BIG DATA for Healthcare: A Survey [Электронный ресурс]. – Электронные данные. – Режим доступа: <https://ieeexplore.ieee.org/document/8585021>.

BIG DATA AND ADVANCED ANALYTICS IN HEALTHCARE

H. A. Bakuryna

Student of engineering and economics at the BSUIR

O.N. Shkor

Senior Lecturer at the Department of Economics at the BSUIR

Belarusian State University of Informatics and Radioelectronics

Minsk, Republic of Belarus

E-mail: hannabakuryna@gmail.com

E-mail: shkor@bsuir.by

Abstract. The concept of Big Data is popular in various fields. The purpose of this paper was to summarize the features, applications, and approaches to Big Data analysis in healthcare. The article discusses the main areas of implementation, application, and analytics of Big Data in medicine with reference to the latest scientific advances in these areas.

The advantages of Big Data in medicine are examined from the perspective of both healthcare organizations and patients. From the patient's perspective, the use of Big Data analysis can lead to improved treatment and cost reduction. In addition to patients, governments, hospitals, and research institutions can also benefit from Big Data.

Keywords: Big Data in healthcare, personalized medicine, public health, healthcare data analytics.

УДК 004.855.5

ВЫБОР ЧИСЛА КЛАСТЕРОВ В ЗАДАЧАХ КЛАСТЕРИЗАЦИИ С ИСПОЛЬЗОВАНИЕМ МЕТОДА СИЛУЭТОВ



В.П. Корячко

Заведующий кафедрой систем
автоматизированного проектирования РГРТУ
им. В.Ф. Уткина, доктор технических наук,
профессор
koryachko.v.p@rsreu.ru

В.П. Корячко

Окончил Рязанский радиотехнический институт. Область научных интересов связана с применением технологий машинного обучения в системах автоматизированного проектирования, разработкой методов представления знаний с использованием нечёткой логики и мягких вычислений.



В.И. Орешков

Доцент кафедры систем
автоматизированного проектирования
РГРТУ им. В.Ф. Уткина, кандидат
технических наук, доцент
vyacheslav.oreshkov@yandex.ru

В.И. Орешков

Окончил Рязанскую государственную радиотехническую академию. Область научных интересов связана с использованием технологий искусственного интеллекта в системах автоматизированного проектирования, машинным обучением, проектированием интеллектуальных информационных систем.

Аннотация. Рассмотрена проблема выбора числа кластеров в задачах кластеризации. Предложена методика выбора числа кластеров на основе метода кластерных силуэтов. В результате эксперимента на реальных данных показано, что лучшее число кластеров с точки зрения минимума внутрикластерных и максимума междукластерных расстояний не всегда соответствует числу естественных групп данных.

Ключевые слова: интеллектуальный анализ данных, машинное обучение, обучение без учителя, обучающие данные, кластер, кластеризация, коэффициент силуэта.

Введение.

Одной из основных задач интеллектуального анализа данных (Data Mining), который называют ядром технологий Big Data [1], является кластеризация - обнаружение в структурированных данных групп наблюдений, близких по своим признакам, и называемых кластерами (от англ. cluster - гроздь, сгусток) [2]. В целом, задача кластеризация похожа на классификацию: в обеих задачах имеет место группировка объектов по близости значений их признаков. Но если классы заранее задаются, т.е. метки классов примеров в обучающем наборе данных должны быть известны (обучение с учителем), то в кластеризации кластеры заранее не определяются, а формируются в процессе построения модели исключительно на основе информации о расстоянии (обычно, евклидовом) между векторами наблюдений в пространстве признаков (обучение без учителя).

Формально, кластер можно определить, как группу объектов такую, что расстояние между любыми двумя объектами в группе будет меньше, чем расстояние между любым объектом группы и любым объектом из другого кластера. Кластеризация является одним из наиболее популярных методов анализа данных. Это связано с тем, что она лучше подходит для обнаружения новизны. Действительно, если в классификации классы задаются заранее, то любой новый объект, предъявленный классификатору, всегда будет отнесён к одному из имеющихся классов, что не позволит обнаружить появление объектов с принципиально новыми, ранее не наблюдавшимися свойствами. В случае кластеризации для объектов с нестандартными свойствами будет создан новый кластер, что позволит их обнаружить.

Практическое применение кластерных аналитических моделей заключается в содержательной интерпретации кластеров. Если свойства объектов внутри кластера известны, то они могут быть обобщены на любой новый объект, распределённой моделью в этот кластер. Таким образом, кластерная

модель любому предъявленному ей объекту должна присвоить номер или иную метку кластера. Например, если при кластеризации клиентов банка в некотором кластере окажутся клиенты с хорошей кредитной историей, то можно предположить, что любой новый клиент, распределённый в этот же кластер, тоже будет иметь хорошую кредитную историю.

Успех решения задачи кластеризации зависит от правильного выбора алгоритма кластеризации и его параметров, особенно числа кластеров. Все алгоритмы кластеризации можно разделить на три группы:

- машинного обучения (сети Кохонена, k -средних, CLOPE, FOREL);
- вероятностные, в которых наблюдениям присваивается не метки кластеров, а вероятность принадлежности объекта кластеру (EM-алгоритм);
- иерархические - формируют кластеры путём объединения (агломеративные) малых групп в большие или наоборот (дивизимные);
- графовые.

Выбор алгоритма не очевиден и может потребовать экспериментов или привлечения сведений о решении аналогичных задач. Тем не менее выбором алгоритма проблемы кластеризации не ограничиваются. Даже если алгоритм подходит, остаётся ещё задача выбора числа кластеров.

На первый взгляд здесь всё очевидно: число кластеров модели должно соответствовать числу групп наблюдений в исходном наборе данных. Целевой функцией, минимизируемой при обучении модели может быть средний квадрат расстояния от объектов, попавших в кластер, до его центра, что показывает насколько похожи объекты внутри кластера. Другой вариант - максимизировать расстояние от объектов одного кластера до центров других кластеров, что показывает степень различия объектов в разных кластерах. Действительно, чем сильнее объекты похожи внутри кластеров и чем выше междукластерные отличия, тем лучше кластеризация, и тем проще интерпретировать модель.

Однако, часто встречается ситуация, когда кластерная модель, оптимальная с точки зрения формальных критериев, оказывается совершенно бессмысленной с точки зрения целей и логики анализа. Причиной этого может быть разная кластеризующая способность признаков. Например, пусть требуется выполнить кластеризацию клиентов банка по трём признаками - доход, стаж работы и возраст. Если кластеризующая сила двух первых признаков значительно лучше, чем возраста, то мы получим кластеры, где по уровню дохода и стажу клиенты сгруппированы хорошо, а по возрасту - нет. Противоречие здесь в том, что по идее, чем выше возраст, тем выше стаж, и, в общем случае, доход. Однако при неудачной кластеризации данная тенденция окажется не выявлена, что может привести к неправильным выводам по результатам анализа. Сделать модель более интерпретируемой и повысить эффективность принятых с её помощью решений можно, подбирая число кластеров таким образом, чтобы оно отвечало, как формальным критериям точности, так и логике задачи, что требует поиска компромисса. Поэтому и в настоящее время число формируемых кластеров, как параметр алгоритма кластеризации, является дискуссионной темой в сообществе аналитиков, которое непрерывно генерирует какие-то новые идеи и решения в контексте данной проблемы.

Наиболее простым решением можно считать эвристическое правило: "не менее пяти, но не более десяти". Оно означает, что число кластеров менее пяти, скорее всего будет плохо отражать разнообразие свойств объектов, а больше десяти - будет очень сложным для интерпретации. Данное правило является, конечно, слишком общим и не учитывает особенностей конкретной задачи анализа, но позволяет получить приемлемый результат во многих практически значимых случаях. Тем не менее, в сообществе аналитиков бытует мнение, что однозначно лучшей кластеризации не бывает, поэтому получение хорошей кластерной модели может потребовать множества итераций выбора алгоритма кластеризации и его параметров, и особенно числа кластеров.

Кроме приведённого общего правила, в литературе можно встретить множество других критериев [3]:

- метод отношения дисперсий (индекс Калински-Карабаш);
- метод статистики разрыва;
- информационные критерии (Акаике и байесовский);
- метод *ISODATA*;
- метод Девиса-Болдуина;
- метод локтя и др.

Следует отметить, что большинство перечисленных методов не является универсальными, и каждый из них способен привести к разным кластерным решениям в рамках одной и той же задачи. В

связи с этим возникает проблема разработки подхода, который опирается не на какой-то формальный критерий или алгоритм выбора числа кластеров, а предоставляет аналитику возможность оперативного подбора данного параметра с хорошим визуальным представлением кластерной структуры, которая позволила бы оперативно подобрать приемлемое кластерное решение. Особенно данный подход представляет интерес в рамках разведочного анализа Тьюки, когда полученные решения впоследствии могут уточняться с помощью более строгих подходов.

В данной работе авторы рассматривают в качестве подхода к выбору числа кластеров использование кластерных силуэтов. Изначально метод разрабатывался для оценки качества кластерных моделей, но может быть использован и для оценки числа кластеров.

Теоретическая часть.

Впервые метод силуэтов в задаче кластеризации был предложен бельгийским статистиком Питером Руссо (Peter Rousseeuw) в 1987 г [4]. Изначально метод силуэта использовался для оценки качества кластеризации, а именно того, насколько результаты кластеризации соответствуют исходным данным. Метод предлагает компактное графическое представление результатов кластеризации. При этом для каждого объекта из обучающего набора данных вычисляется коэффициент силуэта, который отражает степень того, насколько данный объект похож на другие, попавшие в этот же кластер, и не похож на объекты из других кластеров. Значение коэффициента варьируется от -1 до 1, при этом чем выше значение, тем лучше объект "подходит" своему кластеру, и менее "подходит" другим кластерам.

Таким образом, если в результате кластеризации окажется, что большинство объектов имеет высокое значение коэффициента силуэта (0,5 - 1), то и кластерная структура достаточно выражена и соответствует групповой структуре данных. Если же, напротив, большинство объектов имеет низкое значение коэффициента (около 0) или отрицательное, то результаты кластеризации плохие [5].

Метод силуэта можно использовать для большинства алгоритмов кластеризации и с любыми мерами расстояния, используемыми в них (евклидово, степенное, манхэттенское и др.).

Пусть имеются результаты кластеризации набора данных, при этом для каждой точки данных $i \in C_I$, может быть вычислено значение

$$a(i) = \frac{1}{|C_I| - 1} \sum_{j \in C_I, i \neq j} d_{ij},$$

где C_I - кластер с номером I , $|C_I|$ - мощность кластера, d_{ij} - расстояние между разными объектами i и j в кластере. Тогда a_i можно проинтерпретировать как среднее расстояние между объектом i и всеми другими объектами этого же кластера. Эта величина показывает, насколько объект i "подходит" своему кластеру (чем меньше расстояние, тем лучше).

Теперь следует определить, насколько "плохо" объект i подходит к другим кластерам, для чего необходимо вычислить среднее расстояние от этого объекта до всех объектов некоторого кластера C_J , где $C_J \neq C_I$. После того, как среднее расстояние будет вычислено для всех кластеров, следует выбрать тот их них, для которого оно будет наименьшим, т.е.

$$b(i) = \min_{J \neq I} \frac{1}{|C_J|} \sum_{j \in C_J} d_{ij}.$$

Тогда можно сказать, что кластер J является ближайшим к кластеру I , поэтому будет рассматриваться как следующий кандидат чтобы стать "подходящим" для объекта i . Теперь можно вычислить собственно коэффициент силуэта:

$$s(i) = \frac{b(i) - a(i)}{\max(a(i), b(i))},$$

если $C_I > 1$. Если $C_I = 1$, то $s(i) = 0$, $-1 \leq s(i) \leq 1$.

Следует отметить, что для кластеров, которые содержат только один элемент $a(i)$ не определено. В этом случае устанавливаем $s(i) = 0$, что соответствует середине интервала изменения величины.

Чтобы $s(i)$ оказалось близким к 1, необходимо чтобы выполнялось $a(i) \ll b(i)$. Поскольку $a(i)$ показывает, насколько объект i в среднем отличается от других объектов в том же кластере, малое значение этой величины показывает хорошее соответствие объекта кластеру. Большое значение $b(i)$ указывает на плохое соответствие объекта i соседнему кластеру. Таким образом, значение коэффициента силуэта $s(i)$ близкое к 1 указывает на высокую степень соответствия объекта i своему кластеру, и низкую - соседнему, что и является признаком хорошей кластеризации. Значение $s(i)$ близкое к нулю означает, что объект расположен на границе двух естественных кластеров.

Среднее значение $s(i)$ по всем объектам кластера является мерой того, насколько плотно сгруппированы объекты кластере. Таким образом, среднее значение $s(i)$ по всем объектам исходного набора данных показывает, насколько группировка данных в кластеры соответствует их естественной группировке. Если число кластеров больше или меньше числа естественных групп, что имеет место при неудачном выборе числа кластеров в модели, то из-за низких значений коэффициента силуэта, силуэт всего "плохого" кластера будет уже, чем у остальных. Это происходит, когда один естественный кластер оказывается разбит на два или более результирующих кластера, либо два или более естественных кластера окажутся объединены в один. В любом случае, количество естественных и результирующих кластеров оказывается несоответствующим. Таким образом, кластерные силуэты удобно использовать для выбора числа кластеров в модели, соответствующему числу естественных групп а данных.

Для этого используются специального вида диаграммы, которые так и называют - диаграммы силуэтов [5]. На диаграмме для каждого объекта коэффициент силуэта отображается прямоугольником соответствующей длины. Прямоугольники группируются по кластерам (которые обычно выделяются цветом) и в каждом кластере дополнительно ранжируются в порядке убывания. Общий вид диаграммы силуэтов представлен на рис. 1.

Таким образом, на диаграмме становится виден силуэт каждого кластера. По форме силуэтов аналитик оперативно может оценить качество кластеризации. Чем форма силуэтов ближе к прямоугольной, а площадь (средний коэффициент силуэта) ближе к 1, тем лучше кластеризация. Внутри силуэта каждого кластера объекты расположены в порядке убывания их коэффициента силуэта, поэтому легко увидеть, какие именно объекты лучше соответствуют кластеру, а какие хуже.

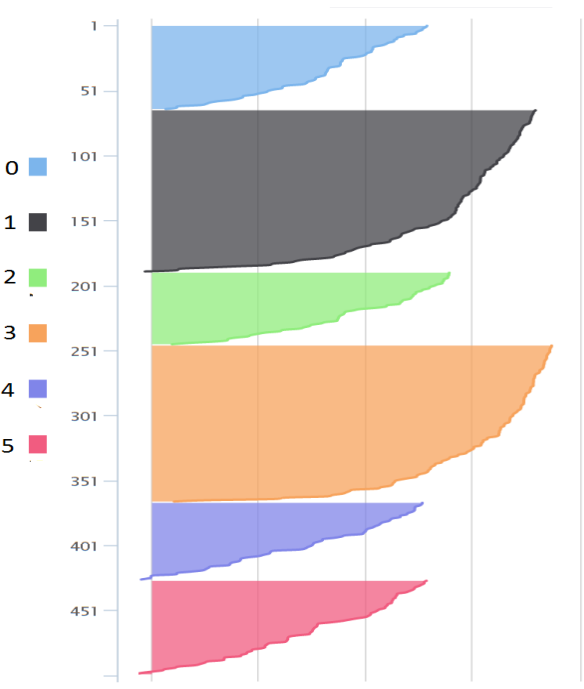


Рисунок 1. Общий вид диаграммы силуэтов

Экспериментальные исследования.

Рассмотрим набор данных, фрагмент которого представлен на рис. 2. Он содержит информацию о заёмщиках кредитной организации, кластеризация которых позволит выявить влияние признаков,

описывающих заёмщика, на вероятность допуска им просрочки по кредиту и оценки её размера.

ID	Возраст	Личный доход	Сумма последнего кредита, руб	Первоначальный взнос, руб	Время проживания по месту пребывания, мес#	Время работы на текущем месте, мес#	Средняя сумма просрочки, руб
1	32	18000	11751	8389	49	27	0
2	62	7000	11160	2790	324	48	0
3	41	60000	79910	19990	5	168	0
4	34	35000	7290	1300	84	13	1439,94
5	31	12000	30198	5752	163	15	5760
6	35	15000	16400	3000	36	24	3070
7	34	6000	3000	790	38	180	0
8	33	10000	7795	7795	36	12	0
9	27	10000	9997	1120	108	12	0
10	57	3000	3280	3000	252	48	0
11	54	11000	22072	8008	18	15	0
12	27	12000	6600	1170	48	12	0

Рисунок 2. Исходный набор данных для кластеризации (фрагмент)

В данной работе не ставится цель провести содержательную интерпретацию кластеров для оценки кредитоспособности клиентов – это задача кредитных менеджеров. Наша цель пояснить на конкретном примере методику выборы числа кластеров с помощью диаграммы силуэтов.

Рассмотрим случай двух кластеров (рис. 3). Для удобства визуального восприятия структура кластеров приведена в двумерной проекции с сохранением топологического подобия, когда на диаграмме показываются расстояния между объектами.

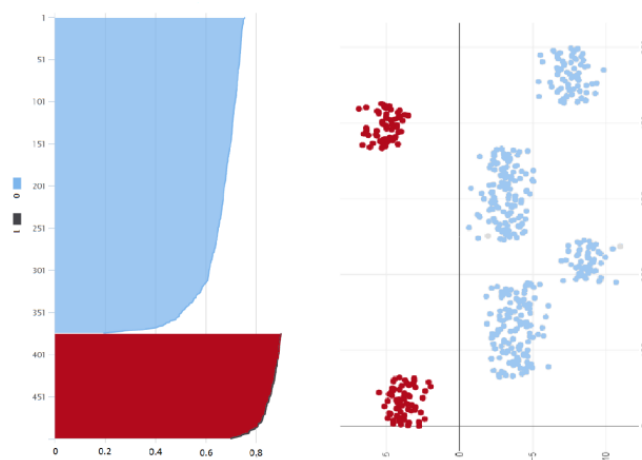


Рисунок 3. Диаграмма силуэтов для случая 2-х кластеров

На рисунке слева представлена диаграмма силуэтов, на которой представлены силуэты двух кластеров, сгенерированных моделью. На правой части рисунка представлена естественная группировка данных, где хорошо различимы 6 групп. Т.е. количество реальных групп не соответствует числу кластеров. В результате на диаграмме мы наблюдаем неравномерность ширины силуэтов.

Случай для 3-х кластеров представлен на рисунке 4.

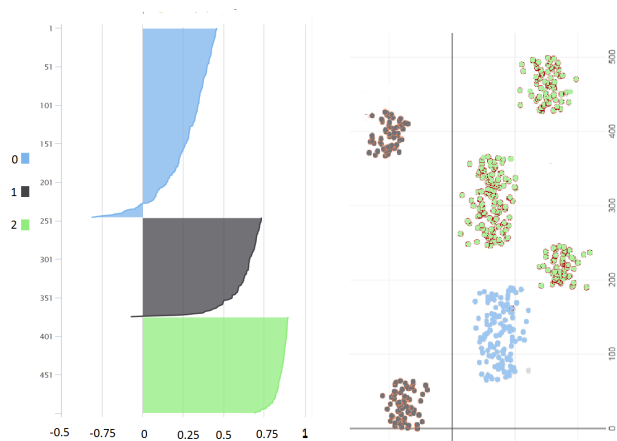


Рисунок 4. Кластеризация для случая 3 кластеров

Для случая 3-х кластеров ситуация усугубляется: силуэт нулевого кластера становится узким, т.е. для него имеют место низкие значения коэффициента силуэта, и даже появляются отрицательные. Это говорит о том, что в среднем, объекты данного кластера слабо связаны и кластеризация по - прежнему не оптимальна.

Для случая 4 кластеров диаграмма силуэтов представлена на рис. 5. На ней силуэты кластеров достаточно широкие и равномерные, что говорит о хорошей кластеризации, когда объекты в среднем хорошо соответствуют своим кластерам, и плохо - остальным кластерам. Следует отметить, что данная ситуация имеет место для случая, когда визуально число групп и число кластеров не совпадает, что вызывает определённое противоречие.

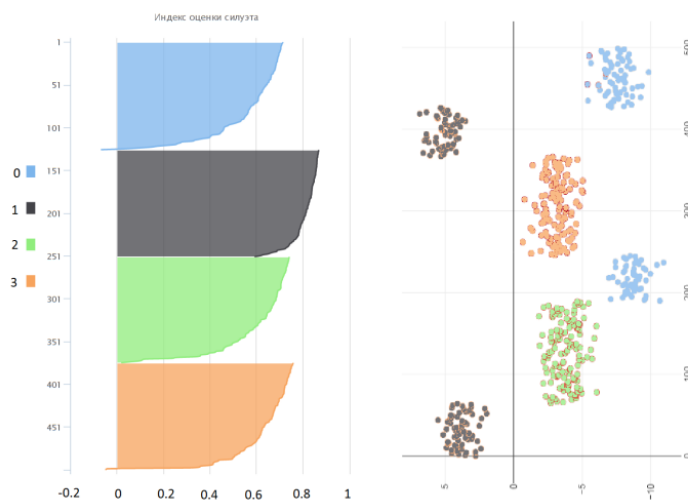


Рисунок 5. Диаграмма силуэтов для случая 4-х кластеров

И, наконец, диаграммы для случаев 5 и 6 кластеров совмещены на рис. 6. Это случай, когда число кластеров наиболее приближено к числу естественных групп данных и ему должна соответствовать лучшая диаграмма силуэтов. Однако на практике это оказалось не так.

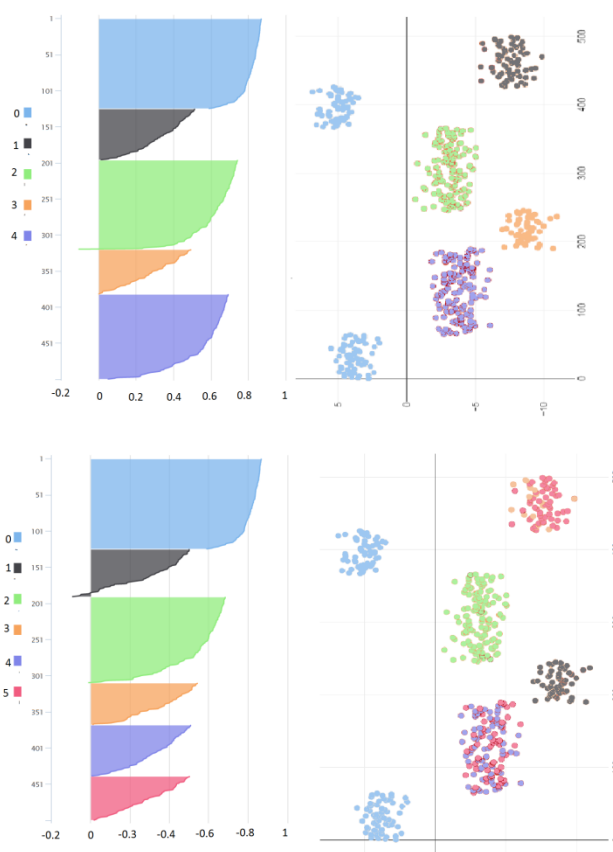


Рисунок 5. Диаграмма силуэтов для случая 5 и 6 кластеров

Для 5 кластеров на диаграмме имеет место появление двух кластеров с узким силуэтом (кластеры 1 и 3), а на диаграмме для 6 кластеров – таковых уже 4 (1, 3, 4 и 5). Напомним, что кластеры с узким силуэтом имеют низкий средний индекс силуэта, что говорит о низком уровне соответствия объектов своему кластеру. Проще говоря, объекты внутри кластера не очень похожи друг на друга.

Поэтому несмотря на то, что случаи 5 и 6 кластеров лучше соответствуют естественной группировке данных, с точки зрения метода силуэта предпочтение имеет модель для 4-х кластеров, которая обеспечивает наименьшее внутрикластерное и наибольшее междукластерное расстояния, что соответствует лучшей кластеризации.

Выводы.

Таким образом, экспериментальные исследования с использованием метода силуэтов на примере кластеризации клиентов банка показали, что лучшая кластеризация с точки зрения минимума средних внутрикластерных расстояний и максимума междукластерных, не обязательно достигается, когда число кластеров как параметра модели соответствует числу естественных групп данных. Это способно вызвать определённые противоречия при содержательной интерпретации кластеров. Применение метода силуэтов позволяет обнаруживать такие противоречия и помогать аналитикам делать более точные выводы о предметной области, описываемой исходными данными.

Вместе с тем, обобщение полученных на конкретном примере результатов даже на аналогичные задачи следует производить с осторожностью. Это связано с тем, что разные алгоритмы кластеризации могут создавать различные кластерные структуры даже на одних и тех же данных. Поэтому подобные исследования должны производиться отдельно в каждом конкретном случае. Хотя сама по себе методика, изложенная выше, является достаточно универсальной.

Заключение.

Таким образом в работе рассмотрена техника применения кластерных силуэтов для выбора числа кластеров в задачах кластеризации. На решении практической задачи с использованием реальных данных показано, что число кластеров модели с точки зрения минимизации внутрикластерных расстояний и максимизации междукластерных, не всегда достигается при соответствии числа кластеров модели

естественной группировки данных.

Список литературы

- [1] Паклин Н.Б. Бизнес-аналитика: от данных к знаниям (+ CD): учеб. пособие. / Паклин Н.Б., Орешков В.И. 2-е изд., испр. – СПб.: Питер, 2013. – 704 с.
- [2] Корячко В.П. Интеллектуальные системы и нечеткая логика / В.П. Корячко, М.А. Бакулева, В.И. Орешков. – М.: КУРС, 2017. – 346 с.
- [3] Орешков В.И. Повышение точности классификации данных с использованием алгоритма k-ближайших соседей на основе прекластеризации обучающих данных / Орешков В.И. Вестник Рязанского гос. радиотехнического университета. Рязань: РГРТУ, – 2021. – № 76. – С. 65–73.
- [4] Peter J. Rousseeuw. Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. Journal of Computational and Applied Mathematics. Volume 20, November 1987, Pages 53-65.
- [5] Паклин Н.Б. Кластерные силуэты / Паклин Н.Б., Орешков В.И. Сб. научн. тр. XX Международной науч.-практ. конф. «Системный анализ в проектировании и управлении». СПб, 29 июня-01 июля 2016 г. Изд-во ФГАОУ ВО «Санкт-Петербургский политехнический ун-тет Петра Великого». – 2016. – с. 314-321.

CHOICE OF THE NUMBER OF CLUSTERS IN CLUSTERING PROBLEMS USING THE SILHOUETTE METHOD

V.P. Koryachko

Head of the Department of Computer-Aided Design Systems, Ryazan State Radio Engineering University named after V.F. Utkin, Doctor of Technical Sciences, Professor

V.I. Oreshkov

Associate Professor of the Department of Computer-Aided Design Systems, Ryazan State Radio Engineering University named after V.F. Utkin, candidate of technical sciences

*Department of Computer-Aided Design Systems
Faculty of Computer Science
Ryazan State Radio Engineering University named after V.F. Utkin
E-mail: vyacheslav.oreshkov@yandex.ru*

Abstract. The problem of choosing the number of clusters in clustering problems is considered. A technique for choosing the number of clusters based on the cluster silhouette method is proposed. As a result of the experiment on real data, it is shown that the best number of clusters in terms of the minimum intracluster and maximum intercluster distances does not always correspond to the number of natural data groups.

Keywords: data mining, machine learning, unsupervised learning, training data, cluster, clustering, silhouette coefficient.

УДК 681.32

МИНИМИЗАЦИЯ ГРАФА ДОСТИЖИМЫХ СОСТОЯНИЙ ПАРАЛЛЕЛЬНОГО АВТОМАТА



Л.Д. Черемисинова

Главный научный сотрудник ОИПИ НАНБ
доктор технических наук, профессор
cld@newman.bas-net.by

Л. Д. Черемисинова

Окончила Томский государственный университет, доктор технических наук, профессор. Работает в ОИПИ НАН Беларуси в должности главного научного сотрудника и Белорусском государственном университете информатики и радиоэлектроники в должности профессора. Круг научных интересов: дискретная математика, логическое проектирование и тестирование дискретных систем управления, реализация параллельных алгоритмов управления.

Аннотация. Рассматривается проблема тестирования управляющих систем, которым присущ параллелизм происходящих в них процессов. В качестве языка задания таких систем используется модель параллельного автомата, представляющая собой подкласс цветных сетей Петри. Анализ поведенческих свойств параллельного автомата проводится на основе его графа достижимости, вершинам которого соответствуют все возможные разметки состояний, число которых растет экспоненциально с ростом степени параллельности состояний. В работе предлагается метод минимизации графа достижимости путем сокращения числа его вершин на основе введения доопределения частичного порядка на множестве параллельно срабатывающих переходов автомата.

Ключевые слова: параллельный автомат, граф достижимых состояний, тестирование дискретных устройств.

Введение.

Развитие микроэлектроники и средств автоматизации проектирования привели к возможности проектирования микроэлектронных управляющих систем значительной сложности. В связи с этим первостепенное значение получила проблема обеспечения надежности и предсказуемости поведения таких сложных систем. Неотъемлемой частью процесса проектирования стало тестирование, в частности, проверка соответствия поведения устройства требованиям, предъявляемым спецификацией на его проектирование. Сложность тестирования возрастает быстрее, чем сложность тестируемых систем, в том числе и в связи с растущими требованиями к качеству [1, 2].

Под тестированием здесь понимается проверка схемной реализации на вход-выходное соответствие модели (input-output conformance), которая состоит в

- генерации проверяющей последовательности входных сигналов на основе заданной спецификации;
- моделировании схемной реализации устройства на полученной тестовой последовательности;
- наблюдении реакций на подаваемые тестовые последовательности сигналов;
- определении, соответствуют ли вход-выходные реакции схемной реализации спецификации на ее проектирование.

Можно выделить следующие два основных способа генерации тестовых входных воздействий (для проверяющей последовательности):

- генерация псевдослучайных наборов воздействий;

– генерация наборов воздействий и проверяющей последовательности в целом, исходя из описания модели.

Минусы первого подхода заключаются в том, что число возможных входных воздействий в каждом состоянии схемы (и модели) экспоненциально зависит от числа переменных, и, главное, не все такие наборы попадают в область определения модели, а значит, и штатного поведения схемной реализации. Второй подход обеспечивает не только наиболее полное тестирование схемы на области, определяемой спецификацией на ее проектирование, но и позволяет сократить длину проверяющей последовательности.

В настоящее время наиболее разработанным направлением в области тестирования является верификация на основе моделей, представляемых конечными автоматами [3, 4]. Менее изучена проблема тестирования систем, которым присущ параллелизм происходящих в них процессов [5, 6]. К ним относятся системы, обеспечивающие управление взаимодействующими компонентами, которые работают параллельно и асинхронно. Параллелизм, присутствующий в объектах управления, отражается в функциональной модели цифровых систем, управляющих данными объектами. Обычно рассматриваются системы, которые состоят из параллельно работающих компонентов и моделируются сетями конечных автоматов. Менее изучена задача тестирования систем с использованием моделей «истинного параллелизма», таких как сети Петри [7, 8]. В таких системах процессы выполняются параллельно и независимо, и нет необходимости контролировать порядок их выполнения. Для управляющих систем рассматриваемого класса характерно также и то, что управляющие воздействия и сигналы о состоянии объектов управления описываются булевыми переменными, лишь небольшой процент всей информации является числовым. Основным инструментом анализа поведенческих свойств моделей этого класса является граф достижимых состояний, в процессе его обхода генерируется тестовая последовательность [7, 8, 9].

В настоящей работе рассматривается задача тестирования систем управления с «истинным параллелизмом» происходящих в них процессов. В качестве языка задания спецификации используется язык параллельных алгоритмов логического управления, а точнее задание описаний в стандартном виде – моделью, названной параллельным автоматом [10]. Алгоритмы в таком виде представляют собой подкласс цветных сетей Петри – расширенные сети свободного выбора [11, 12]. Тестовая последовательность для анализа схемной реализации параллельного автомата строится в процессе обхода его графа достижимости, вершинам которого соответствуют все возможные достижимые состояния автомата.

Построение графа достижимости (в явном или неявном виде) всегда возможно для моделей с конечным пространством состояний, однако проблемой является экспоненциальный рост числа вершин графа достижимости. Основной причиной такого роста (кроме большого числа состояний) является наличие параллельно происходящих в исследуемой системе процессов, что отражается в степени параллелизма состояний. Большое пространство состояний негативно влияет на время тестирования. Более того, параллельный автомат может быть настолько сложным, что становится невозможным в приемлемое для практики время построить граф достижимости и получить даже какую-нибудь тестовую последовательность путем обхода его вершин.

В настоящей работе описываются пути сокращения графа достижимости на основе введения отношения частичного порядка на множестве параллельно срабатывающих переходов. Предлагается модификация метода построения графа достижимости параллельного автомата, предложенного в работе [13]. Целью модификация является сокращение числа вершин графа, в основе которого лежит частичное упорядочение пар переходов автомата на основе учета их информационного взаимодействия.

Параллельный автомат. Параллельный автомат можно рассматривать как динамическую дискретную систему, допускающую параллельно протекающие процессы, которым соответствуют параллельные частичные состояния. В параллельном автомате [10] (в

отличие от классического конечного автомата) рассматриваются частичные состояния, названные так потому, что в любой момент времени параллельный автомат может одновременно находиться в нескольких таких состояниях. Множество частичных состояний в момент времени определяет полное (глобальное) состояние S_t параллельного автомата, которое можно интерпретировать как маркировку сети Петри. Частичные состояния, составляющие полное состояние автомата, соответствуют параллельно протекающим процессам и называются параллельными. Следует заметить, что количество полных состояний параллельного автомата растет экспоненциально с ростом числа частичных состояний и переходов соответствующего параллельного автомата [10].

Параллельный автомат задается совокупностью множеств $\{S, X, Y, T\}$ частичных состояний, входных переменных, выходных переменных и переходов между подмножествами частичных состояний $S_{i1}, S_{i2} \subseteq S$. Переход параллельного автомата задается в форме

$$\tau_i = (\gamma_i^1, k_i^1) \rightarrow (\gamma_i^2, k_i^2),$$

где γ_i^1 и γ_i^2 трактуются как подмножества частичных состояний из S , в которых автомат находится перед и после срабатывания i -го перехода; k_i^1 и k_i^2 – элементарные конъюнкции входных и выходных булевых переменных, трактуемые как условие перехода и выходные сигналы, сопровождающие переход.

Переход срабатывает, когда текущее полное состояние S_t автомата включает все частичные состояния из γ_i^1 ($\gamma_i^1 \subseteq S_t$) а переменные автомата принимают значения, обращающие элементарную конъюнкцию k_i^1 в единицу ($k_i^1 \wedge W_t = 1$) на множестве W_t значений логических переменных системы управления, которое было достигнуто к моменту времени t . Таким образом, динамика параллельных автоматов описывается в пространстве достижимых полных состояний и значений логических сигналов.

Переходы параллельного автомата могут происходить последовательно или параллельно, в последнем случае в случае корректного автомата (рассматриваются только они) для любой пары переходов τ_i и τ_j ($i \neq j$) выполняются условия: если $\gamma_i^1 \cap \gamma_j^1 \neq \emptyset$, то $\gamma_i^1 = \gamma_j^1$, и, если переходы τ_i и τ_j параллельны, то $\gamma_i^2 \neq \gamma_j^2$.

Чтобы вычислить следующее полное состояние S_{t+1} автомата, необходимо определить множество T переходов τ_i , которые срабатывают в состоянии S_t и состоянии W_t на множестве переменных автомата. Следующие состояния S_{t+1} и W_{t+1} задаются следующим образом:

1) $S_{t+1} = (S_t \setminus \gamma^1) \cup (\gamma^2)$, где $\gamma^1 = \cup_i \gamma_i^1$ и $\gamma^2 = \cup_i \gamma_i^2$ есть объединения множеств частичных состояний всех параллельно выполняемых переходов $\tau_i \in T$;

2) состояние W_{t+1} на множестве переменных получается из W_t таким изменением значений переменных, входящих в конъюнкции k_i^2 всех переходов $\tau_i \in T$, что все $k_i^2 = 1$.

Ниже приведен пример описания простого параллельного автомата из работы [13], определенного на множествах $S = \{1, 2, \dots, 11\}$ частичных состояний, $X = \{x_1, x_2\}$ и $Y = \{y_1, y_2\}$ входных и выходных (управляющих) переменных и девяти переходах:

$$\tau_1 = (1, \underline{x_1} \underline{x_2}) \rightarrow (10, \underline{y_1} \underline{y_2});$$

$$\tau_2 = (10, \underline{x_2}) \rightarrow (2.3.4, -);$$

$$\tau_3 = (2, -) \rightarrow (5.6, \underline{y_1});$$

$$\tau_4 = (3.5, \underline{x_2}) \rightarrow (8, -);$$

$$\tau_5 = (4, \underline{x_1}) \rightarrow (7, \underline{y_1});$$

$$\tau_6 = (4, \underline{x_1}) \rightarrow (9, \underline{y_2});$$

$$\tau_7 = (7, \underline{x_2}) \rightarrow (9, -);$$

$$\tau_8 = (6.8.9, -) \rightarrow (11, \underline{y_2});$$

$$\tau_9 = (11, \underline{x_1}) \rightarrow (1, -).$$

Генерация графа достижимых состояний. Граф достижимости автомата представляет собой ориентированный граф, вершинам которого соответствуют все возможные полные состояния P_i автомата, а дугам – переходы между этими состояниями. Дуга графа помечается

символом перехода τ_i и связывает вершины графа, помеченные множествами полных состояний P_p и P_q , если результат срабатывания перехода τ_i меняет состояние P_p автомата на P_q . Граф достижимых состояний представляет собой в общем случае мультиграф, в котором могут быть петли и кратные дуги, различающиеся присвоенными им метками. Кроме того, очевидно, что каждый переход алгоритма управления может повторяться много раз в качестве метки разных дуг графа достижимости.

Граф достижимости получается путем вычисления всех полных состояний P_i (начиная с начальной), которые ставятся в соответствие вершинам графа, и дуг, связывающих вершины, которые помечаются теми переходами автомата, которые вызывают изменение соответствующих полных состояний [13].

Множество полных состояний $P = \{P_1, P_2, \dots, P_l\}$ автомата строится последовательно. Начиная с заданного начального состояния $P = \{P_1\}$, выбирается очередное уже сформированное полное состояние $P_i \subseteq P$, и для него производится последовательный просмотр всех переходов τ_j автомата. Если для текущего выбранного перехода $\tau_j = (\gamma_j^1, k_j^1) \rightarrow (\gamma_j^2, k_j^2)$ выполняется $\gamma_j^1 \subseteq P_i$, множество полных состояний P заменяется на $P \cup \{(P_i \setminus \gamma_j^1) \cup \gamma_j^2\}$ (увеличивается в общем случае). Процесс заканчивается после просмотра всех состояний из P .

В результате работы алгоритма для приведенного выше параллельного автомата получается множество, состоящее из 12 полных состояний: $\{\{1\}, \{10\}, \{2, 3, 4\}, \{3, 4, 5, 6\}, \{2, 3, 7\}, \{2, 3, 9\}, \{4, 6, 8\}, \{3, 5, 6, 7\}, \{6, 8, 7\}, \{3, 5, 6, 9\}, \{6, 8, 9\}, \{11\}\}$.

Найденные полные состояния P_1, P_2, \dots, P_l являются вершинами ориентированного графа достижимых состояний. Из вершины P_g исходит дуга, заходящая в вершину P_h , если в исходном задании имеется переход $\tau_j = (\gamma_j^1, k_j^1) \rightarrow (\gamma_j^2, k_j^2)$, такой, что $P_h = (P_g \setminus \gamma_j^1) \cup \gamma_j^2$. Метод построения графа достижимости, лежащий в основе данного алгоритма [13], заключается в переборе пар полных состояний (P_g, P_h) и поиске перехода τ_j переводящего P_g в P_h или P_h в P_g . Из вершины P_g графа достижимости исходит дуга в вершину P_h , если $P_g \setminus P_h = \gamma_j^1$ и $P_h \setminus P_g = \gamma_j^2$. Если же $P_h \setminus P_g = \gamma_j^1$ и $P_g \setminus P_h = \gamma_j^2$, то дуга исходит из вершины P_h в вершину P_g . Иначе вершины P_g и P_h не связаны.

Граф достижимости удобно задать перечнем дуг и приписанных им меток переходов автомата, так как графы данного вида обладают сравнительно небольшим числом дуг. Полученный граф достижимости имеет 12 вершин и 19 дуг: $(\{1\}, \{10\}, \tau_1)$, $(\{10\}, \{2,3,4\}, \tau_2)$, $(\{2,3,4\}, \{3,4,5,6\}, \tau_3)$, $(\{2,3,4\}, \{2,3,7\}, \tau_5)$, $(\{2,3,4\}, \{2,3,9\}, \tau_6)$, $(\{3,4,5,6\}, \{4,6,8\}, \tau_4)$, $(\{3,4,5,6\}, \{3,5,6,7\}, \tau_5)$, $(\{3,4,5,6\}, \{3,5,6,9\}, \tau_6)$, $(\{2,3,7\}, \{3,5,6,7\}, \tau_3)$, $(\{2,3,7\}, \{2,3,9\}, \tau_7)$, $(\{4,6,8\}, \{6,7,8\}, \tau_5)$, $(\{4,6,8\}, \{6,8,9\}, \tau_6)$, $(\{3,5,6,7\}, \{6,7,8\}, \tau_4)$, $(\{3,5,6,7\}, \{3,5,6,9\}, \tau_7)$, $(\{2,3,9\}, \{3,5,6,9\}, \tau_3)$, $(\{6,7,8\}, \{6,8,9\}, \tau_7)$, $(\{3,5,6,9\}, \{6,8,9\}, \tau_4)$, $(\{6,8,9\}, \{11\}, \tau_8)$, $(\{11\}, \{1\}, \tau_9)$.

Более наглядно граф задается в матричном виде, где строкам и столбцам соответствуют полные состояния P_i автомата, и на пересечении i -й строки и j -го столбца указывается переход из состояния P_i в P_j . Для полученного графа эта матрица имеет следующий вид:

Таблица 1. Генерация графа достижимых состояний

	P_1	P_2	P_3	P_4	P_5	P_6	P_7	P_8	P_9	P_{10}	P_{11}	P_{12}
$P_1=\{1\}$		τ_1										
$P_2=\{10\}$			τ_2									
$P_3=\{2,3,4\}$				τ_3	τ_5	τ_6						
$P_4=\{3,4,5,6\}$							τ_4	τ_5	τ_6			
$P_5=\{2,3,7\}$							τ_7	τ_3				
$P_6=\{2,3,9\}$									τ_3			
$P_7=\{4,6,8\}$										τ_5	τ_6	
$P_8=\{3,5,6,7\}$									τ_7	τ_4		
$P_9=\{3,5,6,9\}$											τ_4	
$P_{10}=\{6,7,8\}$											τ_7	
$P_{11}=\{6,8,9\}$												τ_8
$P_{12}=\{11\}$	τ_9											

Сокращение числа вершин в графе достижимости.

В настоящее время известны подходы к сокращению размерности графа достижимости (числа вершин и дуг) без потери важной информации для решаемой задачи (обзор таких методов можно найти в работе [9]). Основным недостатком предложенных правил редукции графа является то, что их применимость ограничена относительно специфическими структурами. Эффективным подходом к редукции графа достижимости является доопределение частичное упорядочение параллельных переходов (partial order reduction [14]). В основе методов сокращения графа лежит предположение о коммутативности асинхронных одновременно происходящих процессов. Такие методы применяются в методах формальной верификации сетей Петри, однако эти методы не учитывают информационное взаимодействие переходов, что неприемлемо при решении задачи генерации тестов на основе графа достижимых состояний.

Доопределение частичного порядка на множестве параллельных переходов основано на том факте, что некоторые параллельно выполняемые переходы не зависят друг от друга, в том смысле, что порядок выполнения одних переходов не влияет на условия срабатывания других. Соответственно, можно рассматривать не все возможные порядки выполнения таких переходов, а упорядочить их произвольным образом, в том числе и совместив их выполнение. Получаемый в результате граф достижимых состояний эквивалентен исходному по степени соответствия спецификации системы, но содержит значительно меньше вершин и дуг.

Пара переходов $\tau_i = (\gamma_i^1, k_i^1) \rightarrow (\gamma_i^2, k_i^2)$ и $\tau_j = (\gamma_j^1, k_j^1) \rightarrow (\gamma_j^2, k_j^2)$, одновременно срабатывающих в состоянии $S_i(W_i)$, являются совместимыми (коммутативными) в рамках этого состояния, если они могут выполняться в любом порядке, приводя в итоге к одному и тому же полному состоянию S_{i+1} автомата. Совместимость переходов τ_i и τ_j имеет место, если:

- 1) условия их срабатывания совместимы, т.е. если $k_i^1 \wedge k_j^1 \neq 0$;
- 2) изменение значений переменных в результате срабатывания каждого из этих переходов не приводит к нарушению условия срабатывания другого, т.е. в том случае, если $k_j^1 \wedge k_i^2 \neq 0$ и $k_i^1 \wedge k_j^2 \neq 0$;
- 3) изменения значений выходных сигналов, вызываемые переходами, совместимы, т.е. если $k_i^2 \wedge k_j^2 \neq 0$.

При любом порядке выполнения совместимых переходов полное состояние S_i автомата заменяется на состояние $(S_i \setminus (\gamma_i^1 \cup \gamma_j^1)) \cup (\gamma_i^2 \cup \gamma_j^2)$. Определение совместимости переходов может быть обобщено и на случай более двух переходов: переходы, возможные в некотором состоянии, могут быть выполнены одновременно, если они попарно совместимы.

Построение сокращенного графа достижимости параллельного автомата.

Если в процессе построения графа достижимости для очередной достигнутой маркировки возможно срабатывание пары (или более) параллельных переходов, не находящихся в отношении следования, то следующее полное состояние строится как результат одновременного срабатывания тех из этих переходов, которые являются совместимыми.

Предложенный в работе [13] алгоритм генерации полных состояний параллельного автомата модифицируется следующим образом. Множество полных состояний $P = \{P_1, P_2, \dots, P_l\}$ автомата строится последовательно, начиная с заданного начального состояния $P = \{P_1\}$. Также в алгоритме [13] на каждом шаге выбирается очередное уже сформированное полное состояние $P_g \subseteq P$, и для него производится последовательный просмотр всех переходов τ_i автомата. В результате

- формируется множество T_g переходов $\tau_i = (\gamma_i^1, k_i^1) \rightarrow (\gamma_i^2, k_i^2)$, достижимых из полного состояния P_g , т. е. переходов, для которых $\gamma_i^1 \subseteq P_g$;
- из множества T_g выделяются максимальные по включению подмножества $T_g^l \subseteq T_g$ совместимых переходов, образуя покрытие множества T_g этими подмножествами;
- для каждого подмножества T_g^l совместимых переходов формируется порожаемое ими полное состояние $P_{g+1}^l: P_{g+1}^l = (P_g \setminus (\cup_i \gamma_i^1 / \tau_i \in T_g^l)) \cup (\cup_i \gamma_i^2 / \tau_i \in T_g^l)$, которое объявляется новым

достижимым полным состоянием и порождает вершину графа достижимости (если оно отлично от полученных ранее);

– в граф достижимости вводится дуга из g -й вершины P_g во вновь введенную вершину P_{g+1} , порожденную подмножеством переходов из T_g^1 .

Этот процесс повторяется для каждого из вновь полученных полных состояний. Процесс заканчивается построением графа достижимых состояний, когда не получается новых полных состояний, отличных от уже полученных.

Например, из описания приведенного выше автомата видно, что в состоянии $P_3 = \{2, 3, 4\}$ возможно срабатывание трех переходов $\tau_3 = (2, -) \rightarrow (5,6, y_1)$; $\tau_5 = (4, x_1) \rightarrow (7, y_1)$ и $\tau_6 = (4, x_1) \rightarrow (9, y_2)$, соответственно $T_3 = \{\tau_3, \tau_5, \tau_6\}$. Максимальными совместимыми подмножествами переходов являются $T_3^1 = \{\tau_3, \tau_5\}$ и $T_3^2 = \{\tau_3, \tau_6\}$, они дают минимальное покрытие множества T_3 . Срабатывание переходов из $T_3^1 = \{\tau_3, \tau_5\}$ порождает новое полное состояние $P_4 = \{3, 5, 6, 7\}$, которое достижимо при срабатывании переходов τ_3 и τ_5 в любом порядке (в зависимости от сигналов, поступающих на входы автомата). Срабатывание переходов из $T_3^2 = \{\tau_3, \tau_6\}$ порождает новое полное состояние $P_5 = \{3, 5, 6, 9\}$. Соответственно, в граф достижимости вводятся вершины P_4 и P_5 и две дуги: (P_3, P_4) и (P_3, P_5) .

Полученный граф достижимости имеет восемь вершин и 10 дуг: $(\{1\}, \{10\}, \tau_1)$, $(\{10\}, \{2,3,4\}, \tau_2)$, $(\{2,3,4\}, \{3,5,6,7\}, \tau_3, \tau_5)$, $(\{2,3,4\}, \{3,5,6,9\}, \tau_3, \tau_6)$, $(\{3,5,6,7\}, \{3,5,6,9\}, \tau_7)$, $(\{3,5,6,7\}, \{6,7,8\}, \tau_4)$, $(\{3,5,6,9\}, \{6,7,8\}, \tau_4)$, $(\{6,7,8\}, \{6,8,9\}, \tau_7)$, $(\{6,8,9\}, \{11\}, \tau_8)$, $(\{11\}, \{1\}, \tau_9)$.

В матричное задание полученного графа достижимости имеет следующий вид:

Таблица 2. Генерация сокращенного графа достижимости параллельного автомата

	P_1	P_2	P_3	P_4	P_5	P_6	P_7	P_8
$P_1 = \{1\}$		τ_1						
$P_2 = \{10\}$			τ_2					
$P_3 = \{2,3,4\}$				τ_3, τ_5	τ_3, τ_6			
$P_4 = \{3,5,6,7\}$					τ_7	τ_4		
$P_5 = \{3,5,6,9\}$							τ_4	
$P_6 = \{6,7,8\}$							τ_7	
$P_7 = \{6,8,9\}$								τ_8
$P_8 = \{11\}$	τ_9							

Таким образом, в результате выполнения предложенного алгоритма удалось сократить более чем на четверть число вершин (с 12 до 8) и почти в два раза число дуг (с 19 до 10) графа достижимых состояний параллельного автомата.

Заключение.

Предложен метод построения сокращенного графа достижимых состояний параллельного автомата, который основан на предположении независимости срабатывания параллельно происходящих переходов автомата. Сокращение графа достигается доопределением частичного порядка на множестве переходов. Получаемое сокращение графа достижимости позволяет решать задачу построения тестов для сложных систем управления с параллелизмом поведения.

Список литературы

- [1]. Ebert, C. Embedded Software: Facts, Figures, and Future / C. Ebert, C. Jones // Computer. – 2009. – Vol. 42, no. 4. – P. 42–52.
- [2]. Валидация на системном уровне. Высокоуровневое моделирование и управление тестированием: пер. с англ. Е. Б. Махияновой / М. Чэнь [и др.]. – М. : Техносфера, 2014. – 296 с.
- [3]. Lee, D. Principles and methods of testing finite state machine – a survey / D. Lee, M. Yannakakis // Proceedings of the IEEE. – 1996. – Vol. 84, no. 8. – P. 1090–1123.
- [4]. Kanso, B. Compositional testing for FSM-based models / B. Kanso, O. Chebaro // Intern. J. of Software Engineering & Applications (IJSEA). – 2014. – Vol. 5, no. 3. – P. 1–20.
- [5]. Ponce de Leon, H. Model-based Testing for Concurrent Systems with Labeled Event Structures / H. Ponce de Leon, S.H. Delphine Longuet // Software Testing, Verification & Reliability. – 2014. – Vol. 24, no. 7. – P. 558–590.

- [6]. Tretmans, J. Model based testing with labelled transition systems / J. Tretmans // Formal Methods and Testing: Lecture Notes in Computer Science. – Springer, 2008. – Vol. 4949. – P. 1–38.
- [7]. Zhu, H. A Methodology of Testing High-level Petri Nets / H. Zhu, X.D. He // Information and Software Technology. – 2002. – Vol. 44. – P. 473–489.
- [8]. Liu, J. I/O Conformance Test Generation with Colored Petri Nets / J. Liu, X. Ye1, J. Zhou, X. Song // Applied Mathematics and Information Sciences. – 2014. – Vol. 8, no. 6. – P. 2695–2704.
- [9]. Karatkevich, A. Dynamic Analysis of Petri Net-based Discrete Systems / A. Karatkevich. – Berlin : Springer-Verlag, 2007. – Vol. 358. – 166 p.
- [10]. Закревский, А. Д. Параллельные алгоритмы логического управления / А. Д. Закревский. – Минск: Ин-т техн. кибернетики НАН Беларуси, 1999. – 202 с.
- [11]. Питерсон, Дж. Теория сетей Петри и моделирование систем: пер. с англ. М. В. Горбатовой, В. Л. Торхова, В. Н. Четверикова / Дж. Питерсон. – М.: Мир, 1984. – 264 с.
- [12]. Hack, M. Analysis of production schemata by Petri nets / M. Hack // Project MAK-94. – Cambridge, 1972. – 119 p.
- [13]. Поттосин, Ю. В. Верификация систем с параллелизмом поведения на основе графа достижимых состояний / Ю.В. Поттосин, В.И. Романов, Л.Д. Черемисинова // Информатика. – 2019. – Т. 16, № 2. – С. 62–72.
- [14]. Lluch-Lafuente, A. Partial Order Reduction in Directed Model Checking / A. Lluch-Lafuente, S. Edelkamp, S. Leue // Proceedings of the 9th International SPIN Workshop on Model Checking of Software. Berlin, Heidelberg: Springer-Verlag, 2002 April. – Vol. 11–13. – P. 112–127.

MINIMIZATION OF THE REACHABILITY GRAPH FOR A PARALLEL AUTOMATON

L.D. Cheremisinova

*Principal researcher of UIIP of NAS of Belarus, doctor
of technical sciences, professor*

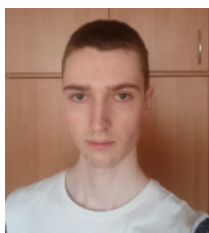
*United Institute of Informatics Problems of National Academy of Sciences of Belarus,
Republic of Belarus
E-mail: cher, cld}@newman.bas-net.by*

Abstract. The problem of testing control systems, which are inherent in the parallelism of the processes occurring in them, is considered. The parallel automaton model, which is a subclass of colored Petri nets, is used as a language for specifying such systems. The analysis of the behavioral properties of a parallel automaton is carried out on the basis of its reachability graph, the vertices of which correspond to all possible states markings, the number of which grows exponentially with an increase in the degree of parallelism of states. The paper proposes a method for minimizing the reachability graph by reducing the number of its vertices based on the introduction of the extension of the partial order on the set of parallel firing the automaton transitions.

Keywords: parallel automaton, reachability graph, discrete device testing.

УДК 004.62:316.472.4

ПРИМЕНЕНИЕ BIG DATA ДЛЯ АНАЛИЗА СОЦИАЛЬНЫХ СЕТЕЙ



Е.И. Баяк

Студент специальности ИиТП
факультета компьютерных
систем и сетей БГУИР
ibayak@gmail.com



С.Н. Нестеренков

Декан факультета компьютерных
систем и сетей БГУИР, кандидат
технических наук, доцент
s.nesterenkov@bsuir.by



Д.А. Жалейко

Инженер-программист
ОИТ БГУИР
d.zhaleyko@bsuir.by

Е.И. Баяк

Окончил среднюю школу №13 города Гродно. Обучается в Белорусском государственном университете информатики и радиоэлектроники.

С.Н. Нестеренков

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

Д.А. Жалейко

Окончил в Белорусский государственный университет информатики и радиоэлектроники в 2021 году по специальности "Вычислительные машины, системы и сети".

Аннотация. Большие данные или Big Data современный термин, описывающий огромные объемы как структурированной, так и неструктурированной информации. Социальные сети, как источник больших данных, имеют отличительные черты в их сборе и методах анализа. А полученные результаты применяются в различных сферах, начиная от обеспечения информационной безопасности и заканчивая улучшением финансовых показателей.

Ключевые слова: Прогнозная аналитика, информационная безопасность, интеллектуальный анализ данных, социальная сеть.

Введение.

Термин большие данные появился в 2008 году. Big Data – это огромное разнообразие данных, которые постоянно поступают в хранилище в больших объемах. К большим данным относят аббревиатуру VVV (объем, скорость, разнообразие). То есть большие данные характеризуются многообразием, быстрым поступлением и большим объемом данных [1, 2]. Но почему этот термин зачастую связывают с социальными сетями?

По состоянию на январь 2023 года за год количество заходящих в социальные сети людей увеличилось на 137 миллионов человек. Среднее ежедневное использование составляет 2 часа 31 минуту. Большая часть населения мира (4,76 миллиарда человек) пользуются социальными сетями. Следовательно, к источникам больших данных можно отнести социальные сети или мессенджеры. В этом случае Big Data представляет информацию о зарегистрированных пользователях в социальной сети, их информацию, картинки профилей и сообщения которыми обмениваются миллионы людей ежедневно. В некоторых случаях платформа может собирать информацию о кликах и разнообразных действиях пользователей. Различные изображения аудио-, видеофайлы поступают в неструктурированном виде и не могут быть сохранены в таблицах традиционных реляционных баз данных. Поэтому различные мессенджеры являются наглядным источником больших данных, так как информация разнообразна, постоянно изменяется, добавляется и удаляется.

Аспекты применения Big Data для анализа социальных сетей.

Большие данные могут быть применены в различных технологиях. Для социальной сети сбор и анализ информации о действиях пользователей, количестве зарегистрированных, активных за определенный период времени участников является обязательным для улучшения и разработки. Настоящие технологии позволяют получить полезные данные о качестве обслуживания клиентов гораздо эффективнее, чем раньше. С помощью Big Data можно извлечь полезные сведения из социальных сетей, таким образом повысив качество взаимодействия с клиентами и сделав предложения максимально полезными.

Информационная безопасность – это практика предотвращения несанкционированного доступа, раскрытия, использования, искажения, изменения, записи или уничтожения информации в физическом или электронном виде. Её задача – обеспечивать защиту конфиденциальности, целостности и доступности определенных данных. Поиск вредоносной активности обеспечивается путем сбора и анализа огромных объемов данных. Так, в социальных сетях частой проблемой является взлом аккаунта, спам, создание многочисленных так называемых интернет-ботов и различные интернет-мошенничества. Большие данные распознают шаблоны, характерные для мошенников, и собирать значительные объемы данных, чтобы ускорить предоставление нормативной отчетности о безопасности [3].

Big Data также играет роль в финансовой эффективности компании. Среди двух основных способов финансирования выделяют рекламные интеграции и подписку на услуги. Результаты исследований данных повышают эффективность финансовых решений и планирования.

Проведение анализа Big Data.

Существует три способа классификации методов анализа для социальных сетей: по типу данных, по цели и в зависимости от характера задачи. Чаще всего анализируют текст, изображения, видео, аудио. По цели анализа выделяют: предиктивный, описательный, предписывающий и диагностический метод. А в зависимости от характера задач можно проводить визуализацию либо веб-аналитику. Существуют различные методы анализа больших данных, такие как: интеллектуальный анализ данных, прогнозная аналитика, машинное обучение, глубокое обучение, интеллектуальный анализ текста [4].

Алгоритмы интеллектуального анализа данных, такие как нейронные сети, можно научить распознавать нормальную активность пользователя и определять подозрительное поведение. Интеллектуальный анализ данных относят к диагностической аналитике. Такой метод подходит для обнаружения аномалий и обеспечения информационной безопасности, но работает только с неструктурированными данными.

Интеллектуальный анализ текста также является популярным методом для работы с данными социальных сетей. Он заключается в изучении больших коллекций документов для обнаружения новой информации или помощи в ответах на конкретные исследовательские вопросы. К его преимуществу можно отнести то, что он работает как со структурированными, так и с неструктурированными данными, а также то, что существует множество техник и алгоритмов для работы с текстом.

Предиктивная или прогнозная аналитика наиболее актуальна в социальных сетях. Процесс прогнозной аналитики основывается на анализе исторических и текущих данных с целью создания прогнозов на будущее с помощью передовых математических и статистических методов и машинного обучения. Такая аналитика отлично подходит для оценки финансовой эффективности компании. Полученные результаты подскажут о возможных появляющихся трендах и помогут провести планирование [5]. К технологиям прогнозного анализа относят: деревья решений, анализ регрессии, анализ временных рядов и нейронные сети глубокого обучения.

Заключение.

Big Data, термин, который появился относительно недавно, но уже сейчас имеет огромную значимость. Он тесно связан с социальными сетями и затрагивает различные аспекты. Также

нельзя забывать о важности Big Data в информационной безопасности. А для анализа огромного количества данных будут разрабатываться различные методы и технологии, так как с ростом клиентской базы скорость поступления данных будет только расти.

Список литературы

[1] Md. Saifur Rahman. A Systematic Review Towards Big Data Analytics in Social Media / Md. Saifur Rahman, Hassan Reza // Big Data Mining and Analytics / ed. Y. Pan. – Shenzhen, 2022. – P. 228–244.

[2] Зязюлькин, С.П. Использование DQN для обучения агентов игр (Atari 2600) / С.П. Зязюлькин, С.Н. Нестеренков // BIG DATA and Advanced Analytics = BIG DATA и анализ высокого уровня : сб. материалов VI Междунар. науч.-практ. конф. (Республика Беларусь, Минск, 20-21 мая 2020 года): в 3 ч. Ч. 2 / редкол. : В. А. Богуш [и др.]. – Минск : Бестпринт, 2020. – С. 274–280.

[3] Беляк, А. А. Анализ производительности технологии Hadoop / А. А. Беляк, С. Н. Нестеренков // BIG DATA and Advanced Analytics = BIG DATA и анализ высокого уровня: сб. научных статей VII Междунар. науч.-практ. конф. (Республика Беларусь, Минск, 19-20 мая 2021 года): / редкол. : В. А. Богуш [и др.]. – Минск : Бестпринт, 2021. – С. 343–346.

[4] Jayakanna, H. S. A Study on Deep Learning / H. S. Jayakanna, Mrs. M. Raju // International Journal for Research in Applied Science and Engineering Technology – 2022 – Vol. 10, № 11 – P. 961–964.

[5] Kelleher, J. D. Fundamentals of Machine Learning for Predictive Data Analytics: Algorithms, Worked Examples, and Case Studies / J. D. Kelleher, B. Mac Namee, A. D'Arcy. – 2nd ed. – Boston : The MIT Press, 2020. – 856 p.

APPLICATION OF BIG DATA FOR SOCIAL NETWORKS ANALYSIS

Y.I. Bayak

Student of the specialty "Informatics and Programming Technologies" of the Faculty of Computer Systems and Networks of BSUIR

S.N. Nesterenkov

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

D.A. Zhalejko

Software Engineer of BSUIR Department of Information Technology

Department of Computer Science

Faculty of Computer Systems and Networks

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

E-mail: ibayak@gmail.com

Abstract. Big Data or Big Data is a modern term that describes huge volumes of both structured and unstructured information. Social networks, as a source of Big Data, have distinctive features in their collection and analysis methods. And the results obtained are applied in various areas, ranging from information security to improving financial performance.

Keywords: Predictive analytics, information security, data mining, social network.

УДК 004.42: 007.52

ПРИЛОЖЕНИЕ ДЛЯ ДЕМОНСТРАЦИИ ДАННЫХ В ВИДЕ ВИРТУАЛЬНЫХ 3D-ОБЪЕКТОВ С ДОПОЛНЕННОЙ РЕАЛЬНОСТЬЮ



А.И. Калько

Старший преподаватель кафедры ИТиФМД БарГУ,
магистр информатики и вычислительной техники
lexa170594@gmail.com



О.Д. Хадарович

студент 3 курса БарГУ,
lexa170594@gmail.com

А.И. Калько

Окончил Барановичский государственный университет г.Барановичи. Область научных интересов связана с разработкой методов и алгоритмов построения информационно-компьютерных систем, нейросетевых архитектур, распознавание образов.

О.Д. Хадарович

Обучается на третьем курсе в Барановичском государственном университете г.Барановичи. Область научных интересов связана с разработкой мобильных приложений и дополненная реальность.

Аннотация. В данной статье выполнен анализ различных технологий построения дополненной реальности, основанных на маркерах и безмаркерных методах, а также описаны требования к точности отслеживания поз 6DOF для мобильных устройств. На основе проведенного анализа разработано приложение для демонстрации данных в виде виртуальных 3D-объектов с дополненной реальностью, которое учитывает все вышеупомянутые требования и использует маркерные технологии для достижения наилучшей точности отображения виртуальных объектов на реальном мире. Представленное приложение позволяет пользователям с легкостью взаимодействовать с 3D-объектами, получать дополнительную информацию об объектах и сценах, и в целом обогащать свой опыт взаимодействия с реальным миром.

Ключевые слова: Дополненная реальность, виртуальные 3D-объекты, маркерные технологии, мобильные устройства.

Введение.

С самого зарождения в 1950-х годах технологии дополненной реальности развиваются и всё больше проникают в повседневную жизнь. На данный момент дополненная реальность в основном встречается в телефонах: маски на фото в социальных сетях, помещение персонажей в пространство и, конечно же, игры, например, ставшая настоящим феноменом Pokemon Go. Хотя всё вышеперечисленное и относится к сфере развлечений, всё больше компаний понимают важность ниши технологий дополненной реальности и представляют более утилитарные приложения, к примеру:

- AirMeasure – виртуальная рулетка, способная определять расстояния и размеры в 3D-окружении;

- Google Translate умеет переводить текст, который видит камера, в реальном времени;
- Sun Seeker помогает увидеть траекторию солнца на местности в любой день года;
- Google Sky Map помогает узнать, какие звезды сейчас видно на небе.
- IKEA Place позволяет «примерить» мебель из каталога прямо к интерьеру комнаты.

Необходимо также отметить перспективность AR технологий в сфере образования (возможность в будущем увидеть в вузах и школах виртуальные интерактивные иллюстрации) и медицины (максимально наглядное обучение студентов медвузов; визуализация данных прямо на пациенте, вместо расставленных вокруг экранов; максимально наглядное УЗИ).

Актуальность.

В последние годы дополненная реальность становится все более популярной, и это вызывает необходимость создания новых технологий для улучшения ее функциональности и привлекательности для пользователей. Одной из таких технологий является приложение для демонстрации данных в виде виртуальных 3D-объектов с дополненной реальностью. Это приложение позволяет визуализировать данные в форме объемных объектов, что может быть особенно полезно для обучения и понимания сложных концепций.

Кроме того, разработка приложений с дополненной реальностью требует высокой точности отслеживания поз 6DOF объектов реального мира, что приводит к развитию новых технологий для повышения точности и удобства использования. Такие приложения могут быть использованы в различных отраслях, таких как образование, медицина, игровая индустрия и др.

Таким образом, разработка приложения для демонстрации данных в виде виртуальных 3D-объектов с дополненной реальностью является актуальной темой и может иметь многообещающие перспективы в будущем.

Перспективы использования дополненной реальности для отображения данных.

Дополненная реальность – воспринимаемая смешанная реальность, создаваемая с помощью компьютера с использованием «дополненных» элементов воспринимаемой реальности, когда виртуальные объекты проецируются на реальное окружение [1].

Существует несколько видов построения дополненной реальности, которые основаны на:

- Маркерах (таргетах);
- Координатах местоположения в пространстве.

В настоящее время существует множество технологий, позволяющих создавать виртуальные 3D-объекты с дополненной реальностью и использовать их в различных приложениях. Одним из способов демонстрации данных с помощью виртуальных 3D-объектов является использование безмаркерных технологий, которые основываются на получении информации от датчиков в мобильном устройстве пользователя, таких как GPS-приёмник, акселерометр, гироскоп и т.д. Эти данные позволяют достаточно точно определить положение пользователя в пространстве и создать виртуальные 3D-объекты с дополненной реальностью, отображающиеся на экране мобильного устройства.

В отличие от безмаркерных технологий, маркерные технологии используют специальные объекты, называемые маркерами, для определения положения и ориентации в пространстве. Эти маркеры распознаются специальными приложениями на мобильных устройствах и позволяют создавать виртуальные 3D-объекты, которые будут отображаться на маркере с учетом его положения и ориентации. Такой подход позволяет достичь высокой точности и реалистичности визуализации, так как виртуальный объект будет отображаться на реальном физическом объекте.

В обоих случаях, при использовании дополнительных приспособлений, таких как графические фильтры и/или e-модели высокого качества, можно добиться еще более реалистичной визуализации и создать виртуальные 3D-объекты, которые будут трудно отличимы от остальных элементов в реальном мире. Такие технологии могут использоваться в различных приложениях, например, в маркетинге и рекламе, образовании и туризме. Дополненная реальность на мобильных устройствах требует точного отслеживания поз 6DOF (6 степеней свободы) объектов реального мира. 6DOF (6 степеней свободы) относятся к свободе перемещения твердого объекта в трехмерном пространстве [2]. В частности, тело может свободно двигаться вперед / назад, вверх / вниз, влево / вправо в сочетании с вращением вокруг трех перпендикулярных осей.

Для достижения высокой надежности и производительности мобильных приложений дополненной реальности необходимо обеспечить удобную позу слежения, которая должна быть адаптивной и не требовать высоких технических характеристик. Обычно для этого достаточно одной камеры, установленной на мобильном устройстве, которая одновременно используется в

качестве видефона и для отслеживания позы камеры относительно окружающей среды в режиме реального времени.

Для обеспечения надежного отслеживания позы в условиях ограниченных вычислительных ресурсов мобильных устройств, часто используются прямоугольные доверительные маркеры, которые позволяют Vuforia, одной из наиболее распространенных библиотек отслеживания маркеров, обнаруживать и отслеживать визуальные объекты в режиме реального времени на широком спектре мобильных платформ.

Однако для того, чтобы маркер мог быть обнаружен и отслежен надежно, он должен иметь большое количество признаков и сбалансированное распределение объектов в изображении. Важным критерием отслеживания является локальный контраст маркера, который должен быть достаточно сильным, чтобы обеспечить его обнаружение и отслеживание в условиях различных освещенных и текстурных фонов. Поэтому при разработке приложений дополненной реальности необходимо учитывать эти факторы для обеспечения высокой надежности и производительности приложения.

Даже если на изображении много объектов с яркими контурами, повторяющиеся паттерны могут затруднить процесс отслеживания, так как они не имеют уникального паттерна для обнаружения.

После обнаружения доверительных маркеров, мы получаем замкнутые полигоны. Однако, чтобы создать маркер, они должны быть преобразованы в прямоугольные формы. На следующем этапе используются контуры для определения прямоугольных форм. Vuforia предоставляет несколько типов трекеров (например, Image Tracker), которые расширяют базовый класс трекера и используются для проверки паттернов.

Для каждого кадра камеры в приложении вызывается функция начала класса *Renderer*, которая возвращает объект состояния. Это состояние отображает отслеживаемые объекты, которые в данный момент отслеживаются трекером (например, Image Tracker). Когда обнаруживается паттерн, который нужно найти, последним шагом перед рендерингом является оценка позы относительно маркера. Подобно доверительному обнаружению и проверке паттернов, существует несколько различных алгоритмов оценки позы, которые все используют одну и ту же базовую концепцию: сначала создается начальная догадка, которая оценивает приблизительное положение и ориентацию отслеживаемого объекта относительно камеры. Затем эта первая оценка уточняется итеративно до тех пор, пока не будут выполнены конкретные критерии качества или не будет достигнуто максимальное число итераций. Полученное матричное представление затем преобразуется в матрицу OpenGL, чтобы иметь возможность визуализировать модель в представлении пользователя [3].

Для использования возможностей Vuforia необходимо зарегистрироваться на официальном сайте <https://developer.vuforia.com/>, создать базу данных для таргетов (целей, при обнаружении которых будет происходить определенное событие) и добавить необходимое изображение, которое будет являться таргетом, представлено на рисунке 1.



Рисунок 1. Таргет для приложения

Определённые сайтом маркеры геометрии, представлены на рисунке 2.



Рисунок 2. Маркеры геометрии на таргете

Данный таргет получил 4 балла из 5 по методам оценки Vuforia.

Далее необходимо скачать .package файл базы данных и добавить его в проект Unity, после чего необходимо ввести код-лицензию в конфигурацию Vuforia [4]. Затем нужно добавить определённые компоненты проекта для AR приложения – AR Camera и ImageTarget – и назначить действия на события OnTargetFound() и OnTargetLost() второго компонента.

Для удобства создадим пользовательский интерфейс и назначим кнопкам необходимые действия. Интерфейс представлен на рисунке 3.

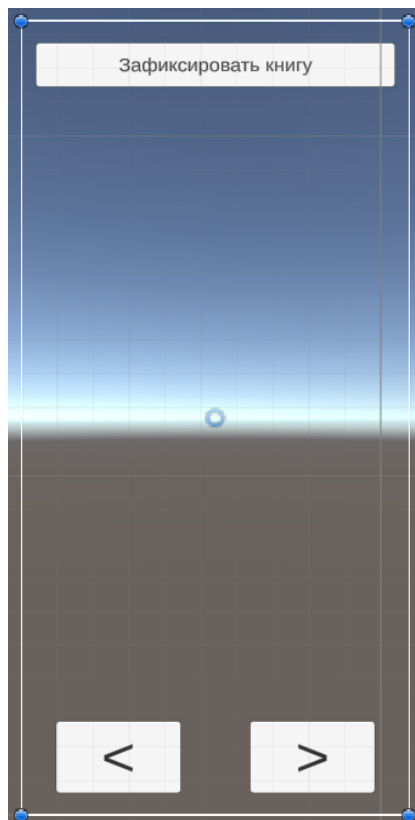


Рисунок 3. Пользовательский интерфейс

Также в ходе разработки приложения для реализации 3D-объекта книги с анимацией пролистывания страниц был использован бесплатный ассет «Book – Page Curl Version 2.0», разработанный Abdullah Aldandarawy, представлен на рисунке 4 [5].



Рисунок 4. Официальное изображение на странице «Book – Page Curl» в Unity Asset Store

Данный ассет модифицируется путем добавления всех страниц книги и создания интерактивного объекта, который позволит пользователям взаимодействовать с ним, перелистывая страницы не только с помощью кнопок, но и нажатием на экран в нижней части страницы. После чего объект книги полностью принял необходимый нам вид, представленный на рисунке 5.

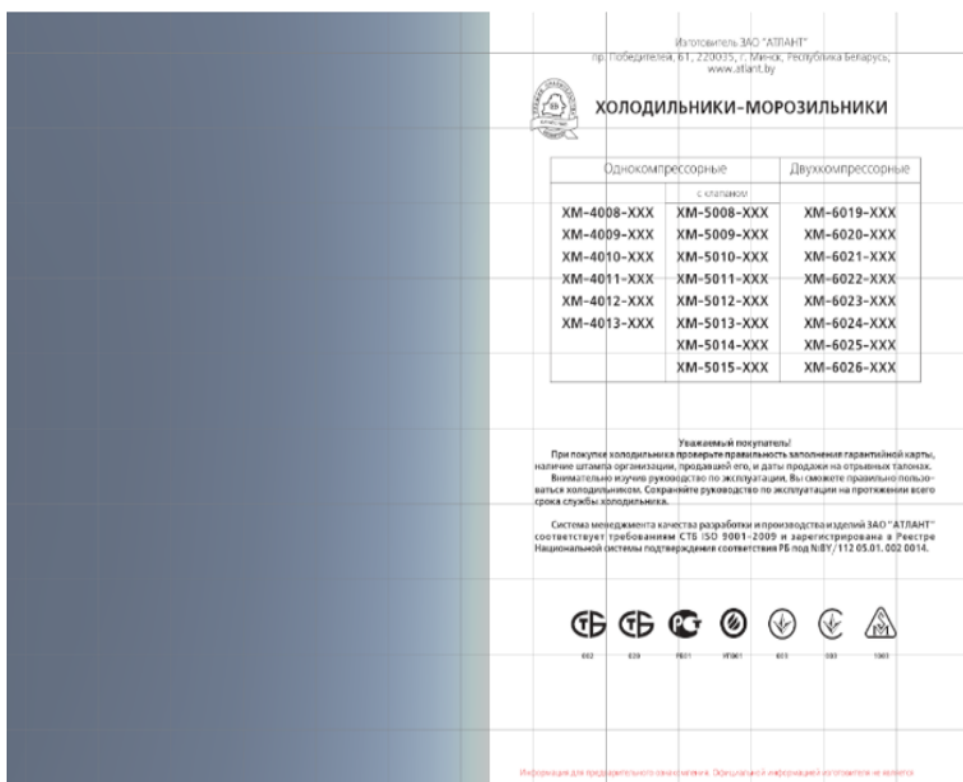


Рисунок 5. Итоговый вид 3D-объекта книги

В качестве книги было выбрано «Руководство по эксплуатации холодильников-морозильников» производителя «Atlant».

Также реализована возможность закрепления книги на экране в 2D-формате для удобства чтения.

Руководство пользователя Android приложения.

Приложение разработано для системы Android. Итоговый размер приложения составляет 57.0 МБ.

После успешной установки приложения на рабочем столе появится ярлык для запуска приложения, представленный на рисунке 6.



Рисунок 6. Ярлык для запуска приложения

При первом запуске приложение запросит доступ к камере устройства. После получения прав доступа к камере приложения готово к работе. При обнаружении таргета приложением в поле действия камеры устройства, на месте этого таргета на экране будет отображён 3D-объект книги.

При запуске приложения на экране появится изображение с камеры устройства. При наведении камеры устройства на таргет, на экране отобразится 3D-объект книги. Таргетом также будет считаться чёрно-белая (и любой другой цветовой гаммы) версия изображения, так приложения реагирует лишь на геометрию объектов, изображённых на таргете, представленных на рисунке 7. Если в поле видимости камеры приложение не находит таргет, то ничего отображено не будет.

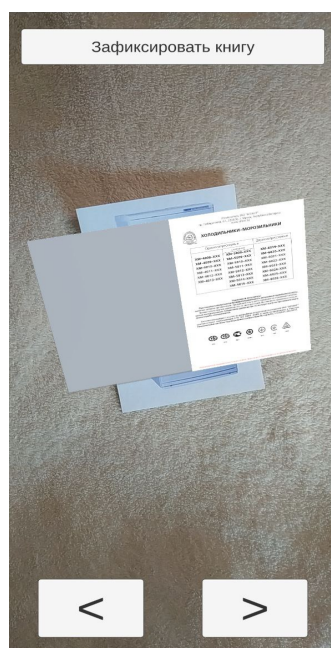


Рисунок 7. Отображение 3D-объекта книги при обнаружении приложением таргета

При умеренных отклонениях камеры устройства от нормали к таргету 3D-объект книги не будет пропадать, а в такой же степени отклонится. Тем самым будет соблюдено отношение нормали камеры устройства по отношению к таргету и отображения 3D-объекта по отношению к экрану устройства, представленное на рисунке 8.



Рисунок 8. Отклонение камеры устройства от нормали к таргету

В результате исследования разработано мобильное приложение с дополненной реальностью в Unity на языке C#, которое включает в себя трёхмерную анимацию книги-инструкции.

Заключение.

В заключении статьи хочется подчеркнуть, что приложения для демонстрации данных в виде виртуальных 3D-объектов с дополненной реальностью представляют собой мощный инструмент для визуализации и анализа данных. Проанализированы различные технологии создания дополненной реальности, основанные на маркерах и без маркеров, и сформированы требования к позе слежения для мобильных устройств.

Кроме того, разработан прототип приложения для демонстрации данных в виде виртуальных 3D-объектов с дополненной реальностью, используя маркерную технологию. Созданное приложение предоставляет возможность визуализации данных в удобном и интерактивном формате, а также может быть использовано для образовательных и научных целей.

Таким образом, работа демонстрирует перспективность использования технологий дополненной реальности для визуализации данных и создания новых приложений в области образования, науки и бизнеса.

Список литературы

[1] Дополненная реальность [Электронный ресурс]. URL: https://ru.wikipedia.org/wiki/Дополненная_реальность (Дата обращения: 21.03.2022).

[2] Понятия виртуальной, дополненной и смешанной реальности [Электронный ресурс]. URL: <https://studfile.net/preview/16387411/page:10/> (Дата обращения: 21.03.2022).

[3] Маркеры и их роль в технологии дополненной реальности для мобильных устройств на примере фреймворка Vuforia [Электронный ресурс]. URL: <https://science-engineering.ru/ru/article/view?id=1340> (Дата обращения: 21.03.2022).

[4] Мальчиков, С. Ю. Разработка оболочки для прохождения тестирования на языке java / С. Ю. Мальчиков, А. И. Калько // Проблемы развития регионов в условиях модернизации экономики, общества и образования : Сборник трудов IV Международной научно-практической конференции, Балаково, 05 апреля 2018 года. – Балаково: Национальный исследовательский ядерный университет "МИФИ", 2018. – С. 68-71.

[5] Мазур, С. А. Разработка приложения SQUARES для развития памяти / С. А. Мазур, С. Н. Шапутько, А. И. Калько // Содружество наук. Барановичи-2019 : материалы XV Междунар. науч.-практ. конф. молодых исследователей, Барановичи, 16 мая 2019 г. : в 2 ч. / М-во образования Респ. Беларусь, Баранович. гос. ун-т, Студенч. науч. о-во БарГУ ; [редкол. : В. В. Климук (гл. ред.) и др.]. – Барановичи, 2019. – Ч. 1. – С. 77–78.

AN APPLICATION FOR DEMONSTRATING DATA IN THE FORM OF VIRTUAL 3D OBJECTS WITH AUGMENTED REALITY

A.I. Kalko

*Senior Lecturer of the Department of IT and FM at
BarSU, Master of Computer Science and Computer
Engineering*

O.D. Khadarovich

3rd year student BarSU,

*Department of Information Technologies and Physical and Mathematical Disciplines
Faculty of Engineering
Baranavichy State University, Republic of Belarus
E-mail: lexa170594@gmail.com*

Abstract. This article analyzes various augmented reality technologies based on marker and markerless methods and describes the requirements for 6DOF pose tracking accuracy for mobile devices. Based on the analysis, an application for demonstrating data in the form of virtual 3D objects with augmented reality has been developed, which takes into account all the aforementioned requirements and uses marker-based technologies to achieve the best accuracy of displaying virtual objects in the real world. The presented application allows users to easily interact with 3D objects, obtain additional information about objects and scenes, and overall enhance their experience of interacting with the real world.

Keywords: Augmented reality, virtual 3D objects, marker-based technologies, mobile devices.

УДК: 621.391

МЕТОДЫ ОЦЕНКИ РИСКА НАРУШЕНИЙ ИНФОРМАЦИОННОЙ БЕЗОПАСНОСТИ ПРИ ОБЛАЧНЫХ ВЫЧИСЛЕНИЯХ В ФИНАНСОВЫХ И БАНКОВСКИХ СИСТЕМАХ



Ж.Б. Балтаев
Доцент кафедры
«Информационных систем и
цифровых технологий» ТФИ,
PhD, доцент
j_baltayev@tfi.uz



С.Д. Арзикулов
Старший преподаватель
кафедры «Информационные
системы и цифровые
технологии» ТФИ,
arzikulov.sunnatulla@gmail.com



М.А. Миржамолова
Ассистент кафедры
«Информационных систем и
цифровых технологий» ТФИ,
dizafa@icloud.com

Ж.Б. Балтаев

Обучался в Ташкентском университете информационных технологий. Закончил бакалавриатуру в 2010 году, а магистратуру в 2013 году. 2020 году защитил диссертацию по специальности «05.04.01 - Телекоммуникационные и компьютерные системы, телекоммуникационные сети и информационные устройства. Распределение информации». Получил научную степень доктора философии по техническим наукам (PhD). Трудовую деятельность начал в Ташкентском университете информационных технологий (2013-2020), продолжил работу в Ташкентский государственный транспортный университет (2020-2023), в данный момент работает в Ташкентском финансовом институте (2023-н.в.).

С.Д. Арзикулов

Обучался в Ташкентском государственном университете (ныне Национальный университет Узбекистана имени Мирзо Улугбека) факультет прикладная математика и механика. Закончил в 1989 году, в специальности прикладной математика. Трудовую деятельность начал в Научном исследовательском институте «Алгоритм» при НПО «Кибернетика» (1989-2004), продолжил работу в Ташкентском университете информационных технологий (2004-2021), в данный момент работает в Ташкентском финансовом институте (2022-н.в.).

М.А. Миржамолова

В 2011 году я училась в Ташкентском Государственном Педагогическом Университете (бакалавр). В 2017 году я училась в Ташкентском университете информационных технологий имени Мухаммад ал Хоразми (степень магистра). С 5 сентября 2022 года работаю ассистентом кафедры «Информационных систем и цифровых технологий» Ташкентского финансового института.

Аннотация: Оценка рисков информационной безопасности (ИБ) в среде облачных вычислений (СОВ) являются важным методом в обеспечении ИБ в целом. Управление рисками ИБ представляет собой непрерывный процесс идентификации, оценки и минимизации рисков от реализации угроз ИБ. За последние несколько лет в отрасли информационных технологий получила развитие новая парадигма – «облачные вычисления». Облачные вычисления -это модель предоставления повсеместного и удобного сетевого доступа (по мере необходимости) к общему пулу конфигурируемых вычислительных ресурсов (например, сетей, серверов, систем хранения, приложений и сервисов), которые могут быть быстро предоставлены и освобождены с минимальными усилиями по управлению и необходимостью взаимодействия с провайдером услуг (сервис-провайдером). Далее по после того как будет произведена оценка, необходимо сопоставить уровень риска и контрмеры которые на данный момент используются владельцем. После чего принимается решение о критичности положения и необходимости тех или иных действий по безопасности

Ключевые слова: Информационной безопасности, среде облачных вычислений, управление риск, эталонная модель, модель с нарушением, AnyLogic, NetworkEnter, NetworkMoveTo, NetworkRelease, Delay, Selectoutput, Sink

Введение.

В финансовых и банковских системах облачные вычисления - это модель, обеспечивающая (по мере необходимости) повсеместный и удобный сетевой доступ к общему пулу настраиваемых вычислительных ресурсов (например, сетей, серверов, систем хранения, приложений и сервисов), которые могут быть быстро предоставлены и освобождены с минимальными усилиями по управлению и необходимостью взаимодействия с провайдером услуг (сервис-провайдером).

Анализ результатов исследований, ведущихся в направлении обеспечения ИБ в среде облачных вычислений (СОВ) показывает, что в настоящее время не до конца решены вопросы, связанные с оценкой ИБ. В первую очередь это касается используемых моделей и методов оценки ИБ в среде облачных вычислений на основе риск – ориентированного подхода.

Отсутствие показателей оценки ИБ создает неопределенность в понимании того:

- в каком состоянии находится ИБ СОВ;
- какова эффективность методов и средств защиты в отношении повышения ИБ.

Среди первоочередных задач по выработке подходов к обеспечению информационной безопасности облаков необходимо отнести следующие:

1. Исследование вопросов наличия уязвимостей виртуальных сред;
2. Построение классификации уязвимостей виртуальных сред;
3. Разработку типовой модели угроз для виртуальных сред;

На сегодняшний день существуют два подхода к оценке ИБ СОВ:

- качественный;
- количественный.

В системе управления процесс оценки риска является составной частью процесса управления рисками. Под управлением рисками понимается оценка и уменьшение рисков, которые могут воздействовать на СОВ.

Качественный подход к оценке рисков ИБ СОВ. Качественные методологии, используемые для оценки ИБ СОВ, исходят из того, что зачастую потенциальные потери неосязаемы, поэтому опасность или риск ИБ СОВ в результате реализации той или иной угрозы нельзя представить в каком-либо выражении.

В настоящее время известно множество методов качественной оценки рисков, большинство из которых построено на использовании табличных методов. Другие подходы предлагают графическое изображение дерева решения, которое показывает распределение вероятностей самых общих случаев.

Ни одно из указанных направлений не способно в отдельности обеспечить получение объективной, представительной и конструктивной оценки ИБ СОВ.

Количественный подход к оценке рисков ИБ СОВ. Рациональное применение разнообразных средств защиты информации в СОВ с учетом критерия экономической эффективности – это сложная научно-техническая задача. Умножая величину ущерба на число вероятных реализаций угрозы, получают грубый уровень риска от данной угрозы.

При оценивании рисков учитываются такие факторы, как ценность ресурсов, значимость угроз и уязвимостей, эффективность имеющихся и планируемых средств защиты.

Распространенные методики анализа рисков:

- методики, использующие оценку риска на качественном уровне (например, по шкале «высокий», «средний», «низкий»). К таким методикам, в частности, относится FRAP;
- количественные методики (риск оценивается через числовое значение, например размер ожидаемых годовых потерь). К этому классу относится методика RiskWatch;
- методики, использующие смешанные оценки (такой подход используется в CRAMM, методике Microsoft и т.д.).

созданию имитационной модели для оценки рисков ИБ СОВ, основной идеей которой является реализация эксперимента, являющегося одним из методов оценки риска ИБ. Было принято решение произвести оценку риска через атаки, реализуемые в данном эксперименте.

В ходе анализа существующих угроз, уязвимостей и атак было принято решение для сравнительного анализа произвести оценку СОВ с точки зрения ее безопасности (защищенности) смоделировать модель и рассмотреть три случая:

1. «Эталонная модель», указывающую на нормальное функционирование СОВ с отсутствием всяческих воздействий со стороны злоумышленника;
2. «Модель с нарушением ИБ» не удовлетворяющая нормативным уровням защищенности, заданным в ходе моделирования и эксперимента;
3. «Модель с нарушением ИБ» удовлетворяющая нормативным уровням защищенности, заданным в ходе моделирования и эксперимента.

Рассмотрим процесс мультиагентного моделирования в среде имитационного моделирования AnyLogic с целью анализа наиболее распространённых угроз ИБ по средствам таких механизмов, как атаки типа «отказ в обслуживании».

В краткости рассмотрим методику реализации данных атак в среде AnyLogic.

Архитектура исследуемого фрагмента простейшей сети показана на рисунке, и включает в состав: 3 ISP, к которым обращаются 6 пользователей, в том числе 3 злоумышленника, также присутствует система обеспечивающая информационную безопасность СОВ.

Для построения модели используем библиотеку сетевые элементы палитры Enterprise Library. Рассмотрим подробно сеть пользователей и злоумышленников. В пользовательской сети присутствует три клиента, которые генерируют запрос на среду облачных вычислений. Ресурсы каждого клиента отличаются, однако они объединяются в пользовательскую сеть и создают общую нагрузку на систему облачных вычислений.

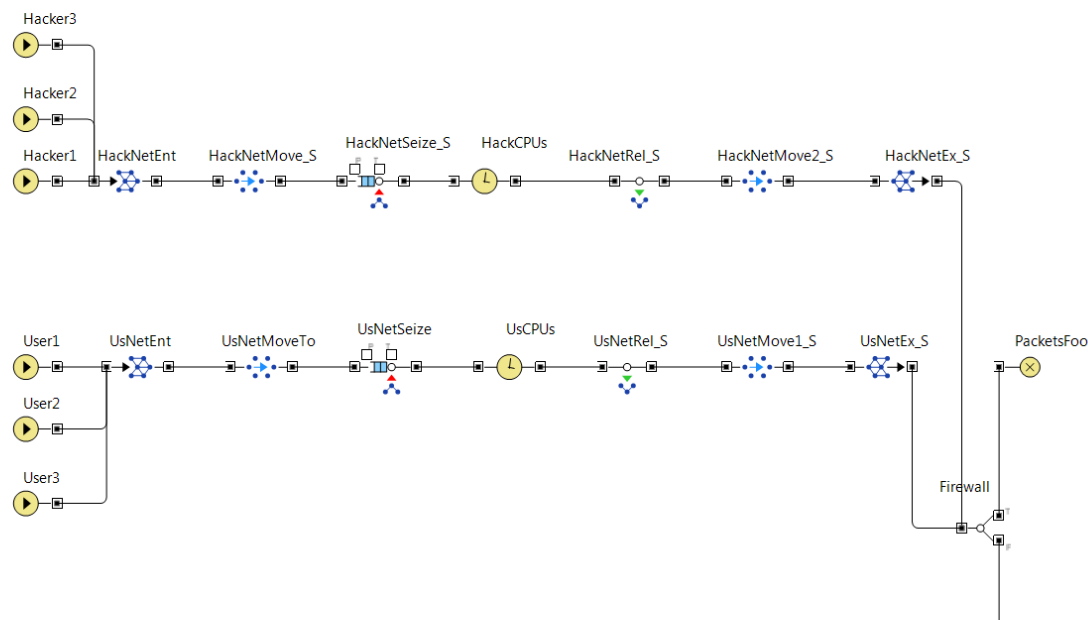


Рисунок 1. Пользовательская сеть и сеть злоумышленников в AnyLogic

Для генерации запросов используется элемент Source, который генерирует запросы. При помощи данного объекта можно сконфигурировать объект так, чтобы он создавал заявки других типов, указав конструктор нужного класса в параметре Новая заявка, а также задать действие, которое должно выполняться перед тем, как новая заявка покинет объект, и связать с заявкой

определенную фигуру анимации. Есть несколько способов задания того, сколько заявок и когда должен создавать этот объект.

Элемент NetworkEnter. Регистрирует заявку в сети и помещает ее в заданный узел сети. После добавления в сеть заявка может перемещаться по сети и использовать сетевые ресурсы. Заявка не может одновременно находиться сразу в нескольких сетях, поэтому перед добавлением в другую сеть она должна быть вначале удалена из текущей сети с помощью объекта NetworkExit.

NetworkMoveTo. Перемещает заявку в новое место сети. Если к заявке присоединены какие-то ресурсы, то они перемещаются вместе с заявкой. При этом независимо от скорости ресурсов перемещаться такая группа из заявки и ее ресурсов будет со скоростью заявки. Время, которое заявка проведет в этом объекте, будет равно длине кратчайшего из возможных путей из текущего местоположения заявки в узел назначения, поделенной на скорость заявки.

NetworkSeize. Объект захватывает для заявки заданное количество сетевых ресурсов. При необходимости может пересылать захваченные ресурсы в заданное место сети и присоединять их к заявке. Объект можно рассматривать как очередь Queue заявок, ожидающих ресурсов, за которой могут следовать объекты NetworkSendTo и NetworkAttach.

Delay. Задерживает заявки на заданный период времени. Время задержки вычисляется динамически, может быть случайным, зависеть от текущей заявки или от каких-то других условий. Одновременно могут быть задержаны сразу несколько заявок. Заявки задерживаются независимо друг от друга – время задержки вычисляется отдельно для каждой заявки. Как только время задержки истекает, заявка тут же покидает объект. Если объект Delay заполнен полностью, то новую заявку он не примет, и в этом случае.

NetworkRelease. Освобождает все или какие-то определенные сетевые ресурсы (ранее захваченные заявкой с помощью объекта NetworkSeize). В случае освобождения каких-то определенных ресурсов, они выбираются из общего числа захваченных ресурсов согласно заданному Списку ресурсов. Если освобождается движущийся ресурс, Вы можете как оставить его в текущем узле сети, так и вернуть его в его базовое местоположение. Сразу после освобождения сеть проверит, не были ли только что освобожденные ресурсы запрошены другими заявками, и если да, то движущийся ресурс будет тут же захвачен другой заявкой и не отправится в свое базовое местоположение, даже если у объекта будет выбрана соответствующая опция.

Selectoutput используется как Firewall. Используется как система безопасности к ресурсам облачных вычислений. Обрабатывает запросы и с определенной вероятностью перехватывает хакерский трафик и блокирует его.

Sink. Используем для сбора статистических данных. Уничтожает поступившие заявки. Обычно используется в качестве конечной точки потока заявок. Для того, чтобы заявки удалялись из модели и уничтожались, нужно соединить выходной порт последнего блока процессной диаграммы с портом объекта Sink.

Оценка временного риска по результатам исследования системы в трех случаях:

- «Эталонная модель», указывающую на нормальное функционирование СОВ с отсутствием всяческих воздействий со стороны злоумышленника;
- «Модель с нарушением ИБ» не удовлетворяющая нормативным уровням защищенности, заданным в ходе моделирования и эксперимента;
- «Модель с нарушением ИБ» удовлетворяющая нормативным уровням защищенности, заданным в ходе моделирования и эксперимента.

Полученные результаты запишем в сводную таблицу соответственно.

Таблица 1. Результаты исследования

Параметр	Характеристика		
$R_{УТР\ 1}$	0,03	0,12	0,46
Всего полезных запросов	70	62	46
Перехваченные firewall'ом хакерские запросы	2	18	38
Всего пройдено хакерских запросов		17	24
$R_{риск}$	0,05	0,2	0,77
$R_{без}$	0,95	0,8	0,23

По результатам моделирования можем наблюдать, что в условиях реализации угроз через модель нарушителя можно получить количественные оценки ИБ, в частности риски нарушений, реализуемость угрозы. Благодаря построенной модели можно оценить риск с учетом временных характеристик. Полученные результаты довольно критичны, их необходимо более детально изучить. Полученная количественная оценка рисков ИБ позволило установить категории риска и установить наиболее приемлемые характеристики.

Таким образом, в таблице укажем категории риска, то есть признаки, по которым можно будет в дальнейшем численно характеризовать ИБ СПД. Введем четыре базовых категорий, выраженные через ущерб и угрозы соответственно.

Таблица 2. Категорирование рисков нарушения ИБ СПД

Качественная градация	Количественная градация, $R_{риск}$	Управление	Количественная градация, $R_{без}$
Риск отсутствует	до 0,05	допустимый	от 0,95 до 0,999
Низкий уровень	от 0,0501 до 0,1	допустимый	до 0,9
Средний уровень	от 0,101 до 0,2	не допустимый	от 0,899 до 0,8
Высокий уровень	от 0,2 до 0,779	не допустимый	0,799 до 0,221
Критический Риск	от 0,779 до 0,999	не допустимый	0,221 до 0,001

Далее по после того как будет произведена оценка, необходимо сопоставить уровень риска и контрмеры которые на данный момент используются владельцем. После чего принимается решение о критичности положения и необходимости тех или иных действий по безопасности.

Список литературы

- [1]. Baltayev, J., Boltayev, S. (2023). Model and Methods for Detecting Undetected Errors Us-ing a Signature Analyzer. In: Guda, A. (eds) Networked Control Systems for Connected and Automated Vehicles. NN 2022. Lecture Notes in Networks and Systems, vol 509. Springer, Cham. https://doi.org/10.1007/978-3-031-11058-0_112
- [2]. Djuraev R.X., Baltaev J.B., Xasanov O.A. Increasing the efficiency of diagnosing micro-processor devices based on multichannel signal analysis means. Тошкент, ICISCT2020.
- [3]. Djuraev R.X., Baltaev J.B., Badalov J.I. Study of the method of compact testing of technical means of data transmission networks Toshkent, ICISCT2020
- [4]. Amirsaidov U.B., Abbasxanova X.Yu., Baltaev J.B. Metodi otsenki nadejnosti seti peredachi dannix s uchuyotom vozdeystviya vneshnix faktorov VESTNIK TashGTU, 4/2014g., str 27-31
- [5]. Djuraev R.X., Djabbarov Sh.Yu., Baltaev J.B. «Sistemi texnicheskogo obslujivaniya i ek-spluatatsii setey telekommunikatsii». Uchebnik.-T.: "Aloqachi".2019, 234 s.
- [6]. Djuraev R.X., Baltaev J.B., Alimov U.B. Methods of Determining Reference Signals for One and Multichannel Signatural Analyzer of Microprocessor Systems, Science Publishing Group, Communications USA № 6(1) 2018. –R. 20-24
- [7]. Djuraev R.X., Baltaev J.B. Investigated Methods of Improving the Yefficiency of Diagnos-ing Microprocessor Devices of Data Transmission Systems Based on Multi-Channel Signa-ture Analysis, Science Publishing Group, Communications USA № 7(1) 2019. –R. 13-24 rr.
- [8]. Djuraev R.X., Djabbarov Sh.No, it's not., Baltaev J.B. It is a modeling program for diagnos-ing digital devices using an eight-channel signature method. Agency of property of the Re-public of Uzbekistan. Certificate of official registration of the program created for electronic computing machines. № DGU 04123. 16.12.2016 y.

[9]. Djuraev R.X., Djabbarov Sh.Yu., Baltaev J.B. It is a modeling program for diagnosing digital devices using the sixteen-channel signature method. Agency of property of the Republic of Uzbekistan. Certificate of official registration of the program created for electronic computing machines №DGU 04124. 16.12.2016 y.

[10]. Djuraev R.X., Baltaev J.B., Toshtemirov T.Q. Program of detection of errors not detected by the signature analyzer "Agency of real estate of the Ministry of Justice of the Republic of Uzbekistan" DGU 10922.

[11]. Анкудинов И.В. Микропротессорные системы. Архитектура и проектирование. Учебное пособие. Санкт-Петербург 2003

[12]. Антошина И.В., Котов Ю.Т. Микропротессоры и микропротессорные системы (аналитический обзор). Москва 2005 г

[13]. Безуглов Д.А., И.В. Калиенко И.В. сифровие устройства и микропротессоры. – Ростов – на – Дону: Феникс, 2006. – 480 с.

[14]. Заиналабедин Наваби, Дигитал Систем Тест анд Тестапле Десигн: Усинг ХДЛ Моделс анд Ар-читестурес, Спрингер, 2011.

[15]. Бестугин А. Р., Богданова А. Ф., Стогов Г. В. Контроль и диагностирование телекоммуникационных сетей - СПб: Политехника, 2003. 174 с.: ил.

METHODS FOR ASSESSING THE RISK OF INFORMATION SECURITY BREACHES IN CLOUD COMPUTING IN FINANCIAL AND BANKING SYSTEMS

J.B. Baltayev

PhD, Associate Professor

"Information Systems and Digital Technologies" of the TFI

S.D. Arzikulov

Senior lecturer of the Department

"Information Systems and Digital Technologies" of the, TFI

M.A. Mirjamolova

Assistant of the Department of

"Information Systems and Digital Technologies" of the TFI

Department of Information Systems and Digital Technologies

Faculty of taxes and insurance

Tashkent Financial Institute, Republic of Uzbekistan

E-mail: j_baltayev@tfi.uz

Abstract: Information security risk assessment (IS) in the cloud computing environment (SOV) is an important method in ensuring information security in general. Information security risk management is a continuous process of identification, assessment and minimization of risks from the implementation of information security threats. Over the past few years, a new paradigm has been developed in the information technology industry – "cloud computing". Cloud computing is a model for providing ubiquitous and convenient network access (as needed) to a common pool of configurable computing resources (for example, networks, servers, storage systems, applications and services) that can be quickly provided and released with minimal management effort and the need to interact with a service provider (service provider). Further, after the assessment is made, it is necessary to compare the level of risk and the countermeasures that are currently being used by the owner. After that, a decision is made on the criticality of the situation and the need for certain security actions

Keywords: Information security, cloud computing environment, risk management, Reference model, model with violation, AnyLogic, NetworkEnter, NetworkMoveTo, NetworkRelease, Delay, Selectoutput, Sink

УДК 004.021:004.75

АНАЛИЗ БЫСТРОДЕЙСТВИЯ ПРИ РЕАЛИЗАЦИИ ПАРАЛЛЕЛЬНЫХ ВЫЧИСЛЕНИЙ ДЛЯ РЕШЕНИЯ ЗАДАЧ ОПТИМИЗАЦИИ В СИСТЕМЕ ЛОГИЧЕСКОГО ПРОЕКТИРОВАНИЯ



И.П. Логинова

Старший научный сотрудник ОИПИ НАНБ,
кандидат технических наук, доцент
irilog@mail.ru

И.П. Логинова

Окончила Белорусский государственный университет, кандидат технических наук, доцент. Работает в ОИПИ НАН Беларуси в должности старшего научного сотрудника и в Белорусском государственном университете информатики и радиоэлектроники в должности доцента. Круг научных интересов: программирование, логическое проектирование и верификация цифровых схем, виртуализация, реализация параллельных алгоритмов.

Аннотация. В работе рассматривается возможность реализации параллельных вычислений в системе FLC-2, предназначенной для технологически независимой оптимизации функциональных и структурных описаний логических схем. На примере работы некоторых программ логической оптимизации демонстрируется методика по организации параллельной работы, в том числе и тех программ, которые представлены исполняемым кодом. Проведены сравнения быстродействия различных вариантов запуска параллельной работы программ оптимизации на потоке известных примеров. Приводятся результаты вычислительных экспериментов на многоядерной системе с общей памятью, подтверждающие эффективность использованного подхода при организации параллельных вычислений в системах логического проектирования для решения трудоемких задач.

Ключевые слова: многоядерный процессор с общей памятью, параллельные процесс, планировщик задач, ускорение, дизъюнктивная нормальная форма (ДНФ), логическая оптимизация, СБИС.

Введение.

В настоящее время решение сложных вычислительных задач проводится на многопроцессорных и многоядерных вычислительных системах, что предусматривает разработку параллельных алгоритмов, учитывающих архитектуру вычислительных систем. При этом основной подход состоит в создании новых алгоритмов или переработке существующих последовательных алгоритмов. Последнее хорошо работает, когда задача может быть разделена по данным. Осуществить распараллеливание комбинаторных задач гораздо сложнее, поскольку для таких задач, как правило, невозможно реализовать одновременное выполнение различных частей алгоритма в одной программе или провести разделение исходных данных. Практически все задачи из области логического проектирования сверхбольших интегральных схем (СБИС), в частности, задачи логической оптимизации, носят комбинаторный характер и имеют экспоненциальную сложность. Особенностью логико-комбинаторных задач является то, что основными объектами преобразований для них являются такие структуры, как булевы переменные, булевы векторы, булевы n -элементные матрицы и n -троичные матрицы, графы [1-3]. Для задач этого класса характерно то, что объемы перерабатываемой информации при поиске решений относительно невелики, но сами процессы переработки зачастую сопряжены с необходимостью полного перебора и анализа значительного числа вариантов промежуточных решений. Поэтому становится проблематично решать такие задачи большой размерности за приемлемое время.

Актуальность.

Тривиальные параллельные алгоритмы для задач проектирования СБИС, в которых возможно распараллеливание на нужное количество процессоров, описаны в литературе [4-7]. Так, в книге [4], описаны параллельные алгоритмы для размещения и трассировки, проверки соблюдения правил проектирования, логического синтеза, генерации тестов, моделирования ошибок в схемах, моделирования на поведенческом уровне. Исследования, представленные в работах [5-7] показали, что задачи в данной проблемной области, в первых, не допускают автоматического распараллеливания, во-вторых, распараллеливание посредством преобразования последовательных комбинаторно-логических алгоритмов не позволяет получить эффективный параллельный алгоритм. Поэтому разработка параллельных комбинаторно-логических алгоритмов в этой сфере является деятельностью, в которой успехи довольно редки. Тем не менее, возможности использования в САПР высокопроизводительных вычислений интенсивно исследуются из-за важности практических применений. В данной статье описывается подход, который позволяет реализовать параллельные вычисления для многих проектных процедур в САПР СБИС. Организация параллельных вычислений на многоядерном компьютере с общей памятью иллюстрируется на примере программ оптимизации, используемых в процессе проектирования цифровых схем и входящих в состав программного комплекса FLC-2, разработанного в лаборатории логического проектирования ОИПИ НАН Беларуси [9,10]. Реализацию параллельного запуска программ предлагается организовать с использованием технологии OpenMP. В работе проводится оценка быстродействия следующих программ минимизации, работа которых организована в параллельном режиме: отечественной программы минимизации булевых функций в классе ДНФ, входящей в систему FLC-2 [10] и зарубежных, свободно распространяемых программ Espresso ПС [12] и ABC [13]. Предложенная методика распараллеливания, основана на имеющейся в FLC-2 возможности разделения исходных данных проектными процедурами. Её использование приводит к сокращению времени выполнения проектных процедур, для которых применима что в итоге позволяет увеличить размерности решаемых задач логического проектирования.

Описание используемой программной среды и объектов для экспериментов.

В работах [9, 10] приведено описание системы логической оптимизации функционально-структурных описаний цифровых устройств FLC-2. В качестве исходных данных для системы выступает функциональное либо иерархическое функционально-структурное описание проектируемой цифровой схемы. Результатом является оптимизированное описание, на основе которого можно получить логическую схему меньшей сложности и большего быстродействия. В настоящее время существенно расширился инструментарий системы и возможности базового функционального обеспечения, частности, в системе FLC-2 проектные процедуры позволяют проводить оптимизацию исходных описаний используя разнообразные методы, как на основе структур данных разработанных для FLC-2, так и с привлечением программ логической оптимизации, имеющихся в свободном доступе. В качестве языка описания логических схем в FLC-2 используется язык SF [9], который ориентирован на иерархическое описание логической схемы. Иерархия описания схемы представляется в виде дерева, вершине дерева соответствует отдельный блок. Любой не листовой блок, соответствующий узлу дерева иерархии, выражается заданием связей входящих в него подсхем. Описание схемы на языке SF представлено последовательностью функционально-структурных описаний подсхем (блоков). Функциональные описания листовых блоков дерева иерархии представляют собой либо логические уравнения – скобочные выражения в булевом базисе И, ИЛИ, НЕ (так называемый LOG-формат), либо матричные формы представления систем булевых функций в виде ДНФ (так называемый SDF-формат). Система ДНФ булевых функций задается парой матриц: строки троичной матрицы задают элементарные конъюнкции, входящие в систему ДНФ, в булевой матрице единицы задают вхождения конъюнкций в ДНФ функций. Если головной блок описан на функциональном уровне, то в этом случае весь проект описания схемы представляет собой один листовой блок. Форматные преобразования в FLC-2 позволяют выполнять преобразование многоуровневых представлений в виде уравнений из LOG-формата в матричную форму системы ДНФ (SDF-формат). Таблицы истинности систем булевых функций также представляются в матричной форме (SDF-формат). Методы разбиения логической схемы на отдельные подсхемы (блоки)

реализуемые посредством применения соответствующих проектных процедур FLC-2, преследуют различные цели. Для всей схемы и каждого из блоков, в отдельности, могут быть применены методы оптимизации, в том числе методы минимизации в классе ДНФ. Возможность работы с матричными формами и иерархическими описаниями является важной отличительной особенностью языка SF и системы FLC-2 в целом. Конвертация в объекты, представленные на языке описания аппаратуры интегральных схем [14] (и обратно) используется для импорта данных в FLC-2, а также для экспорта оптимизированных SF-описаний с целью их моделирования и схемной реализации. Одним из ключевых моментов при проведении преобразований в рамках программного комплекса FLC-2 является верификация проектных решений на всех этапах маршрута [15,16]. Программы верификации работают для любых пар проектных состояний, как на уровне SF-описаний, так и на уровне объектов, представленных на языке описания аппаратуры.

Программы логической оптимизации, используемые в экспериментах.

Логическая минимизация в классе ДНФ [1, 17-19] реализует алгоритмы совместной и раздельной минимизации по различным критериям (числу элементарных конъюнкций, суммарному числу литералов в конъюнкциях и др.). Ниже перечислены программы, с которыми проводились эксперименты по организации параллельных вычислений.

Программа Espresso [12] является самой известной программой минимизации и предназначена для совместной и раздельной минимизации систем полностью определенных и частичных булевых функций (и систем многозначных функций) в классе ДНФ по различным критериям – числу элементарных конъюнкций и числу литералов. Существуют варианты и модификации данной программы. В экспериментах, результаты которых приводятся далее, использовалась программа Espresso ПС (далее – Espresso) совместной минимизации системы полностью определенных функций.

Программа ABC основана на графовых моделях многоуровневых BDD-представлений систем функций. В программе использованы специальные структуры данных, разработанные для решения этой задачи. Алгоритмы ABC эффективно работают для раздельной минимизации ДНФ булевых функций большой размерности [13].

Программа Minim [17] включена в систему FLC-2, предназначена для совместной и раздельной минимизации систем как полностью определенных, так и частичных булевых функций, заданных в интервальной форме, использует в качестве входных и выходных данных описание логической схемы в формате SDF [9]. В проведенных экспериментах программа Minim использовалась для выполнения раздельной минимизации полностью определенных булевых функций, в которой критериями минимизации являются число конъюнкций в кратчайшей ДНФ каждой компонентной функции. Программа реализует обобщение алгоритма [18], использует специальные структуры данных для векторно-матричных объектов (булевых и троичных векторов, матриц) [2,3]. При проведении раздельной минимизации программой Minim использовался итерационный алгоритм, ориентированный на получение решения лучшего качества.

Примеры систем функций, используемые в экспериментах.

В работе [11] были представлены результаты экспериментов с программами минимизации систем булевых функций в классе дизъюнктивных нормальных форм, а также проведен подробный анализ эффективности и качества результатов оптимизации. Эксперименты в исследовании [11] проводились на четырех наборах систем булевых функций. Для экспериментов, представленных в этой работе, были отобраны 25 примеров из комплекса примеров [11]. В таблице 1 приведены основные параметры примеров, с которыми проводились эксперименты. Первую группу примеров (1-14) составляют системы функций, описывающие функционирование устройств перемножения и устройств возведения в квадрат целых неотрицательных чисел. Подобные примеры систем функций появляются, при решении задач перепроектирования логических схем из одного базиса логических элементов в другой. Вторую группу примеров (15-19) составляют системы булевых функций, взятые из практики проектирования управляющей логики заказных СБИС: примеры S12, ..., S16 – таблицы истинности, задающие SF-описания систем некоторых функций: пример 25 – система ДНФ, которая задает блок управляющей

логики в заказной СБИС. Третья группа примеров (20-24) сформирована из псевдослучайных систем ДНФ [11].

Таблица 1. Параметры систем ДНФ для экспериментов

Исходные данные – моноблок (SF-описание, моноблок)					Исходные данные – сеть (SF-описание, 2Connect- формат)		
№ п/п	Имя примера	n	m	k	Число блоков	Размер файла (Мб)	$\sum_{i=1}^m k_i$
Первый набор примеров							
1	square_12_DNF	12	24	5 612	24	0,214	5 902
2	square_12_TABL	12	24	4 096	24	0,156	39 962
3	square_13_DNF	12	26	11 259	26	0,462	11 904
4	square_13_TABL	14	26	8 191	26	0,336	87 928
5	square_14_DNF	14	28	22 867	28	0,98	24 151
6	square_14_TABL	15	28	16 384	28	0,720	192 100
7	square_15_DNF	15	30	45 568	30	2,08	48 045
8	square_15_TABL	14	30	32 768	30	1,50	416 811 2
9	mult_7_DNF	14	14	13 060	14	0,395	13 252
10	mult_7_TABL	16	14	16 384	14	0,496	93 334
11	mult_8_DNF	16	16	52 810	16	1,76	53 621
12	mult_8_TABL	18	16	65 536	16	2,18	434 660
13	mult_9_DNF	18	18	208 598	18	7,75	212 047
14	mult_9_TABL	18	18	262 144	18	9,75	1 302 835
Второй набор примеров							
15	S12	13	12	4 096	13	12	27 238
16	S13	14	13	8 192	14	13	58 616
17	S14	15	14	16 384	15	14	125 497
18	S15	16	15	32 768	16	15	267 551
19	S16	16	16	65 536	16	16	568 043
Третий набор примеров							
20	Psevdo1	15	20	1 000	20	0,0315	2 103
21	Psevdo2	20	20	5 000	20	0,180	10 676
22	Psevdo3	25	20	10 000	20	0,410	26 657
23	Psevdo4	30	20	25 000	20	1,12	92 300
24	Psevdo5	25	20	30 000	20	1,48	150 620
25	Sistem8	25	20	45 548	20	2,08	45 947

Последовательным применением ряда проектных процедур и форматных преобразований в системе FLC-2 по исходным описаниям всех примеров из таблицы 1 получено разбиение каждого исходного описания, представленного в виде одной ДНФ (далее моноблока), в описания логической сети (формат 2-CONNECT). В таблице 1 приведены параметры логических сетей, блоками которых являются однокомпонентные функции в виде ДНФ. В предпоследнем и последнем столбцах таблицы 1 для примеров 1-25 дополнительно приведена следующая информация: размер файла сети (Мб) и общее число элементарных конъюнкций в блоках логической сети.

Описание предлагаемых подходов и алгоритмов.

Далее рассматриваются нескольких подходов по организации параллельных вычислений с участием упомянутых выше программ минимизации. В работе [20] представлена модификация алгоритма раздельной минимизации программы Minim, в котором параллельное выполнение реализовано путем использования директив OpenMP непосредственно в коде программы Minim. Также в [20] приведен фрагмент алгоритма программы раздельной минимизации, в котором показана организация этого подхода. Псевдокод этого фрагмента программы представлен на рисунке 1.

```
MiRi(Limit, Crit, CSOP &Sop1, CSOP & ResSop)
{
  n=Sop1.foo1(), m=Sop1.foo2(), k=Sop1.foo3(); // n, m, k - входы/выходы/конъюнкции
  CSOP SopR(...); // создать объект для ДНФ
  num_core=omp_get_num_procs (); // определить число ядер
#pragma omp parallel for shared(Limit, Crit, Sop1, SopR, n, m) num_threads(num_core)
  for (i = 0 ; i < m ; i++)
  {
    #pragma omp task firstprivate(i) shared(Sop1, k, SopR, n, m)
    {
      CTM Tm, Tm1; // создать два объекта: Tm, Tm1 - троичные матрицы
      CSOP SopRR(...); // в каждой task объект для ДНФ
      .....
      Tm1=foo4 (...);
      .....
      if(Tm1.foo5 ())
      {
        MiOi (Limit, Crit, Tm1, Tm); // минимизация одной полностью определенной функции
        for (j = 0; j < Tm.foo4(); j++)
          { SopRR.foo5(Tm.foo6(j),...); }
        #pragma omp critical
        {
          SopR.foo7 (...SopRR);
        } // конец критической секции
      } // конец проверки условия в 15-й строке
    } // конец кода в блоке task
  } // конец кода в блоке parallel
  ResSop.foo8 (...SopR); // результирующая система ДНФ
}
```

Рисунок 1. Псевдокод параллельного алгоритма раздельной минимизации (MinimPar)

Другой подход заключается в использовании директив OpenMP для организации параллельной работы группы процессов, каждый из которых запускает программу минимизации со своим набором исходных данных. Такие условия в полной мере соответствуют проведению минимизации для блоков, входящих в состав логической сети, если предварительно провести разделение логической сети на файлы с отдельными блоками. Поэтому в проводимых экспериментах перед параллельным запуском процессов для каждого примера из таблицы 1 проводится предварительное разделение сети на отдельные блоки.

Параллельный алгоритм раздельной минимизации системы булевых функций.

Как было показано в [20] распараллеливание алгоритма раздельной минимизации, когда каждая из функций системы минимизируется независимо от других, может быть сравнительно просто реализовано с использованием директив OpenMP. Объектом распараллеливания является программный цикл (по числу компонентных функций), в котором выполняется минимизация для одной функции. Нативный алгоритм минимизации однокомпонентной функции впервые описан в работе [17]. В экспериментах, результаты которых представлены ниже, используется одна из разновидностей этого алгоритма – т.н. итерационный метод. Как было упомянуто выше, на рисунке 1 представлен псевдокод параллельного алгоритма раздельной минимизации. Описание структур данных для участвующих в алгоритме объектов, а также назначение фрагментов кода более подробно рассмотрены в [20]. В алгоритме задействован механизм так называемых задач, который определяется заданием в коде директив OpenMP – **task** и **parallel for**. Эти директивы проводят с отдельными фрагментами кода следующие действия. После OpenMP-директивы **parallel for** образуются параллельные регионы. В конце параллельного региона группы потоков останавливаются, а выполнение основного потока продолжается. Далее внутри отдельного параллельного региона посредством OpenMP-прагмы **task** в текущем потоке выделяется отдельная независимая задача, с которой ассоциирован соответствующий блок операторов. Такая задача помещается внутрь параллельной области и может выполняться в отдельном потоке. Однако задача **task** не образует новый поток, а помещается в пул, из которого ее может взять один из свободных потоков. Задача может выполняться немедленно после создания или быть отложенной на неопределенное время и выполняться по частям. Размер таких частей определяется размером фрагмента операторов, а порядок выполнения частей текущих или отложенных задач определяется системным планировщиком задач и условиями синхронизации. Динамическое распределение задач по потокам осуществляется алгоритмами планировщика задач [20]. С помощью OpenMP-директивы **critical** оформляется критическая секция

программы, в которой собираются минимизированные ДНФ компонентной функции, полученные отдельными задачами. В каждый момент времени в критической секции может находиться не более одной задачи. Все другие задачи будут заблокированы, пока вошедшая задача не закончит выполнение. Как только вошедший в критическую секцию поток выйдет из нее, один из заблокированных потоков туда войдет. Если на входе стоит несколько потоков, то случайным образом выбирается один из них, а остальные продолжают ожидание. Естественно, введение критических секций снижает степень параллелизма. Задачи **task** являются альтернативой традиционным потокам **threads**. Механизм задач позволяет реализовать лучший баланс нагрузки, эффективное использование имеющихся ресурсов и предлагает высокий уровень абстракции, благодаря которому можно не заботиться о непосредственном управлении параллельным исполнением. Время нахождения задачи в критической секции, где формируется объединение ДНФ текущих решений, определяется временем нахождения минимальной ДНФ для одной компонентной функции. Модификацией программы **Minim**, которая проводит параллельные вычисления согласно алгоритму на рисунке 1 является программа **MinimPar**. Исходными данными для программы **MinimPar** служит система ДНФ, представленная одним блоком в формате **SDF**.

Алгоритм запуска параллельных Qt-процессов.

В основу алгоритмов, которые реализуют параллельную работу программ минимизации, положено создание и использование процесса. Процесс – это экземпляр программы, загруженной в память компьютера для выполнения. Каждый процесс имеет свою область памяти. Когда программа производит запуск другой программы, средства ОС всегда создает новый процесс (дочерний). В среде разработки Qt [21] можно создавать процессы с помощью класса **QProcess**, который используется для запуска внешних программ и связи с ними, содержит методы для манипулирования системными переменными процесса. Работа с объектами класса **QProcess** может осуществляться в синхронном и асинхронном режимах. Запуск процесса выполняется методом **start()**, в который необходимо передать имя команды и список ее аргументов. В поле аргументов может быть введена любая команда, в том числе это может быть командная строка, запускающая внешнюю программу. Фрагмент программы **StartParBlock**, псевдокод которого приведен на рисунке 2, организует запуск в параллельном режиме процессов с исполняемыми модулями запускаемых программ. На рисунке 2 показано, что для каждого процесса формируются командные строки для запуска программ минимизации, участвующих в экспериментах: **Minim**, **MinimPar**, **ABC** и **Espresso**. Входными данными являются блоки, полученные в результате разбиения логической сети. Каждый блок содержит описание ДНФ однокомпонентной функции.

```
// n , m - число входов, число выходов (блоков), num_core - число ядер
#pragma omp parallel for firstprivate(..)shared(..,NameProg,m) num_threads(num_core)
for(i=1; i<=m; i++) // цикл по числу блоков (m)
{
#pragma omp task untied firstprivate(i, ..) shared (..,NameProg,m)
{ .....
QProcess* m_process = new QProcess(); //объект -новый Qt-процесс
m_process->setProcessChannelMode(QProcess::MergedChannels);
// образование канала для вывода протоколов i-го процесса
m_process->setStandardOutputFile("Qt_protocol.log");
QString strCommand = NameProg; // занести в ком. строку имя программы NameProg
QString strCommand1; // аргументы ком. строки для запуска программы
strCommand1+="-i"+W_in.data(); // исходный sf-файл i-го блока
strCommand1+="-o"+W_out.data(); // sf-файл результата минимизации i-го блока
strCommand+=strCommand1;
// формирование аргументов для программ минимизации
if(NameProg=="Minim") // формирование аргументов для программы Minim
strCommand +="-r"+ config.ini";
if(NameProg=="MinimPar") // формирование аргументов для программы MinimPar
strCommand +="-r"+ config.ini";
if(NameProg=="Espresso") // формирование аргументов для программы Espresso
strCommand +="-f"+"Probe.prb"; // дополнительные параметры для Espresso
if(NameProg=="ABC") // формирование аргументов для программы ABC
strCommand +="-r "+Name_config; // для каждого блока свой ini-файл
m_process->start(strCommand); // запуск i-го процесса
m_process->waitForFinished(-1); // режим i-го процесса - асинхронный
rc = m_process->exitCode();
} // конец omp-директивы task
} // конец цикла по числу блоков (Qt-процессов)
foo1(.....); // Собрать время всех Qt-процессов в один протокол
foo2(.....); // Собрать оптимизированные sf-описания в одну логическую сеть
```

Рисунок 2. Псевдокод алгоритма параллельного запуска Qt-процессов (StartParBlock)

Для работы с программами Espresso и ABC описания блоков (формат SDF) преобразуются во входные/выходные описания, используемые этими программами. Такие преобразования проводятся программами конвертирования, предоставляемыми системой FLC-2. В каждом процессе запускается и выполняется программа минимизации для отдельного блока, начиная с чтения исходных данных блока и завершая записью минимизированного описания блока. После размещения прагмы task создается пул задач, каждой задаче назначен свой процесс [20]. В программе StartParBlock организован замер времени начала и окончания каждой задачи, сохранение этих значений в протоколе, что позволяет определить порядок и длительность выполнения каждой задачи. В качестве иллюстрации на рисунке 3 показано распределение процессов по ядрам при параллельном запуске программы MinimPar для блоков примера 22. После обработки полученного протокола определены длительность и порядок выполнения каждой задачи. По этим данным построена интервальная диаграмма, которая показывает, что планировщик задач направил ожидавшие своей очереди процессы-задачи из пула на ядра, которые быстрее всего завершили выполнение процесса.

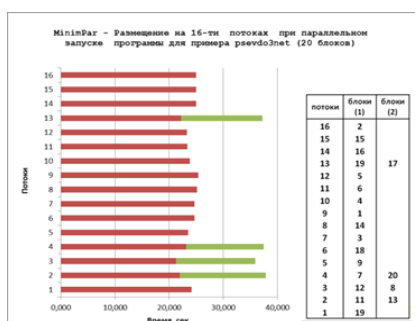


Рисунок 3. Параллельный запуск Qt-процессов. Распределение 20-ти процессов на 16 ядер

Алгоритм параллельного запуска Spawn-процессов.

Системные вызовы среды выполнения Майкрософт также позволяют создавать и исполнять процессы. Системный вызов работает подобно вызову подпрограммы. На рисунке 4 приведен псевдокод алгоритма параллельного запуска с использованием Spawn-процессов. Каждая из запускаемых в цикле системных функций `_spawnlp` создает и запускает новый процесс, который, как и в случае Qt-процесса, можно определить как асинхронный.

```
// n, m - число входов, число выходов, num_core - число ядер
#pragma omp parallel for firstprivate(...) shared(...,NameProg, m) num_threads(num_core)
for(int i=1; i<=m; i++) // запуск spawnlp процессов в цикле
{
    #pragma omp task firstprivate(...,i) shared(... NameProg, m)
    {
        if(NameProg=="ABC") // запуск программы ABC
        {
            W_ini=" -r ";
            W_ini+=Name_config.c_str(); // для каждого блока был создан свой config.ini
            result=spawnlp(...,NameProg,W_in,W_out,W_ini, NULL);
        }
        if(NameProg=="Espresso") // запуск программы Espresso
        {
            W_ini=" -f ";
            W_ini+=TargetPrb.c_str(); // для каждого блока свой config.ini и TargetPrb
            result=spawnlp(...,NameProg,W_in,W_out,W_ini, NULL);
        }
        if(NameProg=="MinimPar") // запуск программы MinimPar или Minim
        {
            result=spawnlp(..., NamePrg, W_in, W_out, NULL);
        }
    } // конец task
} // конец цикла for для всех spawnlp-processes
foo1(.....); // Собрать время всех spawnlp-processes в один протокол
foo2(.....); // Собрать оптимизированные sf-описания в одну логическую сеть
```

Рисунок 4. Псевдокод алгоритма параллельного запуска spawn-процессов (StartParBlock)

Описание экспериментов.

Цель экспериментов: получение временных характеристик параллельных вычислений, анализ быстродействия предложенных алгоритмов организации параллельных вычислений с программами минимизации, которые используются в САПР логического проектирования.

Эксперимент 1. Сравнение времени работы программ Minim и MinimPar.

Программы Minim и MinimPar выполнялись с установкой следующих опций (раздельная минимизация, представление системы ДНФ булевых функций в виде моноблока и 2Connect-сети). В данном эксперименте проводится сравнение временных характеристик для программ Minim и MinimPar для двух представлений исходных данных. В первом разделе таблицы 2 для моноблока организован запуск этих программ в последовательном и параллельном режимах, указаны времена выполнения программ Minim и MinimPar, вычислены коэффициенты ускорения для каждого примера. В примерах 13 и 14 для программы Minim (моноблок, последовательное) указана пометка «нет решения», которая означает, что программой Minim задача не решена за приемлемое время (более 8 часов).

Таблица 2. Время выполнения Minim, MiniPar (моноблок), MinimPar (сеть)

Программа Minim – раздельная минимизация, итерационный метод									
Моноблок (SF-описание, формат SDF)				2Connect сеть(SF-описание, формат SDF)					
№ п/п	Minim последов.	MinimPar паралл.	ускор.	Запуск MinimPar, Qt-процесс			Запуск MinimPar, Spawn-процесс		
	время, с	время, с		последов.	паралл.	ускор.	последов.	паралл.	ускор.
				время, с	время, с		время, с	время, с	
1	16,978	5,043	3,4	16,439	4,914	3,3	19,549	7,55	2,6
2	18,747	5,435	3,4	18,212	5,352	3,4	18,048	5,311	3,4
3	119,625	40,617	2,9	119,758	40,048	3,0	70,817	15,108	4,7
4	97,556	29,031	3,4	97,55	29,702	3,3	74,134	14,944	5,0
5	754,733	259,493	2,9	750,31	256,85	2,9	139,963	29,131	4,8
6	622,005	177,737	3,5	604,298	182,075	3,3	152,655	33,643	4,5
7	2223,512	416,663	5,3	2204,563	413,596	5,3	449,173	125,801	3,6
8	2305,956	421,526	5,5	2286,564	431,915	5,3	533,564	142,666	3,7
9	344,828	130,368	2,6	335,795	130,147	2,6	89,633	20,953	4,3
10	343,124	143,878	2,4	331,32	143,485	2,3	100,427	21,319	4,7
11	2429,447	562,579	4,3	2390,28	543,863	4,4	1024,369	367,759	2,8
12	2964,118	560,962	5,3	2670,287	560,198	4,8	1324,489	417,062	3,2
13	–	7119,39	–	19589,14	7045,246	2,8	22671,134	6919,568	3,3
14	–	7857,82	–	25765,78	7695,867	3,3	25346,841	7637,743	3,3
15	32,725	9,555	3,4	30,541	9,709	3,1	29,843	9,719	3,1
16	176,188	63,125	2,8	175,293	63,793	2,7	69,64	14,435	4,8
17	763,559	262,684	2,9	756,968	264,407	2,9	149,959	34,445	4,4
18	2043,163	218,962	9,3	2020,064	426,186	4,7	549,73	153,355	3,6
19	2429,447	708,848	3,4	3587,953	711,828	5,0	2351,599	667,248	3,5
20	0,676	0,151	4,5	1,249	0,237	5,3	0,898	0,224	4,0
21	43,168	9,114	4,7	44,412	9,669	4,6	42,847	9,174	4,7
22	235,770	48,145	4,9	236,041	45,842	5,1	235,558	63,728	3,7
23	246,861	48,724	5,1	339,52	54,317	6,3	268,335	42,677	6,3
24	844,687	163,83	5,2	928,028	183,873	5,0	610,607	108,797	5,6
25	3503,525	644,533	5,4	2778,772	639,848	4,3	1366,052	441,168	3,1

После анализа результатов эксперимента, представленных в таблице 2, можно сделать следующие выводы. Размещение директив OpenMP **parallel for** и **task** в коде программы MinimPar, дает сопоставимые

временные характеристики с характеристиками, полученными при запуске параллельных процессов с программами Minim и MinimPar для сети, что демонстрирует график на рисунке 5. График на рисунке 6 иллюстрирует значения ускорения для вариантов распараллеливания, реализованного с участием программы MinimPar. На рисунке 6 в помеченный эллипсом диапазоне значения ускорения для примеров 13 и 14 не определены, так как в таблице 2 не представлено время выполнения минимизации в последовательном режиме для этих двух примеров.

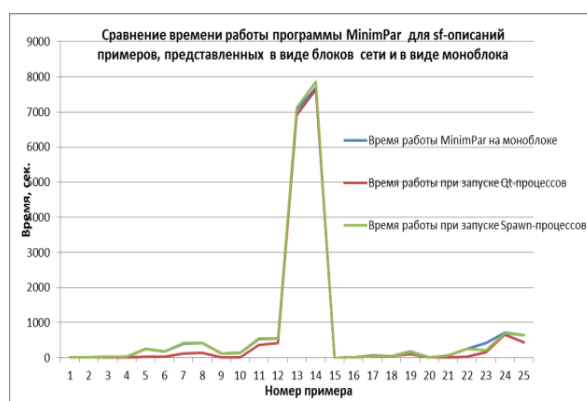


Рисунок 5. Время выполнения параллельных вычислений с MinimPar

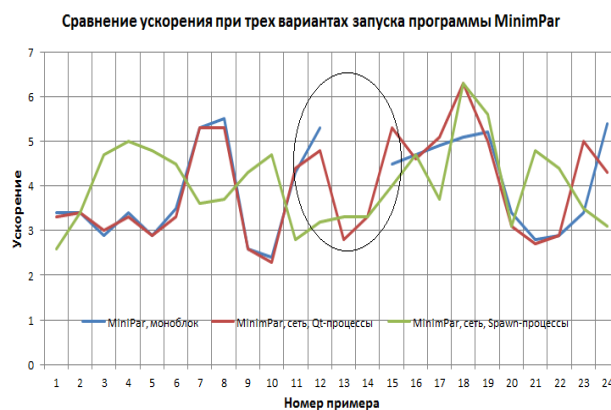


Рисунок 6. Ускорение при параллельных вычислениях с MinimPar

Анализ графиков на рисунках 5 и 6 показывает, что 1) при задании исходных данных в формате моноблока время выполнения MinimPar близко к значениям времени выполнения этой программы, запускаемой с использованием параллельных процессов; 2) график 6 иллюстрирует сопоставимые значения ускорения для программы MinimPar при работе на моноблоке и на сети (Qt-процессы); 3) График 6 показывает, что при запуске параллельных Spawn-процессов получены другие значения ускорения. В целом, разброс значений ускорения от 3 до 6 обусловлено параметрами примеров: размерами файлов и числом блоков. Вывод по результатам этого эксперимента следующий – MinimPar, а также параллельный запуск процессов с этой программой дает увеличение быстродействия в среднем в 4 раза.

Эксперимент 2. Оценка времени работы программы Espresso – два механизма запуска процессов.

Выполнение Espresso осуществляется посредством запуска последовательных и параллельных процессов. Проводится сравнение временных характеристик, полученных при использовании двух механизмов запуска процессов – Qt- и Spawn- процессов. Для Espresso исходными данными являются конвертированные описания блоков сетей примеров 1-25. В таблице 3 указаны: время минимизации программой Espresso для каждой сети при запуске процессов двух типов, вычислены значения ускорения. На рисунке 7 приведены графики, которые построены согласно временным характеристикам, полученным при запуске последовательных и параллельных процессов с программой Espresso. На рисунке 8 показаны графики ускорения для всех вариантов распараллеливания, реализованных с участием Espresso. Графики показывают практически схожие временные характеристики, полученные при минимизации сетей программой Espresso. По результатам этих экспериментов можно сделать следующий вывод, что параллельный запуск с использованием программы Espresso для всех примеров не зависит от вида процессов и показывает сопоставимые характеристики.

Эксперимент 3. Оценка времени работы программы ABC – два механизма запуска процессов.

Программа ABC запускается последовательными и параллельными процессами двух типов: Qt-, Spawn- процессы. Для ABC также проведено конвертирование описаний блоков сетей для примеров 1-25. В таблице 4 указаны времена выполнения минимизации программой ABC, вычислены значение ускорения.

Таблица 3. Время выполнения Espresso (сеть), Qt- и Spawn- процессы

Программа минимизации Espresso						
№ п/п	время, с., Qt-процесс, последов.	время, с., Spawn-процесс, последов.	время, с., Qt-процесс, параллельн.	время, с., Spawn-процесс, параллельн.	ускорение, Qt-процесс	ускорение, Spawn-процесс
1	1,893	1,68	0,4	0,328	4,7	5,1
2	2,105	1,88	0,454	0,42	4,6	4,5
3	2,856	2,597	0,618	0,584	4,6	4,4
4	3,629	3,414	0,862	0,786	4,2	4,3
5	6,224	5,96	1,545	1,348	4,0	4,4
6	8,568	8,491	1,996	1,829	4,3	4,6
7	17,964	17,452	4,712	4,558	3,8	3,8
8	28,335	28,096	6,739	6,325	4,2	4,4
9	3,297	3,197	0,896	0,853	3,7	3,7
10	5,03	4,894	1,223	1,125	4,1	4,4
11	51,134	50,899	15,292	15,223	3,3	3,3
12	85,318	85,913	21,227	21,039	4,0	4,1
13	1862,56	1868,36	692,084	700,964	2,7	2,7
14	2867,787	2853,32	898,208	894,552	3,2	3,2
15	1,336	1,176	0,337	0,272	4,0	4,3
16	2,982	2,8	0,622	0,563	4,8	5,0
17	21,027	20,866	3,895	3,829	5,4	5,4
18	97,641	97,057	21,294	19,982	4,6	4,9
19	462,743	463,262	106,091	105,88	4,4	4,4
20	1,224	1,129	0,295	0,282	4,1	4,0
21	2,294	2,17	0,533	0,472	4,3	4,6
22	6,569	6,445	1,595	1,452	4,1	4,4
23	24,028	23,883	6,019	5,486	4,0	4,4
24	97,262	97,802	22,332	21,653	4,4	4,5
25	91,111	89,757	28,445	28,126	3,2	3,2

Espresso - Сравнение времени получения решения при последовательном и параллельном запуске программы с использованием Qt-процессов и Spawn-процессов

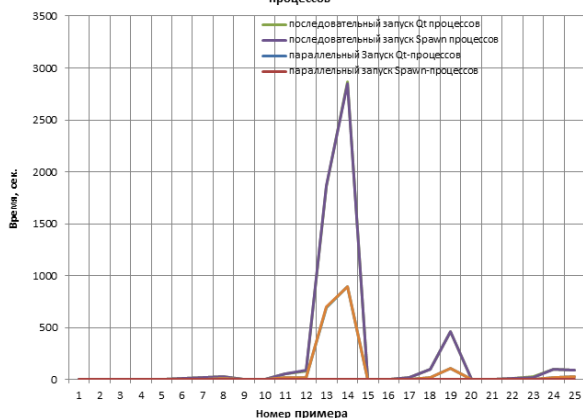


Рисунок 7. Время выполнения программы минимизации Espresso

Espresso - Сравнение ускорения при запуске параллельной программы с Qt-процессами и Spawn-процессами

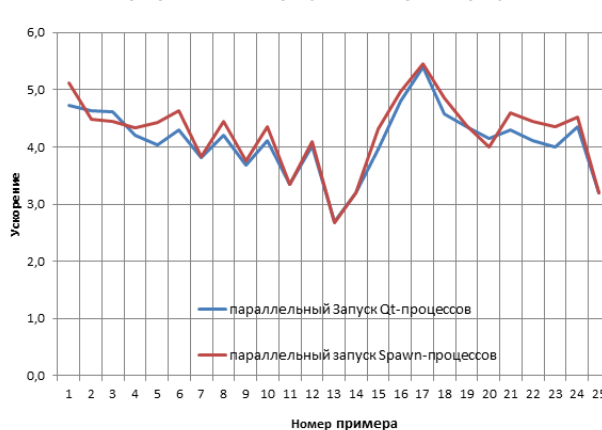


Рисунок 8. Ускорение при параллельных вычислениях с Espresso – два типа процессов

Таблица 4. Время выполнения ABC (сеть), Qt- и Spawn- процессы

Программа ABC – раздельная минимизация						
№ п/п	время, с., Qt-процесс, последов.	время, с., Spawn-процесс, последов.	время, с., Qt-процесс, параллельн.	время, с., Spawn-процесс, параллельн.	ускорение, Qt-процесс	ускорение, Spawn-процесс
1	3,33	4,877	0,548	0,988	4,9	6,1
2	2,673	5,658	0,556	0,945	6,0	4,8
3	3,07	5,449	0,61	1,096	5,0	5,0
4	3,324	6,125	0,647	1,176	5,2	5,1
5	3,457	6,351	0,656	1,391	4,6	5,3
6	4,022	7,719	0,771	1,43	5,4	5,2
7	4,298	8,179	0,878	1,639	5,0	4,9
8	5,09	9,802	1,087	1,917	5,1	4,7
9	1,79	3,309	0,376	0,669	4,9	4,8
10	2,005	3,808	0,388	0,759	5,0	5,2
11	3,127	5,997	0,761	1,394	4,3	4,1
12	3,912	7,67	0,887	1,654	4,6	4,4
13	14,286	28,48	4,359	8,589	3,3	3,3
14	12,873	25,658	3,059	6,209	4,1	4,2
15	2,276	4,097	0,547	0,85	4,8	4,2
16	3,425	6,411	0,719	1,277	5,0	4,8
17	13,464	26,496	2,659	5,458	4,9	5,1
18	28,61	56,9	5,919	11,121	5,1	4,8
19	123,284	245,824	34,06	69,565	3,5	3,6
20	1,524	2,74	0,32	0,535	5,1	4,8
21	1,796	3,332	0,386	0,65	5,1	4,7
22	2,266	4,369	0,455	0,889	4,9	5,0
23	3,124	6,256	0,682	1,267	4,9	4,6
24	4,922	9,677	1,081	2,075	4,7	4,6
25	3,885	7,309	0,845	1,628	4,5	4,6

На рисунке 10 представлены графики, отображающие временные характеристики, полученные при запуске как последовательных, так и параллельных процессов.

На рисунке 11 приведены графики, на котором представлены значения ускорения при параллельных вычислениях с программой ABC для двух видов процессов.

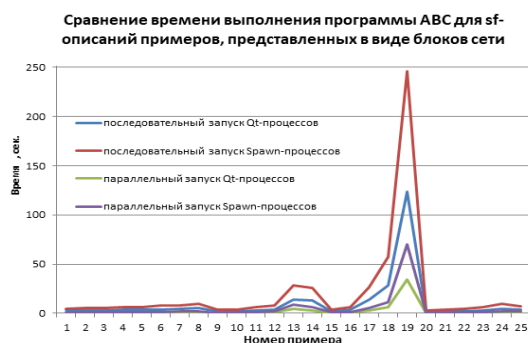


Рисунок 10. Время выполнения программы минимизации ABC



Рисунок 11. Ускорение при параллельных вычислениях с ABC – два типа процессов

Графики на рисунке 10 четко демонстрируют, что временные характеристики минимизации сети программой ABC при использовании Qt-процессов в последовательном и параллельном режимах для всех примеров эффективнее, чем временные характеристики минимизации программой ABC при использовании запуска Spawn-процессов в последовательном и параллельном режимах. График на рисунке 11 показывает, что для всех примеров значения ускорения для Qt- и Spawn-процессов близки. График на рисунке 12 показывает соотношение времени выполнения программы ABC для Qt- и Spawn-процессов. Анализ этого графика позволяет сделать вывод, что проведение параллельных вычислений с программой ABC лучше проводить посредством запуска параллельных Qt-процессов.

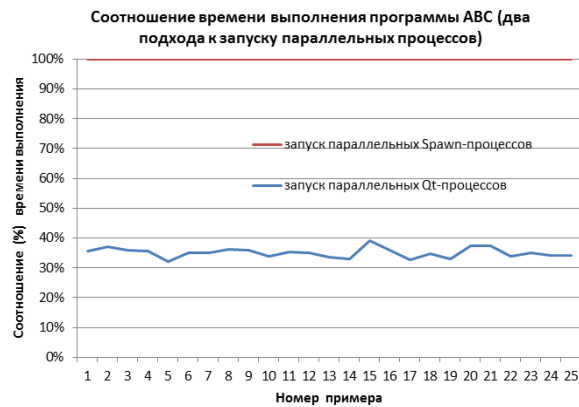


Рисунок 12. Соотношение времени на минимизацию сетей с ABC для Qt- и Spawn-процессов

Эксперимент 4. Организации параллельного запуска для пакета проектов.

В параллельных циклах алгоритмов, приведенных на рисунках 2 и 4, запускаются программы для любого числа блоков. К блокам одного примера можно добавлять блоки, полученные в результате разбиения логической сети другого примера. Программные средства, которые интегрированы с системой FLC-2 и обеспечивают запуск процессов в разных режимах, позволяют проводить объединение блоков нескольких сетей для проведения минимизации одной программой или, наоборот, осуществлять в одном параллельном цикле запуск разных программ, каждая работает со своим пакетом блоков. Как пример, на рисунках 12 и 13 приведены интервальные диаграммы Ганта, построенные после анализа протоколов, полученных при параллельном запуске программ MinimPar и ABC для пакета из блоков примеров 17 и 21.

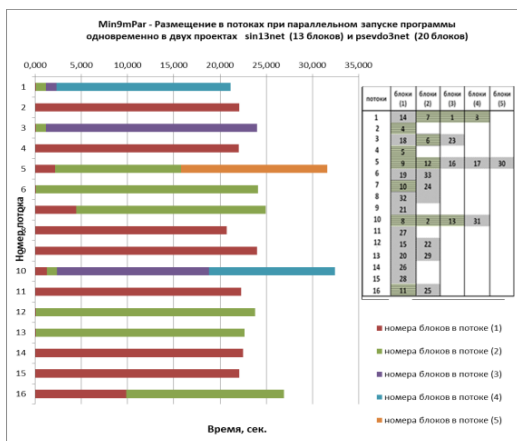


Рисунок 12. Программа MinimPar: распределение задач (процессов) для пакета

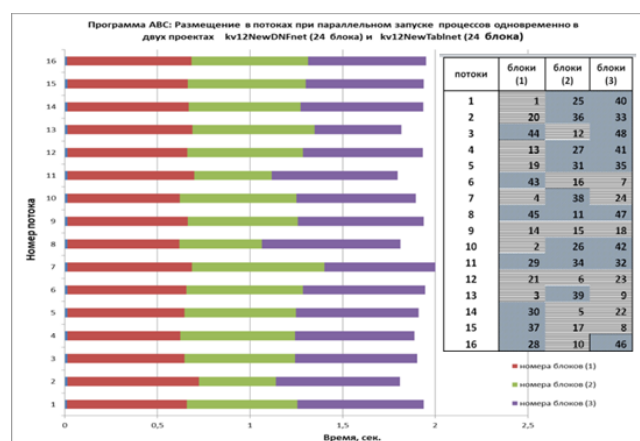


Рисунок 13. Программа ABC: распределение задач (процессов) для пакета

В правой части диаграмм приведены таблицы, в которой показано распределение процессов по ядрам. Из этих диаграмм, видно, что планировщик задач направляет процессы из пула задач вне

зависимости от того, к блокам какого проекта они относятся. В таблице 5 приведены значения ускорения при проведении параллельного запуска процессов для пакета: в 2 раза для программы MinimPar, в 1,89 раза для программы ABC (эксперимент проводился для пакета из 2-х примеров). Таким образом, объединение в пакет дает дополнительное ускорение к тем значениям, которое получается при параллельной минимизации сети отдельного примера.

Таблица 5. Время выполнения программ MinimPar и ABC при объединении блоков 2-х примеров

Параллельная минимизация объединенных блоков двух логических сетей в пакет							
Программа MinimPar				Программа ABC			
Параллельный запуск Spawn-процессов				Параллельный запуск Spawn-процессов			
№ примера	Имя примера	Число блоков	Время, сек	№ примера	Имя примера	Число блоков	Время, сек
17	S13	14	14,435	17	S13	14	0,65
21	psevdo3	20	63,728	21	psevdo3	20	5,458
17 и 21	S13 и psevdo3	34	37,92	17 и 21	S13 и psevdo3	34	3,232
Ускорение (пакет)		2,1		Ускорение (пакет)		1,89	

Заключение.

Предложены алгоритмы по организации параллельных вычислений для программ оптимизации в системе логического проектирования. Данный подход показал эффективность по времени по сравнению с работой этих программ в последовательном режиме. По результатам экспериментов на 8-ми ядерном процессоре Intel i7 получено ускорение в среднем, от 3 до 5. Эксперименты также показали, что проведение параллельного выполнения программ минимизации для пакета примеров дает дополнительное ускорение. Рассмотренные в данной работе подходы предлагают альтернативу разработке новых параллельных алгоритмов для решения комбинаторных задач. Кроме того, этот подход позволяет реализовать в системах проектирования параллельные вычисления с использованием «внешних» программ, которые представлены только исполняемыми кодами.

Список литературы

- [1] Закревский, А.Д. Логические основы проектирования дискретных устройств / А.Д. Закревский, Ю.В. Поттосин, Л.Д. Черемисинова. – М.: Физматлит, 2007. – 589 с.
- [2] Романов, В. И. Булевы векторы и матрицы в C++ / В. И. Романов, И. В. Василькова // Логическое проектирование : сб. науч. тр. – Минск : Ин-т техн. кибернетики НАН Беларуси, 1997. – Вып. 2. – С. 150–158.
- [3] Черемисинов, Д. И. Троичные векторы и матрицы в C++ / Д. И. Черемисинов, Л.Д. Черемисинова // Логическое проектирование : сб. науч. тр. – Минск : Ин-т техн. кибернетики НАН Беларуси, 1998. – Вып. 3. – С. 146–156.
- [4] Banerjee, P. Parallel Algorithms For VLSI Computer-Aided Design / P. Banerjee – Prentice Hall, Englewoods Cliffs, NJ, 1994. – 699 p.
- [5] Черемисинов, Д.И. Проектирование и анализ параллелизма в процессах и программах / Д.И. Черемисинов. – Минск: Беларус. навука, 2011. – 300 с.
- [6] Торопов, Н.Р. Параллельные логико-комбинаторные вычисления в среде MPI / Н.Р. Торопов // Информатика. – 2005. – № 3. – С. 82–90.
- [7] Hamadi, Y. ManySAT: a Parallel SAT Solver/ Y. Hamadi, S. Jabbour, L. Sais // Journal on Satisfiability, Boolean Modeling and Computation, 2009, Vol. 6, pp. 245–262.
- [8] Черемисинов, Д.И. Использование параллельных вычислений при автоматизированном проектировании СБИС / Д.И. Черемисинов, Л.Д. Черемисинова // Проблемы разработки перспективных микро- и наноэлектронных систем – 2016. Сборник трудов / под общ. ред. академика РАН А.Л. Стемпковского. – М.: ИППМ РАН, 2016. Часть – С. 32-39.
- [9] Бибило, П. Н. Логическое проектирование дискретных устройств с использованием продукционно-фреймовой модели представления знаний / П.Н. Бибило, В.И. Романов. – Минск: Беларус. навука, 2011. – 279 с.

- [10] Бибило, П. Н. Система логической оптимизации функционально-структурных описаний цифровых устройств на основе продукционно-фреймовой модели представления знаний / П. Н. Бибило, В. И. Романов // Проблемы разработки перспективных микро- и наноэлектронных систем. – 2020. – Вып. 4. – С. 9–16.
- [11] Бибило, П.Н. Экспериментальное сравнение эффективности программ минимизации систем булевых функций в классе дизъюнктивных нормальных форм / П.Н. Бибило, И.П. Логинова // Информатика. – 2022. – № 2(19). – С. 26–55.
- [12] Logic Minimization Algorithm for VLSI Synthesis / K. R. Brayton [et al.]. – Boston : Kluwer Academic Publishers, 1984. – 193 p.
- [13] Mishchenko, A. An Introduction to Zero-Suppressed Binary Decision Diagrams / A. Mishchenko. – Berkeley : University of California, 2014. – 15 p.
- [14] Бибило, П.Н. Системы проектирования интегральных схем на основе языка VHDL. StateCAD, ModelSim, LeonardoSpectrum / П.Н. Бибило. – М.: СОЛОН-Пресс, 2005. – 384 с.
- [15] Черемисинова, Л.Д. Программные средства верификации описаний комбинационных устройств в процессе логического проектирования / Л.Д. Черемисинова, Д.Я. Новиков // Программная инженерия, 2013, № 7, с. 8–15.
- [16] Логинова, И.П. Верификация с использованием средств formalpro в системе логического проектирования заказных КМОП СБИС / И. П. Логинова // Новые информационные технологии в исследовании сложных структур : материалы Двенадцатой конференции с международным участием. 4–8 июня 2018 г. – Томск : Издательский Дом Томского государственного университета, 2018. – С. 72–73.
- [17] Торопов, Н. Р. Минимизация систем булевых функций в классе ДНФ / Н. Р. Торопов // Логическое проектирование : сб. науч. тр. – Минск : Ин-т техн. кибернетики НАН Беларуси, 1999. – Вып. 4. – С. 4–19.
- [18] Потгосин, Ю. В. Метод минимизации системы полностью определенных булевых функций / Ю. В. Потгосин, Н. Р. Торопов, Е. А. Шестаков // Информатика. – 2008. – № 2(18). – С. 102–110.
- [19] Торопов, Н. Р. Преобразование многоярусной комбинационной сети в двухъярусную / Н. Р. Торопов // Логическое проектирование : сб. науч. тр. – Минск : Ин-т техн. кибернетики НАН Беларуси, 2000. – Вып. 5. – С. 4–14.
- [20] E Ayguadé, E. . The design of Open MP tasks / E. Ayguade [et al.] // IEEE Transactions on Parallel and Distributed Systems. – 2009. – Vol. 20, no. 3. – P. 404–418.
- [21] Шлее, М. Qt 5.10. Профессиональное программирование на C++ / М. Шлее. – СПб.: БХВ-Петербург, 018–1072 с.

PERFORMANCE ANALYSIS FOR IMPLEMENTING PARALLEL COMPUTING TO SOLVE OPTIMIZATION PROBLEMS IN A LOGICAL DESIGN SYSTEM

I.P. Loginova

Senior researcher, UIIP NASB, PhD, associate professor

*United Institute of Informatics Problems of the National Academy of Sciences of Belarus
(UIIP NASB), Belarus
E-mail: irilog@mail.ru*

Abstract. The paper considers the possibility of implementing parallel computing in a system that performs for technologically independent optimization of functional and structural descriptions of logic circuits. Using the example of some logical optimization programs, the methodology for organizing parallel work is demonstrated, including those programs that are represented by executable code. A comparison of the performance of various options for running parallel optimization programs on a stream of known examples is carried out. The results of computational experiments on a multicore system with shared memory are presented. The effectiveness of the approach used in the organization of parallel computing in logical design systems for solving time-consuming tasks is confirmed.

Keywords: multicore system, parallel process, task scheduler, speedup of performance, Disjunctive Normal Form, logical optimization, VLSI.

УДК 004.75

ПРИМЕНЕНИЕ APACHE SPARK ДЛЯ ОБРАБОТКИ БОЛЬШИХ ДАННЫХ В МОБИЛЬНЫХ СИСТЕМАХ



С.Н. Нестеренков
Декан факультета компьютерных систем и сетей БГУИР, кандидат технических наук, доцент
s.nesterenkov@bsuir.by



Д.В. Басак
Инженер-программист, ассистент кафедры ЭВМ БГУИР,
d.basak@bsuir.by



В.В. Куц
Студент 4 курса факультета компьютерных систем и сетей специальности ИиТП БГУИР
mazur.xenia94@gmail.com

С.Н. Нестеренков

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

Д.В. Басак

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

В.В. Куц

Обучается в БГУИР на 4 курсе специальности "Информатика и технологии программирования".

Аннотация. Выполнен краткий анализ роли больших данных в мобильной разработке, возможных способов решения новых сложных задач, поставленных традиционным системам хранения и обработки, и требований к эффективности программного обеспечения, используемого для работы с большими данными. Представлены основы архитектуры популярного фреймворка для разработки с использованием больших данных.

Ключевые слова: большие данные, мобильные приложения, spark.

Введение.

Развитие сферы мобильных приложений, вызванное увеличением количества владельцев смартфонов, стало причиной спроса на услуги компаний-разработчиков мобильных приложений со стороны владельцев бизнеса, нуждающихся в проектировании и разработке персонализированных и эффективных мобильных приложений, способных удовлетворить нужды пользователей.

В современном мире люди полагаются на помощь мобильных приложений при выполнении задач ежедневной рутины. Эти приложения собирают огромное количество данных, эффективная обработка которых требует уникальных средств управления данными. Здесь в дело вступает аналитика больших данных. Собранные данные используются компаниями для улучшения и поддержки приложений, индивидуализации рекламы, создания статистики. Надежное функционирование таких систем возможно при наличии достоверной и своевременной информации о состоянии устройств на объекте, телемеханики и связи.

Актуальность.

Огромные количества изменчивых и разнородных данных, создаваемых людьми и девайсами, нуждаются в развитых технологиях аналитики, хранения и сбора, способных предоставлять обработанную, отфильтрованную и проанализированную информацию, полезную для прогнозирования развития бизнеса и выработки стратегий развития.

Главным фактором в сфере больших данных является скорость, а инструменты этой области предоставляют бизнесу возможность получать данные в реальном времени и эффективно их

анализировать, не полагаясь на неточные или устаревшие наборы данных, что отвечает запросам современных клиентов.

Подробнее о «больших данных».

Исследовательская компания Gartner приводит следующее определение [3]: ««Большие данные» характеризуются объемом, разнообразием и скоростью, с которой структурированные и неструктурированные данные поступают по сетям передачи в процессоры и хранилища, наряду с процессами преобразования этих данных в ценную для бизнеса информацию».

Данное определение позволяет нам выделить четыре основные характеристики больших данных:

1. Объем. Увеличивающееся количество данных, производимых людьми и девайсами, вынуждает ИТ рынок формулировать и обеспечивать новые стандарты в отношении хранения, обработки и предоставления доступа.

2. Разнообразие. Используемые инструменты должны быть способны справиться с неоднородными, неструктурированными, сырыми данными.

3. Скорость. Важна скорость как сбора, так и анализа поступающих данных.

4. Ценность. Качественно обработанные данные могут позволить предсказать трудности или служить решением для уже существующих проблем [2].

Инструменты для обработки больших данных.

Сырые неструктурированные данные требуют качественной аналитики. Инструменты работы с большими данными позволяют разработчикам мобильных приложений собирать, упорядочивать и оценивать различные наборы данных для выявления предпочтений клиентов.

Самые популярные средства для разработки с использованием больших данных Cloudera, Hadoop, Tableau и Spark; они позволяют разработчикам создавать эффективные и современные приложения.

Подробнее об Apache Spark.

Spark предназначен для разработки приложений, нуждающихся в распределённой пакетной и потоковой обработке неструктурированных и слабоструктурированных данных. Основными компонентами архитектуры данного фреймворка выступают driver и executor.

Driver – компонент, отвечающий за:

1. отслеживание состояния обработки, генерацию задач и их перезапуск в случае отказов;
2. оркестровку вычислений – распределение задач среди вычислительных ресурсов с учетом использования ими данных;
3. контроль за вычислительными ресурсами, включая их выделение и возвращение менеджеру ресурсов.

Executor – компонент, ответственный за выполнение вычислений над данными на ресурсах соответствующего вычислительного узла. Его задачи:

1. контроль расхода вычислительных ресурсов;
2. хранение и предоставление доступа к данным, включая загрузку данных с диска или их десериализацию, для последующей обработки;
3. запуск задач и сохранение итогов работы[1], [4].

Заключение.

Компании-разработчики затрачивают большое количество ресурсов для создания интерактивных мобильных приложений, способных подстроиться под пользователя. Качественные приложения находят популярность из-за простоты использования, доступности интерфейса и постоянно растущего числа владельцев смартфонов, что позволяет им собирать огромные количества данных о предпочтениях и локации пользователей. Большие данные поставили перед традиционными системами хранения и обработки новые сложные задачи. Особую ценность имеет исследование возможностей упрощения анализа данных – любые улучшения или оптимизация, помноженные на огромный массив данных для обработки, могут сэкономить много времени. Владельцы программных продуктов, желающие оставаться конкурентноспособными, должны активно вкладываться в эффективные инструменты обработки собираемых данных.

Список литературы

- [1] Абдуллаев А.А., Мамадаев И.М., Червяков Е.В. Большие данные в мобильных приложениях. – Московский государственный технический университет им. Н.Э. Баумана, 2021. – 4 с.
- [2] Клеменков П.А., Кузнецов С.Д. Большие данные: современные подходы к хранению и обработке – 8 с.
- [3] Mark A. Beyer, Douglas Laney. The Importance of «Big Data»: A Definition, 21 June 2012.

[4] Бутаков Н.А., Петров М.В., Насонов Д. Обработка больших данных с Apache Spark. – Университет ИТМО, 2019. – 50 с.

[5] Jaroslav Pokorny. NoSQL databases: a step to database scalability in web environment. Proceedings of the 13th International Conference on Information Integration and Web-based Applications and Services, p. 278-283, ACM New York, NY, USA, 2011.

[6] Одиноченко, М. И. Использование облачных сервисов для решения задач, связанных с применением технологии BIG DATA / М. И. Одиноченко, С. Н. Нестеренков // BIG DATA and Advanced Analytics = BI DATA и анализ высокого уровня: сб. научных статей VIII Междунар. науч.-практ. конф. (Республика Беларусь, Минск, 11-12 мая 2022 года): / редкол. : В. А. Богущ [и др.]. – Минск : Бестпринт, 2022. – С. 252–256.

[7] Беляк, А. А. Анализ производительности технологии Hadoop / А. А. Беляк, С. Н. Нестеренков // BIG DATA and Advanced Analytics = BI DATA и анализ высокого уровня: сб. научных статей VII Междунар. науч.-практ. конф. (Республика Беларусь, Минск, 19-20 мая 2021 года): / редкол. : В. А. Богущ [и др.]. – Минск : Бестпринт, 2021. – С. 343–346.

APPLICATION OF APACHE SPARK TO PROCESSING BIG DATA IN MOBILE SYSTEMS

S.N. Nesterenkov

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

D.V. Basak

*Software engineer,
Assistant of the Department of
Electronic Computers of BSUIR*

V.V. Kuts

*4th year student of the Faculty of
Computer Systems and Networks of
BSUIR*

Faculty of computer systems and networks

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

E-mail: mazur.xenia94@gmail.com

Abstract. A brief analysis of the role of big data in mobile development and the requirements for the efficiency of software used to work with big data is made. The basics of the architecture of the popular framework for development using big data are presented.

Keywords: big data, mobile applications, spark.

УДК 004.654

ОБЛАЧНЫЕ ТЕХНОЛОГИИ И ТЕХНОЛОГИЯ БОЛЬШИХ ДАННЫХ В ОБРАЗОВАНИИ



Е.А. Сальникова
Преподаватель 1 категории УО
БГУИР филиал МРК
магистр
brilic-2@yandex.ru

Е.А. Сальникова

Окончила учреждение образования «Минский государственный высший радиотехнический колледж». Магистрант БГУИР. Работает в учреждении образования «Белорусский государственный университет информатики и радиоэлектроники» филиал «Минский радиотехнический колледж» в должности преподавателя 1 категории цикловой комиссии «Программное обеспечение информационных технологий».

Аннотация. Чтобы повысить мотивацию учащихся к обучению, преподаватели активно применяют различные подходы к организации образовательного процесса. Аналитика больших данных и методы машинного обучения помогают увеличить вовлеченность учащихся и улучшить качество образования. Сегодня большой объем данных создается людьми, машинами и инструментами. И этот набор данных нуждается в новых, улучшенных, инновационных и масштабируемых технологиях для сбора, размещения и анализа. Технология Big Data обрабатывает собранные данные, чтобы в режиме реального времени получить представление о бизнесе, связанное с пользователями, доходами, производительностью, управлением производительностью и рисками. В статье показано, что большие данные в образовании позволяют преподавателям получить разнообразную информацию об уровне подготовки учащихся, усвоении учебной информации, выполненных контрольных заданиях и лабораторных работах.

Ключевые слова: поток данных, образовательный процесс, технологии больших данных, Big Data, облачные технологии.

Введение.

Для обработки больших архивов и больших потоков данных требуются новые технологии, которые часто называют технологиями Big Data (Большие данные). К ним относится информация, которую уже невозможно обрабатывать традиционными способами, в том числе структурированные данные, медиа и случайные объекты [1].

Важность больших данных огромна, поскольку она включает в себя данные, полученные из социальных сетей, машинные данные, записи голоса и видео, структурированные и неструктурированные данные, а также постоянное сохранение этих данных.

Использование больших данных и облачных технологий помогают увеличить вовлеченность учащихся в образовательный процесс и улучшить качество образования.

Big Data и облачные технологии.

В последнее время увеличился интерес к облачным технологиям и большим данным в связи с глобальной цифровизацией во всех сферах общества. А также в связи с большим количеством информации, которую необходимо обработать с помощью современных технологий.

Big Data и облачные сервисы очень быстро внедряются в различные сферы жизни людей: образовательную, медицинскую, экономическую. Учащиеся и преподаватели могут взаимодействовать на одной платформе, обеспечивается более эффективный рабочий процесс, упрощается выполнение различных операций, предоставляется ускоренный доступ к необходимым данным.

Облачные технологии – распределенная обработка данных, в которой доступ к компьютерным программам пользователь получает как онлайн-сервис в режиме реального времени [2].

Облачные технологии устраняют необходимость физического присутствия в учреждении образования. Благодаря внедрению облачных технологий становится возможным объединить преподавателей и учащихся на единой унифицированной платформе. Образовательным организациям, таким как школы, колледжи и университеты, не нужно покупать, владеть и обслуживать собственные серверы и центры обработки данных. Скорее, учреждения образования могут использовать облачные технологии, чтобы использовать вычислительную мощность, базы данных, хранилище и другие услуги, когда они им нужны. Кроме того, учебные учреждения всегда могут быть уверены в безопасности своих ресурсов в облаке [4].

Под термином Big Data понимают огромные наборы структурированных, полуструктурированных или неструктурированных данных. Объемы информации могут быть настолько большими, что их хранение и обработка требуют специальных инструментов [6].

Облачные технологии предоставляют возможности для полноценной работы с Big Data. Легко масштабируемая облачная инфраструктура позволяет хранить и обрабатывать данные без оглядки на технические ограничения. При этом все расходы на управление базовым оборудованием несет поставщик облачных решений, а клиент платит лишь за реально используемые мощности. Также через облако доступны разнообразные программные инструменты для анализа данных. Все это делает анализ больших данных доступным компаниям любых размеров во всех отраслях.

Большие данные отличаются от большого количества информации большим объемом, высокой скоростью накопления информации и ее обработки, многообразием, т.е. наличием одновременно структурированной и неструктурированной разноформатной информации, ценностью и достоверностью. Большой объем предполагает такое количество информации, которое невозможно хранить и обрабатывать классическими методами. Ценность означает, что «большие данные» должны приносить пользу компании.

Чтобы понять всю пользу облачных технологий в современном образовании, стоит проанализировать их преимущества. Первое преимущество, на которое следует обратить внимание – это комфортное управление.

С данным преимуществом сталкиваются преподаватели при внедрении облачных технологий в современную систему образования.

Во-первых, у преподавателя появляется больше шансов привлечь внимание учащегося, сделав процесс обучения увлекательным. Преподаватели университетов могут готовить онлайн-тесты, создавать более качественный контент и общаться со студентами онлайн.

Облачные технологии также помогают охватить более широкую аудиторию студентов и управлять процессом их обучения. Оценивать тесты, домашние задания, проекты, выполненные учащимися, отправлять отзывы и комментарии. С облачными технологиями это стало в разы легче [2].

Вторым преимуществом облачных технологий в сфере образования является быстрый доступ к информации.

Потенциал облака не имеет себе равных, когда речь идет о доступности. По данным сетевых провайдеров, Интернет доступен 99,9% времени, что очень удобно для всех участников образовательного процесса.

Пользователи могут легко получить доступ к содержанию курса, приложениям и данным в любое время и в любом месте. Они могут записаться на курсы и участвовать в групповых мероприятиях. Барьеры места и времени больше не существуют, облако обеспечивает бесперебойную доставку контента в любое время. Более того, он даже отправляет контент на мобильные устройства, чтобы учащиеся могли легко учиться, даже находясь в пути.

Безопасность данных больше не является проблемой, так как все они собираются и хранятся в облаке, и не нужно беспокоиться о том, чтобы хранить все на своем ПК [2].

На облачных сервисах хранятся массивы информации об учебном процессе. Сведения об успеваемости и посещаемости студентов, образовательная и административная деятельность преподавателей, материалы для обучения – целые архивы. Чтобы их обработать, проанализировать, а выводы направить на улучшение образовательных процессов, используют технологию больших данных.

В образовательной сфере выделяют пять типов больших данных для анализа:

- персональные данные;
- данные об обучении студентов: какие используют электронные учебники, онлайн-курсы, просмотры страниц, количество отказов от прочтения;
- данные об использовании учебных материалов: какой контент выбирают студенты для подготовки к занятиям;
- административные данные: посещаемость, пропуски по болезни;
- прогнозы участия студента в той или иной деятельности вуза.

Благодаря способности обобщать и оперировать большими объемами информации технология Big Data помогает сделать учебный процесс интереснее.

Данные аналитики могут влиять на изменение программы учебных курсов и составление такого учебного плана, где будут только те методики, чья эффективность доказана результатами. Только то, что интересно и полезно студентам.

Обработанные данные показывают, какой контент использует студент для обучения, какие учебные материалы его интересуют, от каких он отказывается, как проходит тот или иной курс, где он обращается за помощью.

В будущем такой анализ поможет подбирать каждому учащемуся индивидуальную программу обучения, составлять персональное домашнее задание.

Технология помогает выявить отстающих студентов и сориентировать их в выборе направления, которое больше соответствует их личным качествам и способностям.

Заключение.

Преимущества облачных технологий для сектора образования огромны. Неудивительно, что крупные поставщики в отрасли быстро осваивают облачные технологии, чтобы улучшить предоставляемые ими услуги.

В то же время облако становится лучшим вариантом и для студентов. Ничто не сравнится с удобством доступа к обучению онлайн и облачные технологии делают это возможным. Преподаватели также выигрывают, используя облачные технологии. Они дают им такие плюсы, как живые чаты, потоковые лекции и онлайн-оценки [3].

Вместе с другими формами внедрения облачных технологий может существенно расширить возможности обучения для студентов во всем мире и, в итоге, способствовать обучению будущих поколений навыкам и компетенциям, необходимым для международного карьерного роста.

Большие данные могут стать мощным инструментом для преобразования обучения, переосмыслить подходы, адаптировать опыт для повышения эффективности самой образовательной системы.

Технология больших данных вовремя предоставляет педагогам информацию об успеваемости студентов, помогает контролировать обучение, модифицировать учебные программы, чтобы они были более полезными и интересными для учащихся, рекомендует студентам подходящие курсы.

Когда облачные технологии встречаются с образованием, перед учащимися открывается новый мир возможностей. Более того, доступность, экономичность и совместное обучение могут стать нормой благодаря постоянным инновациям в облаке [6].

Список литературы

- [1] Риз Джордж. Облачные вычисления [пер. с англ.]. – 2011. – С. 288.
[2] Г-н Рэй Дж. Рафаэльс. Облачные вычисления от начала до конца [пер. с англ.]. – 2015. – С. 152.
[3] Джон Ротон – Объяснение облачных технологий // Руководство по внедрению [пер. с англ.]. – 2009 – С. 509.
[4] Эдвард Махон – Переход предприятия в облако // Бизнес-подход [пер. с англ.]. – 2015. – С. 178.
[5] Фиофанова О.А. Анализ больших данных в сфере образования: методология и технологии. М.: Издательский дом «Дело» РАНХиГС, 2020. – 200 с.
[6]. Большие данные в образовании [Электронный ресурс], <http://www.edutainme.ru/post/bolshiedannyev-obrazovanii/>

CLOUD AND BIG DATA TECHNOLOGIES IN EDUCATION

E.A. Salnikova

*Teacher of the 1st category Belarusian State University of
Informatics and Radioelectronics branch Minsk Radio
Engineering College Minsk*

Belarusian State University of Informatics and Radioelectronics branch Minsk Radio Engineering College Minsk

Annotation. To increase students' motivation for learning, teachers actively use various approaches to organizing the educational process. Big data analytics and machine learning methods help increase student engagement and improve the quality of education. Today, a large amount of data is created by people, machines and tools. And this dataset needs new, improved, innovative, and scalable technologies to collect, host, and analyze. Big Data technology processes the collected data to provide real-time business insights related to users, revenue, performance, performance management and risk. The article shows that big data in education allows teachers to obtain a variety of information about the level of preparation of students, the assimilation of educational information, completed control tasks and laboratory work.

Keywords. data flow, educational process, big data technologies, Big Data, cloud technologies

УДК 339.13

BUSINESS DEVELOPMENT В РЕАЛИЗАЦИИ ПРОЕКТОВ ЦИФРОВОГО МАРКЕТИНГА



Л.И. Архипова

Доцент кафедры экономики Белорусского государственного университета информатики и радиоэлектроники, кандидат экономических наук, доцент
l.arkhipova@gmail.com



С.В. Наркевич

Старший преподаватель кафедры экономики Белорусского государственного университета информатики и радиоэлектроники, магистр экономических наук
s.v.narkevich@yandex.ru

Л.И. Архипова

Окончила Белорусский государственный университет и Академию управления при Президенте Республики Беларусь (включая аспирантуру) Имеет многолетний опыт работы в реальном секторе экономики (НПО «Интеграл»). Работает в Белорусском государственном университете информатики и радиоэлектроники. Проводит научные исследования в области современных технологий менеджмента цифрового маркетинга и цифровой трансформации бизнеса.

С.В. Наркевич

Окончил Белорусскую государственную политехническую академию по специальности «Экономика и управление предприятием» и БГУИР (магистр наук). Работал инженером-экономистом в ГП «НИПТИС», экономистом группы ценообразования и договорной работы и отдела бюджетирования в ОАО «Керамин». С 2003 года преподаватель кафедры экономики БГУИР.

Аннотация. В статье исследуется категория и сущность понятия Business development/Развитие бизнеса (BizDev), которая рассматривается, как набор идей, инициатив и видов деятельности, помогающих делать бизнес эффективнее. В статье также рассматривается позиция BizDev в современном маркетинге и бизнесе и обосновывается роль данной категории (вида деятельности) в формировании долгосрочных партнерских отношений и реализации проектов цифрового маркетинга, включая стартапы.

Ключевые слова: business development/развитие бизнеса (BizDev), рост бизнеса, бизнес-решение, маркетинг взаимоотношений, стартап, customer development/развитие клиентов.

Введение. Современный маркетинг, являясь клиентоцентричным, фокусируется на коммуникациях с ориентацией на потребителей и их потребительский опыт (CX), а торговые представители, как правило, направляют свои усилия на поиск квалифицированных потенциальных клиентов (готовых к заключению сделки) и закрытие сделок. Эти две команды решают различные тактические задачи, направленные на достижение целей деловых стратегий организации, которые сегодня обозначены, как стратегии цифровой трансформации бизнеса. Функции Business Development (BizDev), в данном контексте, фокусируются на увеличении дохода, масштабировании бизнеса, а также повышении производительности и прибыльности за счет выстраивания стратегического партнерства и принятия своевременных стратегических бизнес-решений по новым продуктам и рынкам. Необходимость в данной функции определяется природой цифровой трансформации бизнеса и новыми возможностями, которые реализуются применением наборов цифровых технологий (MarTech).

Цифровая трансформация рассматривается, в данном контексте, как процесс использования цифровых технологий для создания новых или изменения действующих бизнес-процессов, культуры и клиентского опыта для удовлетворения меняющихся требований бизнеса и рынка. Такое переосмысление бизнеса в своей основе представляет собой факт

перехода к цифровой трансформации. Бизнес выходит за рамки традиционных подходов и функций, таких как продажи, маркетинг и обслуживание клиентов – цифровая трансформация начинается и заканчивается тем, *что* бизнес знает о клиентах, и *как* взаимодействует с ними.

Исследование и результаты. Место и основные функции BizDev в современной концепции развития и роста бизнеса представлены на рисунке 1 [1-3].



Рисунок 1. Место и роль BizDev в бизнесе

Виды деятельности, связанные с BizDev, встраиваются в современные концепции маркетинга, наиболее значимой из которых, является «маркетинг взаимоотношений – customer relationship management».

Ключевой фигурой клиентоцентричного бизнеса является клиент, именно он определяет будет ли конкретный бизнес «расти или падать». Эта парадигма поддерживает нацеленность бизнеса на «наведение мостов» между бизнесом и потенциальными клиентами, устраняя разрывы между продаваемым продуктом и тем, в котором действительно нуждается клиент, предоставляя клиенту дополнительную ценность и выстраивая более тесные взаимоотношения [4].

Практическая реализация клиентоцентричного подхода является ответственностью BizDev специалистов, которые должны:

- изучать взаимодействие клиентов и продуктов, а также оценивать соответствие продукта рынку или степень, в которой продукт удовлетворяет сильный рыночный спрос (product/market fit);

- взаимодействовать с командами разработчиков для адаптации продукта под полное совпадение с рыночной потребностью – продукты должны решать проблемы клиентов, то есть помогать им выполнять свою работу (JTBD) и повышать производительность;

- работать с командой маркетологов над созданием бренда и брендингом, повышая осведомленность и обучая клиентов;

- обучать и квалифицировать потенциальных клиентов прежде, чем передавать их для дальнейшей работы в отдел продаж;

- формировать стратегические партнерские отношения на базе комплексного решения проблем заказчиков через SPIN-технологии и Solution Selling.

Следует отметить, что BizDev активности реализуются на уровне взаимодействий маркетинга и продаж, не дублируя функции. В таблице 1 даны характеристики каждой из представленных категорий с выделением особенностей и фокусированием на ключевые процессы и принципы [1-4].

Таблица 1. Сравнительная характеристика категорий: маркетинг, продажи, BizDev

Характеристики	Маркетинг	Продажи	Business development
Направления деятельности	<ul style="list-style-type: none"> - повышение осведомленности и активации потребительских сегментов; - формирование спроса; - продвижение бренда; - формирование имиджа 	<ul style="list-style-type: none"> - конверсия потенциальных клиентов в платящих; - стимулирование клиентов, персонала и участников каналов распределения 	<ul style="list-style-type: none"> - создание ценностных предложений; - формирование когорты ключевых клиентов (key accounts); - формирование новых каналов коммуникации и распределения
Инструменты	<ul style="list-style-type: none"> - контент-маркетинг; - e-mail маркетинг; - контекстная реклама; - SEO, SMM; - AARRR и др. 	<ul style="list-style-type: none"> - «воронка» продаж; - CRM, CDP; - автоматизация процесса продаж; - программы лояльности и др. 	MarTech (включая AI и Big Data); SaaS; SPIN; Solution Selling и др.
Фокусирование	<ul style="list-style-type: none"> - создание продуктов, удовлетворяющих конкретные рыночные потребности; - создание условий для «облегчения» процесса продажи 	<ul style="list-style-type: none"> - фокус на клиентов; - фокус на процесс продажи и закрытие сделки; - ориентация на возврат клиентов; - фокус на сделку 	<ul style="list-style-type: none"> - обеспечение конверсии потенциальных клиентов в реальных с высоким уровнем LTV (пожизненная ценность клиента)
Цель	Привлечение клиентов	Транзакции	Партнерские отношения
Метрики	- CTR, CAC (customer acquisition cost)	CR (conversion rate), Retention rate	LTV, ROMI и ROI

Описанные взаимодействия можно представить в виде модели управления формированием отношений с ключевыми клиентами и партнерами (рисунок 2) [5].

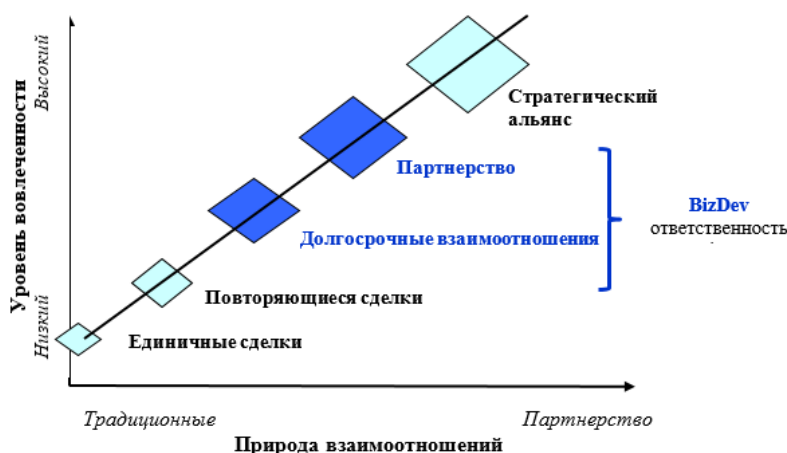


Рисунок 2. Модель формирования взаимоотношений с партнерами

Идентификацию природы взаимоотношений и управления в данном контексте следует понимать, как процесс взаимодействия между покупателями, продавцами и партнерами, в основе которого лежит природа потребительских отношений и уровень их вовлеченности в бизнес. На начальных этапах управления основной целью является установление доверия на основе стабильных показателей качества и открытых коммуникаций. На более высоком уровне

управления число и разнообразие контактов между двумя организациями возрастает (переговоры, бюджеты, совместные действия, др.) К процессу взаимодействия подключаются менеджеры более высокого ранга, уровень доверия повышается, а деятельность фокусируется на решении общих проблем и совместном *поиске рыночных идей и возможностей, включая разработку новой продукции* и ее рыночное тестирование. Особо следует подчеркнуть роль функций и технологий BizDev в создании новых продуктов и стартапов. Как маркетинг, так и business development важны для успеха нового продукта и стартапа. На ранних стадиях стартапа необходимо в большей степени сосредотачиваться на маркетинге, чтобы привлечь клиентов и создать клиентскую базу. По мере того, как стартап растет и утверждается на рынке, приоритетом становится развитие бизнеса (BizDev) – определение и использование возможностей для роста и масштабирования. При выводе новых продуктов на рынок и начинающихся стартапов важно иметь четкое представление о своей целевой аудитории, конкурентной среде и долгосрочных целях, чтобы эффективно распределять ресурсы и расставлять приоритеты в области маркетинга и развития бизнеса. Необходимо также уделять особое внимание работе с клиентской базой, используя инструменты и процедуры customer development (CustDev).

CustDev реально представляет собой тестирование прототипа будущей продукции на потенциальных потребителях и включает в себя серии глубинных интервью по проверке гипотез – выявление проблем продукта и получение обратной связи от будущих клиентов. Цель CustDev состоит в том, чтобы приблизить продукт к потребностям покупателей и их представлениям о полезности и ценности конкретного продукта [6].

В процессе таких взаимодействий с потенциальными покупателями выявляются инсайты по запросам и желаниям, которые способствуют созданию продуктов и сервисов, учитывающих их абсолютную ценность для будущих потребителей. Модель *развития клиента* поддерживает *развитие бизнеса* и реализуется последовательностью этапов (например): выявление потребности, верификация потребителей, моделирование экономики, создание MVP, первая продажа, масштабирование бизнеса [6,7].

Заключение.

BizDev фокусируется на идентификации и поиске возможностей для роста и расширения бизнеса, а также отвечает за увеличение выручки и повышение рентабельности стратегических решений. Развитие бизнеса (BizDev) предполагает действия по достижению долгосрочных стратегических целей:

1. Создание условий для стратегического партнерства и обеспечение успешного взаимодействия с потенциальными партнерами в бизнесе.
2. Изучение существующих и потенциальных рынков и выявление новых возможностей для бизнеса.
3. Поддержание репутации и бренда на базе обоснованного стратегического плана коммуникаций и бизнес-плана развития компании.

Список литературы

[1] What is the difference between Marketing and Business Development? [Электронный ресурс]. – Режим доступа: <https://cmox.co/business-development-vs-marketing/> (Дата обращения: 21.03.2023).

[2] Difference between Business Development, Marketing & Sales [Электронный ресурс]. – Режим доступа: <https://www.linkedin.com/pulse/difference-between-business-development-marketing-sales-kumar/> (Дата обращения: 14.02.2023).

[3] What is the difference between business development and digital marketing? [Электронный ресурс]. – Режим доступа: <https://www.quora.com/What-is-the-difference-between-business-development-and-digital-marketing> (Дата обращения: 07.04.2023)

[4] What Is Business Development? Full Definition & Guide (Plus Tools) [Электронный ресурс]. – Режим доступа: <https://marketsplash.com/business-development/> (Дата обращения: 27.03.2023)

[5] Технологии продаж, деловых переговоров и презентаций: учебн.- метод.пособие / Л.И. Архипова, В.А. Пархименко, Е.А. Олехнович.- Минск: - БГУИР, 2018.- 128с.

[6] Customer development и «кастдев» – что это и в чем разница? [Электронный ресурс]. – Режим доступа: <https://yasno.mobi/blog/customer-development-i-kastdev-cto-eto-i-v-chem-raznitsa> (Дата обращения: 19.03.2023)

[7] Customer development. Трекшн карта этапа customer development [Электронный ресурс]. – Режим доступа: <https://ppt-online.org/27749> (Дата обращения: 15.03.2023)

BUSINESS DEVELOPMENT IN A DIGITAL MARKETING PROJECTS REALIZATION

L.I. Arkhipova

Associate professor, PhD, BSUIR

S.V. Narkevich

*Senior Lecturer of the Department of Economics,
Master of Economic Sciences, BSUIR*

*Belarusian State University of Informatics and Radioelectronics, Republic of Belarus,
E-mail: l.arkhipova@gmail.by.*

Abstracts. The article explores the category and essence of the concept of Business development /Business Development (BizDev), which is considered as a set of ideas, initiatives and activities that help to make business more efficient. The article also discusses the position of BizDev in modern marketing and business, as well as the role of this category (type of activity) in the formation of long-term partnerships and the implementation of digital marketing projects, including start-ups, is substantiated.

Keywords: business development (BizDev), business growth, business solution, marketing relationship, startup, customer development.

УДК 621.365.5

МОДЕЛИРОВАНИЕ ИНДУКЦИОННОГО НАГРЕВА ШАРИКОВ ПРИПОЯ ДЛЯ FLIP-CHIP МОНТАЖА В COMSOL MULTIPHYSICS



В.Л. Ланин

Профессор кафедры электронной техники и технологии, доктор технических наук



А.Д. Хацкевич

*Магистр технических наук, исследователь, инженер-электроник кафедры ЭТТ
dypodt94@mail.ru*

В.Л. Ланин

Профессор кафедры электронной техники и технологии. Имеет 30 летний опыт работы в области технологии ультразвуковой микросварки. Автор 10 монографий в данной области.

А.Д. Хацкевич

Магистр технических наук. Окончил аспирантуру на кафедре ЭТТ. Инженер-исследователь. Имеет 14 публикаций по данной тематике, три из них входят в перечень ВАК.

Аннотация. Выполнено исследование индукционного нагрева шариков припоя. Проведено моделирование процесса в программном пакете COMSOL Multiphysics. Найденны закономерности влияния концентратора и магнитопровода на паяемый образец. Получены термопрофили нагрева в зависимости от частоты индуктора.

Ключевые слова: индукционный нагрев, моделирование.

Введение.

Стремление снижения массогабаритных показателей изделий современной электроники, особенно в секторе высокочастотных и СВЧ-применений, связанных с бурным развитием телекоммуникационных технологий, аэрокосмической техники приводит к развитию новых технологий, в частности “многоэтажных” многокристалльных модулей.

Сборка электронных модулей по технологии 2,5D и 3D-сопряжена с необходимостью использования не только сложного технологического оборудования, но и соответствующих материалов, в частности припойных материалов (шариков) с различной температурой плавления для осуществления последовательного соединения трехмерной конструкции [1].

Одним из методов формирования шариковых выводов является метод Flip Chip это способ монтажа при котором кристалл микросхемы устанавливается на выводы, выполненные непосредственно на его контактных площадках, которые расположены по всей поверхности кристалла микросхемы [2].

Наиболее перспективным методом нагрева является индукционный нагрев – способ нагрева материалов, проводящих электрическую энергию бесконтактным методом при помощи токов высокой частоты. Индукционный нагрев подразумевает нагрев материалов, способных проводить электрическую энергию, при помощи токов высокой частоты, индуцируемых специальным переменным магнитным полем, т.е. нагрев изделий магнитным полем индуктора [3].

Для моделирования процесса индукционной пайки был выбран COMSOL Multiphysics – это программный пакет для анализа, решения и моделирования методом конечных элементов для различных физических и инженерных приложений [4].

Настройка параметров модели

Первым этапом создается геометрия (рисунок 1) и присваиваются материалы компонентов.

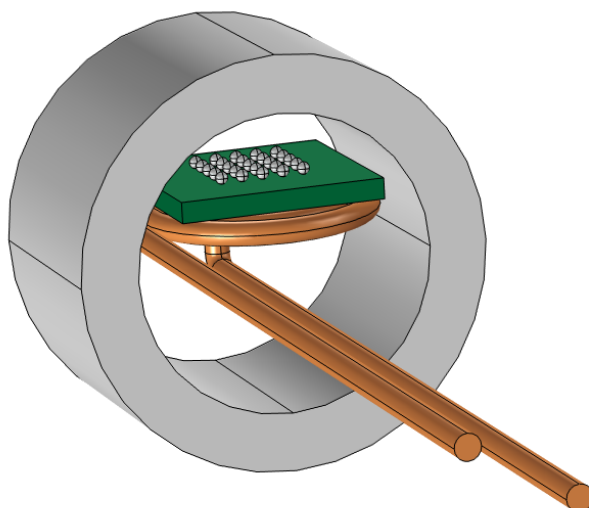


Рисунок 1. Индуктор и печатная плата с массивом шариков припоя в кольцевом ферритовом

Печатная плата с шариками припоя на контактных площадках (рисунок 2а) и концентратором, который прикреплен снизу печатной платы (рисунок 2б), помещена внутрь конструкции, состоящей из плоского индуктора и ферритового сердечника.

Параметры компонентов представлены в таблице 1.

Таблица 1. Геометрические параметры компонентов

Геометрические размеры платы	25x25x2.54 мм
Геометрические размеры концентратора	20x20x0.5 мм
Диаметр трубки индуктора	3 мм
Размер шариков припоя	0.760 мм

Для каждого компонента применены материалы из стандартной библиотеки материалов COMSOL Multiphysics. В качестве материалов катушки использована медь, ферритового сердечника – феррит магнитной проницаемости 2500, стеклотекстолит FR-4 для печатной платы, безсвинцовый припой Sn-3.5Ag-0.5Cu диаметром 0.760мм. В состав модуля Induction Heating входит Magnetic Fields и Heat Transfer in Solids, а также модуль мультифизики Electromagnetic Heating который связывает их между собой.

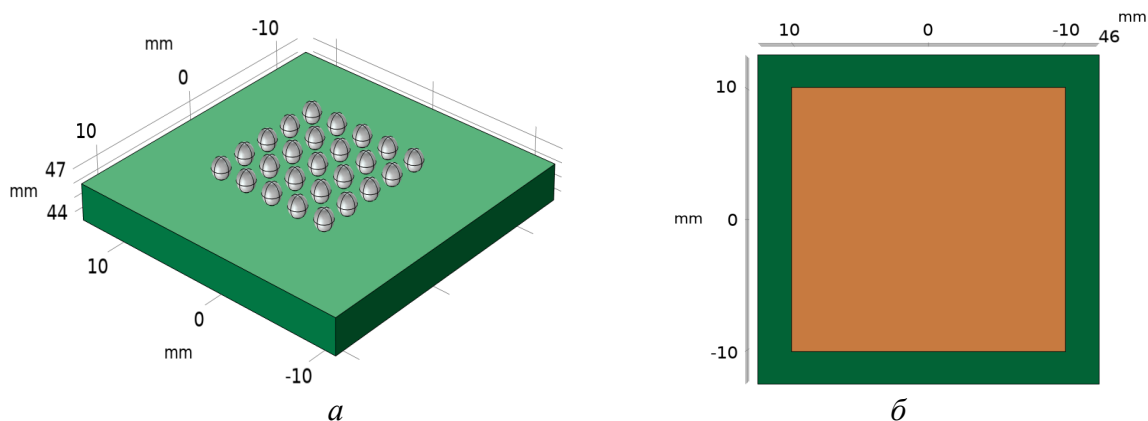


Рисунок 2. Печатная плата с массивом шариков припоя (а) и медный концентратор (б)

Настройку сетки, COMSOL Multiphysics позволяет выполнить как в автоматическом режиме, так и в ручном. В данной задаче автоматический режим не подходит так как сетка получается не оптимальной из-за малого размера шариков припоя. Поэтому для шариков припоя используем Extra Fine размер сетки (рисунок 3), а для оставшихся элементов сетку большей размерности, чтобы оптимизировать модель и тем самым кардинально уменьшить время расчёта модели.

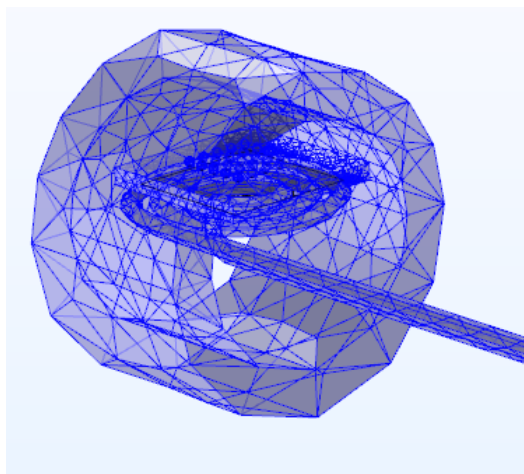


Рисунок 3. Построенная сетка модели

Модуль Coil Geometry Analysis необходимо поместить перед Frequency Domain и Time Dependent. В Frequency Domain указать частоту либо переменную, которой присвоена частота. В Time Dependent указываем время диапазон расчета по времени и шаг. В данном случае временной диапазон выбран от 0 до 120 секунд с интервалом в 1 секунду. Перед расчетом необходимо изменить решатель. Для этого выбрать функцию Get Initial Value и поменять решатель на PARDISO в Solver Configurations.

В результате моделирования получены тепловые поля искомой модели (рисунок 4). Мощность на катушке при частоте 900 кГц составила 29.97 Вт. Средняя температура шариков припоя составила 212 °С, при этом средняя температура медного концентратора составила 276 °С.

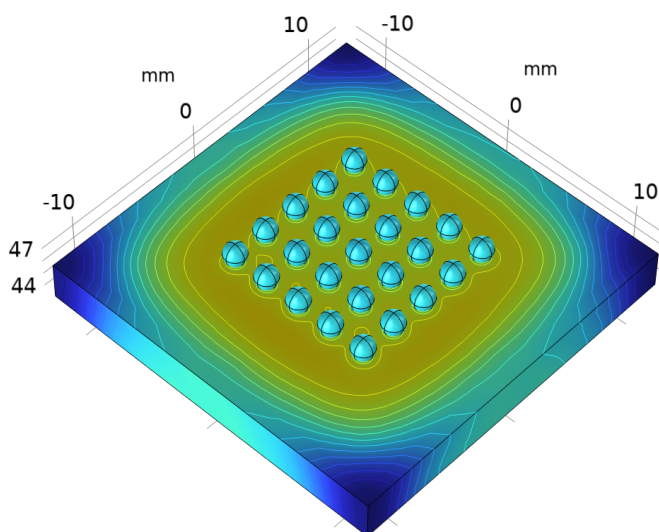


Рисунок 4. Температурные поля печатной платы с нанесенными бессвинцовыми шариками припоя и медным концентратором в ферритовом сердечнике

Термопрофили индукционной пайки печатной платы с нанесенными бессвинцовыми шариками припоя и медным концентратором в ферритовом сердечнике представлены на рисунке 5.

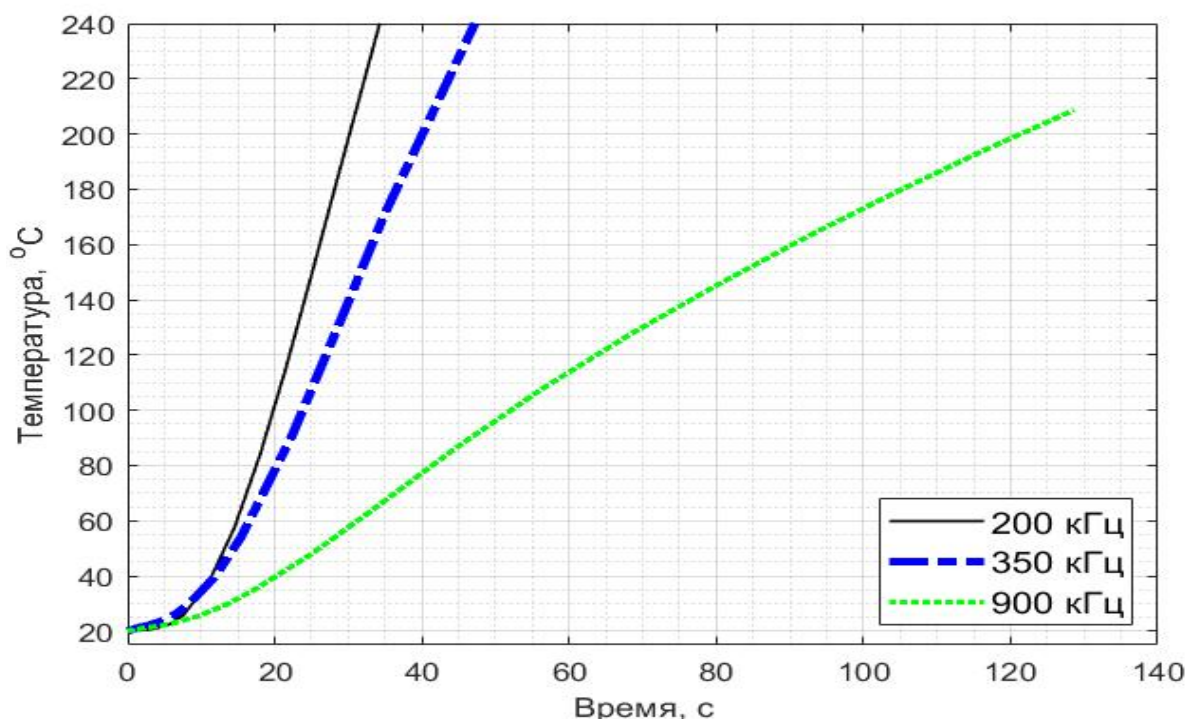


Рисунок 5. Термопрофили с концентратором и ферритовым кольцом на различных частотах

Из графика видно, что уменьшение частоты ведет к увеличению скорости нагрева и температуры шариков припоя. Применение частот ниже 200 кГц нецелесообразно так как это приведет к перегреву медного концентратора, что в свою очередь вызовет перегрев печатной платы с ее последующим разрушением. Поэтому рабочий диапазон частот необходимо выбирать в диапазоне 400 кГц – 1 МГц для данной конфигурации паяемого образца и мощности индуктора.

Закключение.

В результате моделирование получены и оптимизированы параметры индукционного нагрева в программном пакете COMSOL Multiphysics. Использование медного концентратора и ферритового сердечника увеличило эффективность нагрева при этом на паяемый образец кардинально влияет частота катушки индуктора. С уменьшением частоты (менее 200 кГц) медная пластика нагревается до слишком высоких температур (больше 300 градусов), что может привести к подгоранию печатной платы. При этом мощность составила порядка 30 Вт, что доказывает эффективность индукционных систем.

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://ru.wikipedia.org/wiki/Flip_chip. – Дата доступа: 10.04.2023.
- [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.

SIMULATION OF INDUCTION HEATED SOLDER BALLS FOR FLIP-CHIP MOUNTING IN COMSOL MULTIPHYSICS

V.L. Lanin

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

A.D. Khatskevich

*Master of Engineering,
PhD applicant. Electronic Engineer of the Department of ETT*

*Belarusian State University of computer science and Radio Electronics, Republic of Belarus
E-mail: dvpodt94@mail.ru*

Abstract. A study of induction heating of solder balls has been carried out. The process was simulated in the COMSOL Multiphysics software package. Regularities of the influence of the concentrator and the magnetic circuit on the soldered sample are found. Heating thermal profiles are obtained depending on the frequency of the inductor.

Keywords: induction heating, simulation.

УДК 338.46

ТЕОРИЯ И ПРАКТИКА МАРКЕТИНГА БОЛЬШИХ ДАННЫХ В ОРГАНИЗАЦИЯХ ВЫСШЕГО ОБРАЗОВАНИЯ



Е. С. Боженко

*доцент кафедры маркетинга и коммуникаций в бизнесе Южного федерального университета, кандидат экономических наук, доцент
ezhuk@sfedu.ru*



Е. Ю. Полякова

*доцент кафедры маркетинга и коммуникации в бизнесе Южного федерального университета, кандидат экономических наук, доцент
epolyakova@sfedu.ru*

Е. С. Боженко

Окончила Ростовский государственный университет. Область научных интересов связана с трансформацией теории и технологий маркетинга в условиях цифровой трансформации экономики: эволюционный подход.

Е. Ю. Полякова

Окончила Ростовский государственный университет. Область научных интересов связана с внедрением маркетинговых технологий ценностно-ориентированного взаимодействия в современной системе высшего образования.

Аннотация. Система управления в университетах объединяет различные подразделения и коммуникации между ними, а также бизнес-процессы, обеспечивающие развитие образовательной организации, соединяя преподавателей и студентов с необходимыми ресурсами и обеспечивающими службами, а также выстраивает взаимодействия с внешней информационно-коммуникационной средой. Для оптимального управления каналами коммуникаций, формирования эффективных коммуникаций, которые позволят сформировать долгосрочные отношения со всеми стейкхолдерами и выстроить успешную стратегию взаимоотношений, позволяющую достигнуть целей во всех областях деятельности университета, рекомендуется использовать теорию маркетинга больших данных.

Ключевые слова. Big-Data, цифровизация, Цифровая интеллектуальная среда, ценностное предложение, Ценностная модель маркетинга, Большие данные в маркетинге, Система управления в университетах.

Введение.

В настоящее время перед российскими вузами встает новая задача помимо тех традиционных задач, которые они должны реализовывать в области преподавания и исследований – это поддержка развития региона с целью достижения сбалансированного их развития, что станет основой достижения технологического суверенитета.

В связи с этим университеты должны в полной мере использовать механизмы контактов на различных уровнях и в различных областях, а также регулярно проводить углубленные консультации и обмен опытом. Следовательно, в университете увеличивается количество каналов коммуникаций со всеми стейкхолдерами, к которым относятся как абитуриенты, так и вузы-партнеры, СМИ, региональное и международное бизнес-сообщество, региональные и местные власти, предприятия народного хозяйства страны и т.д.

Для оптимального управления каналами коммуникаций, формирования эффективных коммуникаций, которые позволят сформировать долгосрочные отношения со всеми стейкхолдерами и выстроить успешную стратегию взаимоотношений, позволяющую достигнуть целей во всех областях деятельности университета, рекомендуется использовать теорию маркетинга больших данных.

Актуальность. Актуальность данной теории связана с развитием Технологий 4.0. По сравнению с маркетинговыми технологиями 3.0, которые базировались на традиционных подходах маркетинговых исследований и маркетингового анализа, требующих непрерывного изучения рынка, технических достижений, а также определения ценностных установок потребителей, выявления их неосознанных потребностей для дальнейшего соотношения этих данных и выстраивания маркетинговой стратегии, маркетинг 4.0 очень отличается. В традиционном маркетинге маркетинговые исследования базировались на изначальном выдвижении гипотез, в соответствии с которыми проводился сбор данных, т.е. шла фокусировка на известных данных, которая ограничивала поле для получения информации.

В рамках маркетинга 4.0 акцент переносится на определение персонифицированных потребностей потребителей, кастомизацию потребителей с последующей разработкой стратегии маркетинговых коммуникаций в которой происходит интеграция офлайн и онлайн сред при взаимодействии потребителей и организации. В этом случае большие данные делают возможным расширить границы своих представлений о стейкхолдерах и рыночной среде, начиная работу при высоком уровне неизвестности и используя индуктивный метод, который делает возможным проводить эффективную идентификацию данных, выводить гипотезы на этом этапе [14]. Данный инструмент стимулирует развитие новых идей и помогает вырабатывать решения в соответствии с быстроизменяющейся средой окружения.

В этом случае на маркетинг в современном мире большое влияние оказывают большие данные, поскольку данные, которые получают от всех стейкхолдеров и рынка достаточно исчерпывающие, то для их обработки и анализа необходимы современные передовые технологии [11]. Учитывая, что большие данные сами по себе имеют ценность только в том случае, если они используются для извлечения важных знаний, аналитика стала ключевой тенденцией для решения проблем, связанных с большими данными в маркетинге [5]. Здесь могут быть приняты решения по широкому кругу вопросов, таких как определение потенциальных потребителей с положительной реакцией на рекламную кампанию; выявление поведенческих тенденций, складывающихся на рынке; определение отношений потребителей к бренду; реакции на мероприятия стейкхолдеров и прочее. Соответственно с помощью больших данных, которые существуют в структурированных, неструктурированных и полуструктурированных форматах от числовых до аудио- и видеофайлов, можно отслеживать поведение как абитуриентов, так и всех стейкхолдеров, особенно, учитывая, что все данные можно отслеживать в режиме реального времени и с помощью маркетинговых моделей, позволяющих настраивать инструменты маркетинга, влиять на мнение стейкхолдеров на начальных стадиях принятия решения (в процессе поиска информации, сравнения программ и их стоимости).

Следовательно, исходя из актуальности теории маркетинга больших данных и множества подходов к определению содержания данного подхода целью данной статьи является определение содержания, преимуществ и недостатков теории маркетинга больших данных и перспектив ее использования в учебных заведениях высшего образования.

Анализ теории и технологий маркетинга больших данных.

Под Big-Data можно понимать сбор данных как из традиционных источников, так и из цифровых (нетрадиционных) источников, которые представляют собой источник для последующих открытий и анализа [13].

Де Луга в свою очередь трактует большие данные как большие объемы высокоскоростных и гетерогенных данных, которые появляются в результате растущей оцифровки транзакций, взаимодействий, коммуникаций и повседневного опыта. Они включают в себя технологии управления данными, которые обеспечивают несколько ключевых процессов: приобретение и запись; экстракция, очистка и аннотация; интеграция, агрегирование и представление. С помощью аналитики больших данных можно произвести их моделирование, анализ и интерпретацию. Аналитика больших данных часто описывается с точки зрения их подхода и вклада в принятие решений (т.е. описательных, диагностических, прогностических и

предписывающих) и на основе их конкретного домена использования (т.е. текст, речь, веб, сеть и мобильная аналитика). Следовательно, управление большими данными охватывает весь процесс превращения больших данных в ценные идеи, который проходит через управление структурированными и неструктурированными данными, создание специальных структур данных и использование специализированных аналитических алгоритмов [9].

Большие данные можно описать с помощью следующих характеристик: объем (на конец 2020 г. размер создаваемых и хранимых данных составил 40-45 зеттабайт), разнообразие (текст, аудио, изображения, блоги и т.д.) и скорость (быстрота создания), изменчивость и ценность. Соответственно большие данные можно трактовать, как информация большого объема, поступающая с высокой скоростью из разнообразных информационных источников, для которой требуется инновационные формы ее обработки с целью лучшей интерпретации и принятия решений [7].

Ряд ученых к существующему списку характеристик больших данных добавляют также вирусность (процент поступления данных в период времени в определенный узел/пользователю); векторность (направления и размеры распространения больших данных с учетом присущей ей геолокализации и GPS-информации) [10]; волатильность; валидность; достоверность; изменчивость; визуализация [12].

Технологии и инструменты для исследования данных можно разделить на три класса: инструменты пакетной обработки, инструменты потоковой обработки и инструменты интерактивного анализа. Каждая платформа больших данных имеет определенную функциональность и направленность. Например, некоторые платформы предназначены для пакетной обработки, например Apache Hadoop или Pentaho Business Analytics, тогда как другие ориентированы на аналитику в реальном времени, например Apache Kafka или Storm. Что касается архитектуры систем больших данных, логические уровни обеспечивают подход к организации компонентов, выполняющих определенные функции.

Решение для работы с большими данными обычно включает четыре логических уровня: источники больших данных — различные каналы, из которых генерируются данные; передача и хранение данных — уровень, отвечающий за получение и хранение данных; анализ — где извлекаются данные и идеи; и потребление — применение результатов слоя анализа.

В маркетинге большие данные используются относительно недавно. Шайхарихаран считает, что большие данные можно считать новым источником капитала [13]. Для маркетинга они открывают новые огромные возможности. Рингель и Скиера отмечают, что большие данные могут быть использованы для визуализации сложной рыночной структуры (более 1000 товаров) методы картирования; для определения эффекта от рекламной кампании; для распространения рекламной кампании множеству пользователей; для исследования потребительского поведения используется инструмент таргетинга, разработанный Трусовым. Они помогают формировать профили потребителей, предоставлять информацию о неудовлетворенных потребностях стейкхолдеров, удерживать потребителей, сохранять лояльность потребителей при одновременном увеличении скорости получения прибыли и влиять на маркетинговую стратегию.

Также с помощью инструментария больших данных и программного обеспечения открылась возможность определения причинно-следственных связей между рыночными факторами, что повышает эффективность разработки и запуска продуктов-новинок и их доведения до потребителей через персонализированные каналы, а также делает доступным выявление новых целевых сегментов. Данные новации помогают организациям сформировать конкурентные преимущества, выйти за рамки взаимодействия, использовать целевой персонализированный маркетинг.

Соответственно можно указать влияние больших данных на каждом этапе создания ценности в маркетинговой ценностной модели.

На этапе создания ценности как уже отмечалось большие данные необходимы для сбора данных о потребителях (например, онлайн-сервисы для мониторинга социальных сетей: Radian, Atlas.ti, T-LAB), факторах рынка, основных акторов рынка с целью построения причинно-следственных связей с дальнейшей эффективной выработкой целей компании, кастомизирования потребителей и определения конкурентной позиции компании для эффективного формирования маркетинговой стратегии, учитывающей персональный профиль каждого участника рынка.

Этап создания ценности предполагает развитие продукта, сервиса, ценообразования, производства и сбыта.

При развитии продукта инструмент больших данных делает возможным осуществить сбор данных о потребительских потребностях, опыте использования товара, степени неудовлетворенности потребителей. Примером может служить компания Fotd, которая в своей практике воспользовалась Big Data Consumer Analytics и собрала информацию с датчиков автомобилей. Это позволило компании усовершенствовать их товар, улучшив функцию распознавания голоса с одновременным внедрением технологии шумоподавления и изменением места микрофона и решить проблему с системой распознавания голоса.

Решения в сбытовой политике на основе больших данных позволяет компаниям принять решения в выборе каналов продаж, места их расположения, определить оптимальный объем поставок продукции в каждый канал. Например, компания Amazon на основе больших данных, собираемых от своих потребителей и стейкхолдеров может сделать краткосрочный прогноз предпочтений потребителей и объемов их заказов этого товара с учетом географического признака. Этот прогноз служит основой для принятия решений об отправке данного продукта на ближайшей к покупателю склад с целью снижения времени доставки этого товара потребителю.

Ценообразование особенно нуждается в инструментах больших данных поскольку даже небольшое повышение цены может увеличить или уменьшить прибыльность на 20–50 % [15].

Соответственно большие данные помогают устанавливать цены и управлять ими, исходя из данных о потребительском поведении. Изучение потребительского спроса становится основой реализации динамического ценообразования. Это является значимым элементов проведения гибкой ценовой политики.

Также большие данные могут повлиять на изменение процессов, базируясь на данных потребительского пути и опыта; помочь оценить эффективность маркетинговой стратегии и компании, выявить провалы и на этой основе провести корректировку.

На третьем этапе ценностной модели маркетинга «Доставка ценности» осуществляется коммуникация компании с потенциальными участниками рынка. Главная задача на данном этапе предоставить гарантии того, что продукция будет доступной потребителю в нужном месте, в нужное время и по соответствующей цене.

Следовательно большие данные необходимы для создания профиля потребителя, карты клиентского опыта с целью идентификации востребованных товаров и прогноза им приобретения конкретными участниками рынка. Пользуясь следующей информацией компания вырабатывает решения в области маркетинговых коммуникаций, выстраивает оптимальную систему маркетинговых коммуникаций, разрабатывает маркетинговые стратегии и рекламные акции.

Решениями в этой сфере являются следующие. Во-первых, возможно создать геозону, из которой проводится автоматическая отправка сообщений участникам рынка на основе их индивидуальных профилей через сбор данных датчиками и запуска триггера при условии входа потребителем в конкретное место.

Во-вторых, создание мобильных приложений, связывающей бренд компании с потребителем. Данные программы, с одной стороны, направлены на привлечение потребителей на свое приложение, где потребитель может идентифицировать продукт через айдентiku бренда и сформировать лояльность к нему. С другой стороны, они служат платформой сбора данных о

потребителе, их предпочтениях и клиентском пути. Например, концепция drivethru, которая была разработана МакДональдс. Кроме основных составляющих в ней размещено цифровое меню, изменяющееся в зависимости от таких факторов, как время суток, погода, опыт предыдущих действий. Она позволяет удобно, просто и на ходу осуществлять покупки.

В-третьих, использование концепции омниканальности. Востребованность данной концепции объясняется тем, что потребитель в процессе своего опыта взаимодействия с товаром, услугой/ брендом постоянно соприкасается с ним в различных точках контакта: как оффлайн, так и онлайн. За последнее время количество таким точек из-за развития коммуникационных технологий и возникших новых условий существования потребителей из-за пандемии значительно увеличилось. Это привело к тому, что потребитель теперь может выбирать, когда и где он будет взаимодействовать с брендом.

Следовательно, возникшая ситуация актуализировала необходимость интеграции стратегий и тактик обмена сообщениями по нескольким каналам в рамках жизненного цикла клиента, что еще больше актуализирует использование инструментов больших данных. Область применения омниканальности расширилась от управления логистикой до понимания намерения клиента использовать данный канал или внедрить новую технологию, до решения проблем интеграции каналов или изучения поведения клиентов при объединении нескольких каналов на пути к покупке.

«Омниканальность – маркетинговый термин, обозначающий взаимную интеграцию разрозненных каналов коммуникации в единую систему, с целью обеспечения бесшовной и непрерывной коммуникации с клиентом» [1].

Из определений следует, что существует различие в понимании многоканальности и омниканальности по таким критериям, как цель, интеграция каналов, количество субъектов, с которыми формируются связи и происходит обмен информацией, взаимодействие отдельных функций маркетинга с целью формирования непрерывных связей и комфортности [8] взаимодействия.

Одно из ключевых отличий заключается в подходе компании к цифровым каналам. В частности, фирмы, стремящиеся оптимизировать производительность для каждого канала, практикуют многоканальный маркетинг, тогда как фирмы, ориентированные на общую прибыльность клиентов по всем каналам, используют омниканальный маркетинг

Основным и главным компонентом омниканального маркетинга - мобильные платформы. Они стирают традиционные межканальные границы.

Механизм достижения конкурентное преимущество с помощью больших данных в маркетинге можно проиллюстрировать следующим образом (рисунок 1). На первом этапе осуществляется сбор данных об участниках рынка с помощью, например CRM – ресурса организации. Однако, каждый раз, когда мы приступаем к сбору данных, например, в социальных сетях, необходимо на основе потребностей организации определиться с использованием того или иного типа отчета: облако слов, анализ тем и тенденций, анализ новостей или обсуждений и т.д. На втором этапе осуществляет обработка информации, ее анализ и интерпретация, которая делает возможность для создания ценностной маркетинговой модели компании. Задачей организации является не только сбор, хранение и анализ данных, но и их очистка от шумов.



Рисунок 1. Процесс формирования конкурентного преимущества посредством маркетинга больших данных [11]

Таким образом, большие данные в маркетинге открывают организационные возможности, под которыми понимается сложный набор навыков и накопленных знаний, который реализуется через организационные процессы, позволяющие организациям координировать свою деятельность на основе оптимального использования собственных активов [6]. Организационные возможности подразделяют на динамические и операционные. Операционные возможности дают возможность организациям достичь конкурентного преимущества в краткосрочный период времени при высоком уровне неопределенности внешней среды на основе реализации основных функциональных действий. В сравнение со статическими маркетинговыми возможностями динамические маркетинговые возможности это подмножество динамических возможностей. Они являются базой для создания межфункциональной способности организации изменять процессы. Это позволяет реагировать на изменения рынка путем разработки, выпуска и интеграции рыночных знаний и ресурсов с целью достижения эффективности на рынке и устойчивого конкурентного преимущества. Соответственно, возможность осуществления такой реакции трансформируют статические маркетинговые возможности в динамические, которые проявляются в принятии маркетинговых решений, управлении разработкой продуктов, управлении цепочками поставок, управлении брендом и управлении взаимоотношениями с клиентами. Основными инструментами, который может быть использованы на практике, являются: 1. ERP представляющий собой программный пакет, который поддерживает управление компанией чрез архитектуру. Это позволяет автоматизировать и интегрировать большинство или практически все бизнес-процессы организации. 2. CRM, сущность которого можно трактовать как ориентированную на клиента стратегию, направленную на повышение прибыли организации посредством повышения ценности для клиента.

Маркетинг больших данных в системе управления образованием.

В условиях цифровизации перед Вузами открываются новые перспективы и траектории развития, которые способствуют формированию качественно новых систем управления университетом, что отражается на стратегическом планировании развития Вуза: информационная система управления университетом, онлайн-поддержка образовательного процесса, ключевые компетенции цифровой экономики и управление учебным процессом с учетом персонализации и построения индивидуальной образовательной траектории [4].

Как было отмечено ранее, происходит переход от мультиканальности к омниканальности, соединяя в единую систему все доступные инструменты и платформы коммуникации, и делая свободным переход между коммуникационными каналами с фиксацией и сохранением всех шагов на пути взаимодействия.

В условиях данного перехода актуализируется проблема отбора элементов омниканального маркетинга и интеграция их в единый процесс. Для решения этой задачи необходимо автоматизировать всю цепочку создания ценности и ее продвижения и заниматься вопросами цифровизации бизнес-процессов, используя инструменты машинного обучения и big-data.

С помощью big data университет формирует цифровую интеллектуальную среду, в которую встроенные технические устройства, позволяют выстраивать не только цифровые коммуникации, но и организовать традиционные офлайн мероприятия с разнообразными стейкхолдерами, с учетом их запросов, паттернов поведения, спецификой реакции на различные стимулы.

Цифровая интеллектуальная среда – это среда, имеющая огромный потенциал при построении долгосрочных взаимоотношений с потребителями, с учетом их ценностей, строящихся на постоянном вовлечении их в деятельность компании. Благодаря цифровым инструментам современные маркетологи не только формируют коммуникационный посыл, но и анализируют массивы данных относительно поведения потребителей (потенциальных и

постоянных), воздействуют на ценностные установки, выстраивают цифровые стратегии взаимодействия.

Таблица 1. Проблемные поля и возможности формирования ценностного подхода к взаимодействию со стейкхолдерами опыта в условиях цифровой трансформации

Фактор	Возможные проблемные поля	Вариант решения в цифровой среде	Возможности формирования ценностного подхода к взаимодействию со стейкхолдерами
Скорость обслуживания	Длительные ожидания потребителей при обращении в компанию, длительность выполнения заказа/доставки	Переход от мультиканальности к омниканальному маркетингу при формировании сервисного обслуживания	Повышение скорости обслуживания; Создание дополнительного свободного времени для потребителя
Гибкость сервиса	Отсутствие эффективной обработки жалоб и претензий потребителей. Отсутствие учета предпочтений и запросов потребителей	Построение карты «клиентского пути»; сотрудничество с клиентами; быстрое реагирование на изменения рынка; прогнозирование и учет запросов целевой аудитории	Прогнозирование и учет рыночных изменений и следование отзывам потребителей позволяет сделать качественный сервис, с учетом запросов и ожиданий потребителей, повышая ценность предложения
Компетентность сотрудников	Отсутствие/ недостаточность коммуникативных компетенций,	Повышение квалификации и обучение персонала, внедрение системы мотивации персонала	Рост уровня доверия со стороны потребителей, формирование оптимального клиентского пути при принятии решения о покупке
Налаженная обратная связь	Отсутствие актуальных каналов обратной связи, провалы в обработке информации, неработающая клиентская база	Постоянный мониторинг обратной связи; актуальная клиентская база; использование удобных каналов коммуникаций для целевой аудитории	Умение оперативно реагировать на запросы рынка, построение эффективных коммуникаций
Доступность информации	Сложность в организации обратной связи	Организация работы контакт-центров, оптимизация сайта и социальных сетей; общение с потребителем в удобных для него любого канала коммуникации; организация доступности цифрового контента для пользователей с различными ограничениями жизнедеятельности (в соответствии требованиями ГОСТа для цифровых ресурсов)	Повышение доверия со стороны потребителей, рост лояльности и удержание клиентов
Персонализация работы с контактными аудиториями		Таргетинг; геомаркетинг; формирование ценностного контента для целевой аудитории	Получение доверия со стороны потребителей; возможность детализации портрета целевой аудитории; получение конкурентных преимуществ на рынке

Система управления в университетах объединяет различные подразделения и коммуникации между ними, а также бизнес-процессы, обеспечивающие развитие образовательной организации, соединяя преподавателей и студентов с необходимыми ресурсами и обеспечивающими службами, а также выстраивает взаимодействия с внешне информационно-коммуникационной средой.

Важнейшим вопросом, связанным с трансформацией системы управления Вуза, является формирование цифровой экосистемы вуза. Это достаточно длительный и сложный процесс, требующий выстраивание многофакторной системы управления бизнес-процессами организации, научной и исследовательской составляющими. Система управления основными административными процессами ЮФУ построена на базе типового программного продукта «1С:Университет ПРОФ».

В ЮФУ система управления бизнес-процессами включает следующие актуальные сервисы: [2]

- автоматизированная система управления «Деканат»,
- автоматизированная система «Расчет нагрузки преподавателей»,
- балльно-рейтинговая система ЮФУ,
- электронная библиотечная система ЮФУ,
- поступление в вуз онлайн,
- портал online.edu.ru, который обеспечивает доступ к онлайн-курсам,
- облачные сервисы для совместной работы сотрудниками университета и др.

Трансформация сервиса, с все большим включением цифровых технологий, позволяющих дистанционно либо в режиме реального времени решать максимальное количество вопросов, не прибегая к физическому посещению точек взаимодействия, делает возможным формирование «ценностно-ориентированного» подхода [3].

Таким образом, развитие системы образования, цифровизация, актуальные изменения и трансформация внешней среды формируют новую систему координат и пересмотр существующих управленческих практик в системе высшего образования и переносит акценты на реализацию коммуникаций, основанных на больших данных.

Данные процессы требуют переориентации на новые принципы построения информационно-поисковых систем и методов извлечения информации, а также предполагают внедрение рекомендательных систем и прогнозных моделей, лежащих в их основе в повседневную деятельность университета.

Алгоритмы на базе рекомендательных интеллектуальных систем способны проанализировать предпочтения и предлагать пользователям релевантные продукты, позволяют правильно настроить таргетинг, для того, чтобы объявления и активности транслировались только целевым потребителям.

Искусственный интеллект может рекомендовать подходящие программы, курсы, составлять персональные подборки.

Рекомендательная система направлена на обработку массива объектов и на основе заданных критериев, выбрать наиболее подходящие варианты решений, которые с наибольшей вероятностью понравятся пользователю.

Конечная цель рекомендательных систем в университете - ценностное предложение и непрерывные коммуникации при вовлечении в университетскую среду. его о товаре, который ему может быть наиболее интересен в данный момент времени.

Источником сбора информации для рекомендательных систем могут быть:

- Информация, которую сам потребитель вам оставляет (анкета, опрос на сайте, почте и др.),
- Информация, которая фиксируется по итогам наблюдения за потребителем,
- Анализ цифрового следа.

Среди основных барьеров применения цифровых инструментов при анализе и моделировании можно выделить:

- Рассогласованность бизнес-процессов,
- Сложность обеспечения защиты персональных данных клиентов,
- Наличие актуального программного обеспечения
- Нехватка специалистов, готовых работать с данными системами.

Коллаборативные	Основанные на контенте	Основанные на знаниях	Гибридные
<ul style="list-style-type: none"> • Рекомендации, основанные на модели предшествующего поведения пользователей. Учитываются оценки и рекомендации других пользователей системы. • Ценностное предложение формируется в зависимости от ранее выполненных различными пользователями запросов на сайте университета 	<ul style="list-style-type: none"> • Рекомендации основываются на знаниях о пользователях, которые можно собрать. Оценки других пользователей системы не учитываются. • Ценностное предложение формируется по профилю потребителя 	<ul style="list-style-type: none"> • Рекомендации основываются на знаниях о какой-то предметной области: пользователях, товарах и других, которые могут помочь в ранжировании. Оценки других пользователей системы не учитываются • Ценностное предложение формируется в зависимости от региональной потребности специалистов 	<ul style="list-style-type: none"> • Комбинирование нескольких алгоритмов • Ценностное предложение формируется в зависимости от запроса и профиля потребителя

Рисунок 2. Типы рекомендательных систем при формировании ценностного предложения

Если затрагивать научно-исследовательскую составляющую современных университетов, то здесь необходимо остановиться на вкладе университетов в аккумулирование метаданных при формировании базы научных знаний в образовательном и научном пространстве вуза. В этой связи можно привести пример создания Зональной научной библиотекой Южного федерального университета Цифрового репозитория, который подключен к таким информационным системам, как:

- Облако Microsoft Azure,
- Наукометрическая база данных Scopus,
- Регистратор CrossRef,
- Данные для федеративной поисковой системы EBSCO Discovery Service.

В репозитории ЮФУ размещено более 80 тысяч материалов, которые автоматически распределяются в зависимости от типа публикации.



Рисунок 3. Структура размещения материалов в цифровом репозитории ЮФУ

Для систематизации управления с базой данных абитуриентов ЮФУ, внедрена CRM система Битрикс 24, позволяющая свести в одну базу всех потенциальных абитуриентов и повысить эффективность взаимодействия.

Заключение.

В результате проведенного исследования можно сделать вывод, что большие данные полностью изменили способ взаимодействия стейкхолдеров с организациями и наоборот, что

создало новые требования и проблемы для маркетинга в любых организациях, в том числе и образовательных.

В процессе реализации подхода маркетинга больших данных организации могут столкнуться с рядом проблем, что проиллюстрировано примерами в рамках статьи.

Во-первых, существует ограниченный опыт применения инструмента больших данных в компаниях, особенно, в малом и среднем бизнесе.

Во-вторых, в литературе представлено ограниченное количество исследований того, как фирмы преобразуют потенциал больших данных в реальную производительность фирмы в конкурентной маркетинговой среде.

В-третьих, не всегда использование больших данных ведет автоматически к лучшему маркетингу, поскольку они переплетаются с некоторыми важными проблемами и проблемами: невозможностью использования уникального центрального блока и классических средств хранения, необходимостью аналитики в реальном времени, корректностью идеи, сохранение конфиденциальности и т. д. Соответственно при принятии решений применения технологий больших данных в организации руководителям следует понимать маркетинговую ценностную модель организации, цели и задачи компании, ее стратегию, в соответствие с чем требуется разработка цепочки больших данных.

Именно комплексный подход в применение больших данных в маркетинге позволяет организации добиться значительных преимуществ в обеспечение конкурентных преимуществ и тем самым устойчивого развития компании. К открывшимся возможностям можно отнести выявление моделей поведения потребителей, использование амбидекстрации рынка для изучения новых клиентов, использование новой концепции обслуживания через разработку или улучшение внутренних систем для более эффективного как представления, так и продвижения услуг и т.д. То есть маркетинг, чтобы воспользоваться преимуществами больших данных должен охватить дисциплины, как наука о данных, машинное обучение, обработка текста, обработка аудио и видео.

В современных условиях большие данные и маркетинговая аналитика становятся составными частями. Тем самым он сможет развить не только статистические маркетинговые возможности, но и охватить динамические, на основе использования организациями маркетинговых знаний и идеи, полученных из больших данных и маркетинговой аналитики.

Список литературы

- [1]. Омниканальность: панацея от кризиса или модный тренд. URL: <https://roem.ru/01-06-2016/225473/omnikalnost/>
- [2]. Повышение эффективности управления в Южном федеральном университете. URL: Система управления административными процессами ЮФУ (1с.ru)
- [3]. Полякова Е.Ю. Боженко Е.С. Клиентоориентированность: современный маркетинговый подход к деятельности компании. - Монография. Ростов-на-Дону. Издательство «Медиа-Полис». 2019. 124с.
- [4]. Полякова Е.Ю. Новые траектории развития вуза в условиях цифровизации //Ценовой и кредитно-финансовой механизм стимулирования экономического развития России в современных условиях (мировой опыт и отечественная практика). Материалы всероссийской научно-практической конференции. - Карачаевск: КЧГУ, 2022. С.264-270.
- [5]. Amado A., Cortez P., Rita P., Moro S. Research trends on Big Data in Marketing: A text mining and topic modeling based literature analysis // European Research on Management and Business Economics. 2018. Vol. 24, Is. 1. P.p.1-7
- [6]. Cao G., Tian N., Blankson C. Big Data, Marketing Analytics, and Firm Marketing Capabilities // Journal of Computer Information Systems. 2021. №62(1). P.p.1-10
- [7]. Chintagunta P., Hanssens D.M., Hauser J.R. Editorial—Marketing Science and Big Data // Marketing Science. 2016. № 35(3). P.p.341-342
- [8]. Cook G. Customer experience in the omni-channel world and the challenges and opportunities this presents // Journal of Direct, Data and Digital Marketing Practice. 2014. №15. P.262-266.
- [9]. De Luca L.M., Herhausen D., Troilo G. et al. How and when do big data investments pay off? The role of marketing affordances and service innovation // Journal of the Academy of Marketing Science. 2021. №49. P.p. 790–810
- [10]. Ducange P., Pecori R., Mezzina P. A glimpse on big data analytics in the framework of marketing strategies // Soft Computing 2018. №22. P.p. 325–342 (2018).

- [11].Fan S., Lau R., Zhao J. L. Demystifying Big Data Analytics for Business Intelligence Through the Lens of Marketing Mix // Big Data Research. 2015. Vol.1, Iss.2. p.p.28-32
- [12]. Pecori R. S-Kademlia: a trust and reputation method to mitigate a Sybil attack in Kademlia // Comput Network. 2016. №94. Part C. P.p. 205–218
- [13].Rivera S.I.G. Big Data Marketing: una aproximación // Perspectivas. 2015. №35. pp. 147-158.
- [14].Saiharan S. Big Data Consumer Analytics and the Transformation of Marketing, 2020 URL: https://www.researchgate.net/publication/348175860_Big_Data_Consumer_Analytics_and_the_Transformation_of_Marketing
- [15].Vastani S.F. Pricing power has strong impact on profits. – Georgia: International parking institute, 2017 URL: <https://www.parking.org/wp-content/uploads/2017/01/Profitable-Pricing-PDF-Version.pdf>

THEORY AND PRACTICE OF BIG DATA MARKETING IN HIGHER EDUCATION ORGANIZATIONS

E.S. Bozhenko

associate Professor of the Department of Marketing and communications in business of the Southern Federal University, Candidate of Economic Sciences, Associate Professor

E.Y. Polyakova

Associate Professor of the Department of Marketing and communications in business of the Southern Federal University, Candidate of Economic Sciences, Associate Professor

*Department of Marketing and Communications in Business
Faculty of Management
Southern Federal University, Russian Federation
E-mail: ezhuk@sfnu.ru*

Annotation. The management system at universities unites various departments and communications between them, as well as business processes that ensure the development of an educational organization, connecting teachers and students with the necessary resources and support services, and also builds interactions with the external information and communication environment. For optimal management of communication channels, the formation of effective communications that will form long-term relationships with all stakeholders and build a successful relationship strategy that allows you to achieve goals in all areas of the university, it is recommended to use the theory of big data marketing.

Keywords. Big-Data, digitalization, Digital intellectual environment, value proposition, value model of marketing, big data in marketing, management system in universities

УДК 004.89

АНАЛИЗ БИОИНФОРМАЦИИ ДЛЯ МОНИТОРИНГА ОПАСНОЙ КОРОНАВИРУСНОЙ ИНФЕКЦИИ



М.В. Спринджук
Доцент кафедры
экономической
информатики БГУИР,
кандидат технических
наук, лектор
stepanenkomatvei@yandex.ru

М.В. Спринджук

Окончил аспирантуру БГУИР. Область научных интересов связана с разработкой алгоритмов, моделей и программного обеспечения для анализа биоинформационных данных геномной природы.

Аннотация. Выполнен анализ биоинформационных данных Монголии и Беларуси. Показаны волны смертности и выздоровления в Монголии по эпидемиологическим данным временных рядов. Проанализирован профиль геномных вариаций коронавируса в Беларуси и Монголии.

Ключевые слова: медицинская кибернетика, коронавирус, эпидемиология, системы медицинского назначения, прикладная математика, мониторинг опасных инфекций.

Введение.

Современное геномное секвенирование позволяет выделить и амплифицировать крупный геном РНК коронавируса [1-6]. Анализ полученных данных (рисунок 1) может идентифицировать и классифицировать элементарные структуры генома. С усовершенствованиями и доступностью компьютерной техники появились новые возможности более полноценно изучать биоинформационные данные, под которыми понимают геномный текст и соответствующие метаданные.

Актуальность.

Внезапный очаг новой коронавирусной инфекции COVID-19 появился в середине декабря 2019 года в Китае, в городе Ухань, инфекция распространилась на многие города Китая, Юго-Восточной Азии, и по всему миру.

За период с декабря 2019 года по апрель 2023 года в мире умерло около 68370864 больных с подтвержденным диагнозом новой коронавирусной инфекции. За это же время в Беларуси умерло 7118 пациента, в России – 397604, в Монголии – 2136 [<https://www.worldometers.info/coronavirus/>].

Материалы и методы исследования.

В течение интервала времени 2019-2023 гг. нами была выполнена загрузка SARS-CoV-2 геномов из общедоступной базы данных GISAID (*Global Initiative on sharing all influenza data = Глобальная инициатива по обмену всеми данными о гриппе*) [<https://www.epicov.org/epi3/frontend#275474>].

Для анализа данных применялось разработанное программное обеспечение на основе платформы Galaxy, в конвейере анализа данных были использованы модули Pangolin и Nextclade. Дополнительно для визуализации данных применялось программное обеспечение Stata [7] (рисунки 2,3), JMP SAS [8] (рисунки 4-6, 9). Также было разработано программное обеспечение

для визуализации прогноза COVID-19 смертности и выздоровления на основе анализа временных рядов. Для его реализации применялись языки программирования R (рисунки 7, 8) [9] и Python [10].

Pangolin [11,12] был разработан для реализации динамической номенклатуры линий и кластеров-кладов передачи SARS-CoV-2, известной как номенклатура Pango. Pangolin присваивает распознанную линию и имя кластера по принципу, опубликованному A. Rambaut др., 2020 [14].

Nextclade [13] – это инструмент, который определяет различия между загруженными пользователем геномными текстами и эталонной последовательностью и использует эти различия для идентификации, распознавания и присвоения линий передачи и кластеров-кладов, а также сообщает о потенциальных проблемах качества последовательностей в представляемых данных. Руководство пользователя доступно по адресу docs.nextstrain.org/projects/nextclade.

Филогенез нового коронавируса человека.

Филогенез коронавирусных геномов был исследован нами по разработанным ранее методикам анализа биоинформационных данных, опубликованным в источнике [14] и как продолжение вычислительных экспериментов.

В основе филогенетических исследований лежат методы множественного выравнивания нуклеотидных или аминокислотных последовательностей и вычислительные методы построения из результатов выравнивания визуализации, отображающей эволюционные отношения геномов, количественной характеристики эволюционной трансформации (рисунки 8,9).

Филогенетические вычислительные эксперименты ограничены техническими возможностями доступной компьютерной техники, отсутствием в нашем случае метаданных имеющихся геномов, что не позволяет в полной мере выполнить пространственный анализ и подробную кластеризацию. Также результаты такого исследования зависят от выбранных алгоритмов и моделей пред- и постобработки данных, методик фильтрации результатов выравнивания и обрезки ветвей кладограмм. Также очевидно, что имеющиеся в базе данных ошибки секвенирования, повлекут неизбежно и ошибки в результатах филогенетических вычислительных экспериментов.

Результаты исследования филогенеза коронавируса в Беларуси и Монголии.

Результаты филогенетического исследования показывают разнообразие мутационных кладов коронавируса. Отмечается также трансмиссионная активность различных кладов внутри страны и выход их из Беларуси за рубеж. По данным филогенетического исследования имеются основания предполагать, что занос новых вариантов на территорию Беларуси продолжается, однако данных о патогенности различных местных штаммов и вариантов недостаточно, чтобы сформировать строгие научные выводы. Биоинформационный анализ показывает вездесущность коронавируса, его уникальные возможности приспосабливаться мутационной изменчивостью. По сравнению с предыдущими вычислительными экспериментами по данным Беларуси, России и Европы нами наблюдается начальная тенденция к доминированию в Беларуси линий передачи Омикрон над Дельтой и другими кластерами и обнаруживается большое мутационное разнообразие изолятов, которые были идентифицированы как варианты Омикрона. В Монголии, согласно нашим вычислительным экспериментам, доминирует *Альфа штамм* коронавируса (рисунки 4, 5).



Рисунок 1. Конвейер получения и анализа информации антивирусного мониторинга. Разработка автора доклада

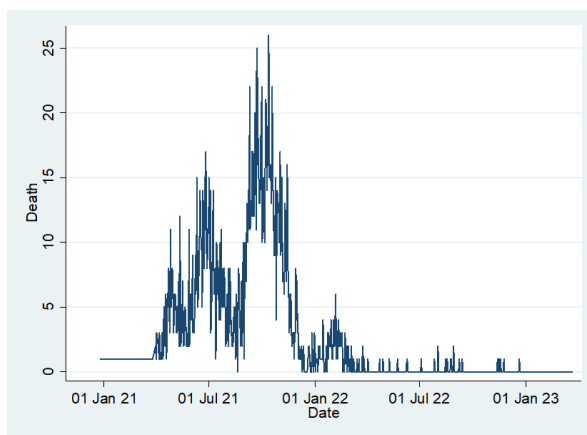


Рисунок 2. Визуализация динамики COVID-19 смертности в Монголии

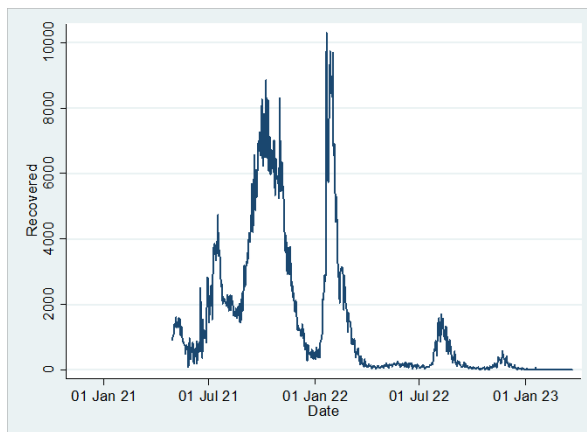


Рисунок 3. Визуализация динамики COVID-19 выздоровления в Монголии

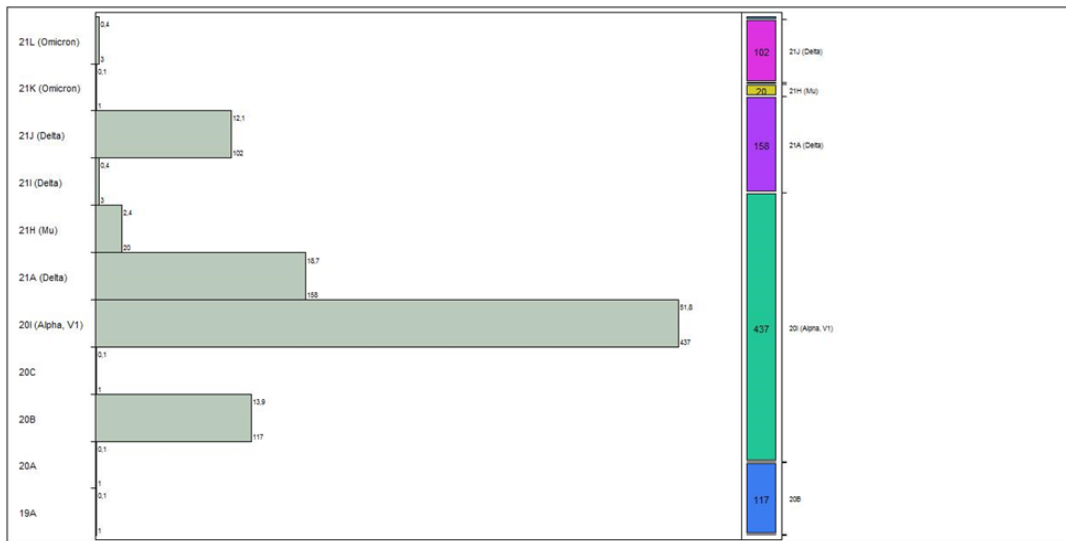


Рисунок 4. Данные Монголии, 844 образца. Показано доминирование *Альфа* штамма коронавируса

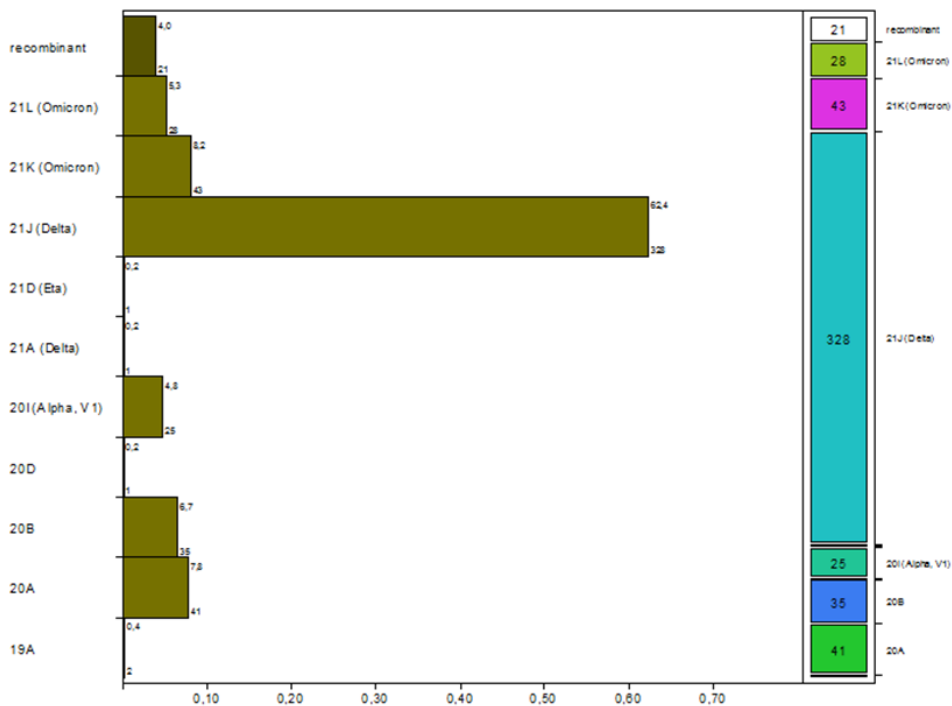


Рисунок 5. Столбчатая диаграмма численного представления результатов идентификации и классификации образцов 526 геномов SARS-CoV-2 из Беларуси. Показано доминирование *Дельта* штамма коронавируса

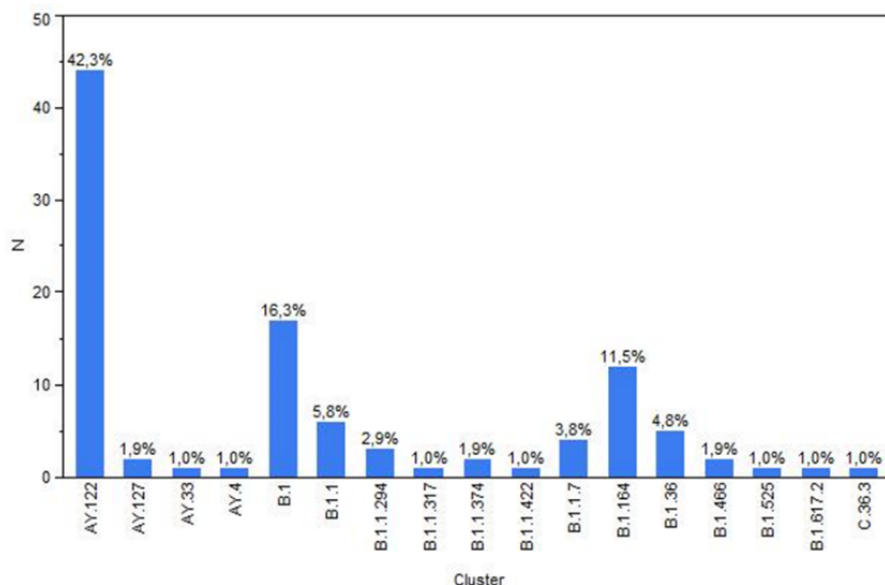


Рисунок 6. Рэнго статистика линий передачи коронавируса (линия AY.122 названа как «Русский» вариант Дельта-штамма коронавируса, причины распространения точно неизвестны)



Рисунок 7 – Фрагмент динамической визуализации множественного выравнивания (модуль языка R Decipher)

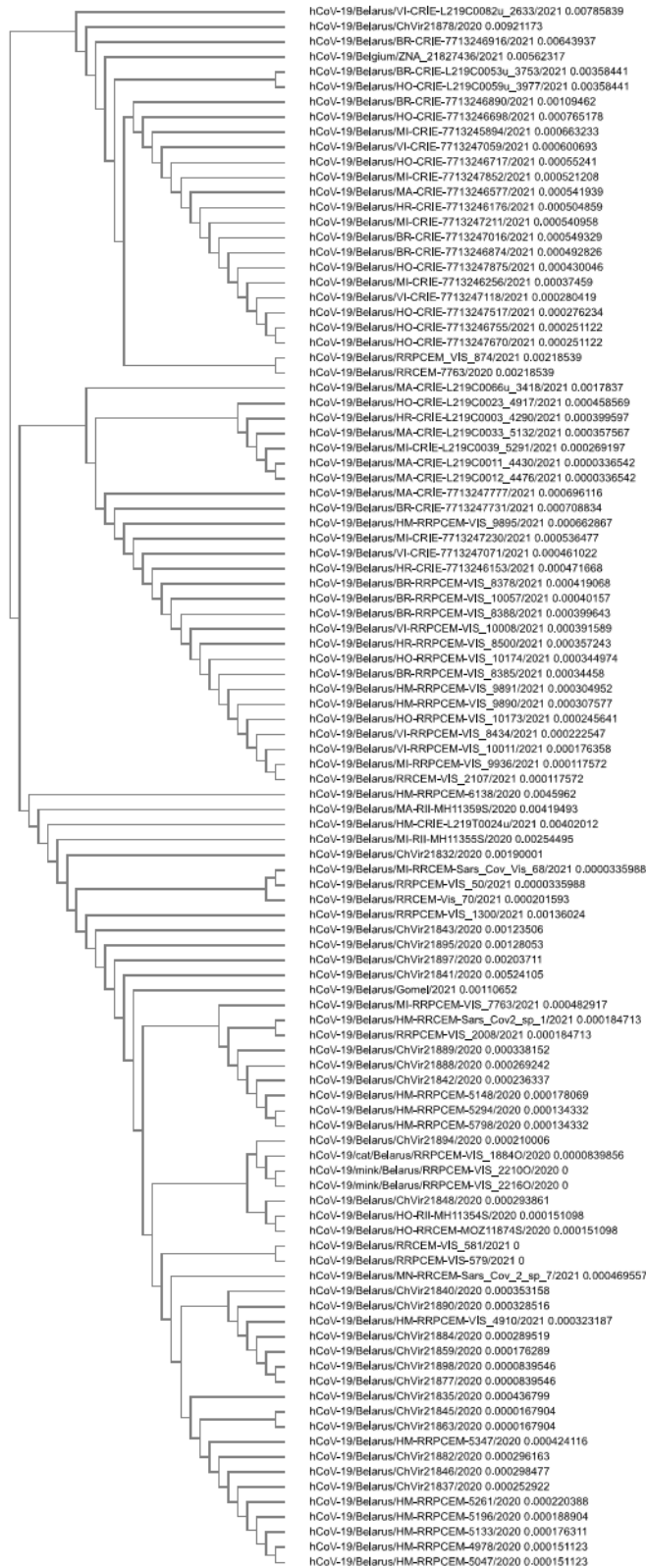


Рисунок 8 Дерево филогенеза 104 геномов коронавируса из Беларуси (модуль Clustal Omega)

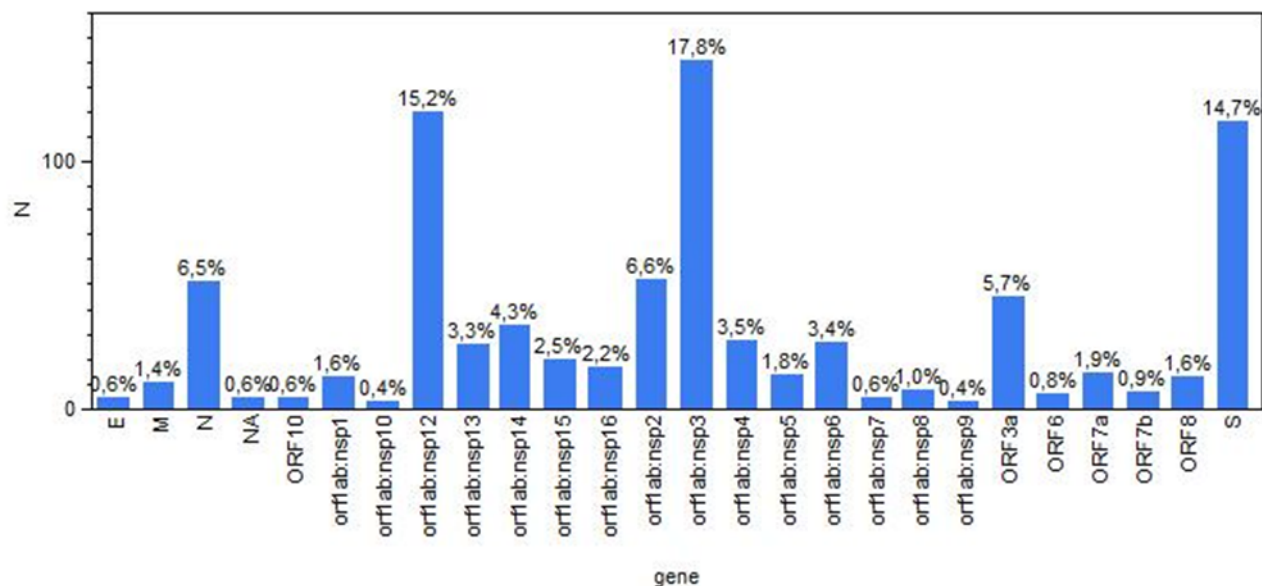


Рисунок 9 Процентное представление частоты мутаций в генах 104 образцов геномов белорусского коронавируса

Заключение.

Вычислительные эксперименты показывают, что геномный профиль коронавируса в Монголии и Беларуси различен, как и эпидемиологические характеристики этой опасной инфекции. Необходимы более крупномасштабные синхронизированные исследования в области геоинформационного изучения эволюции новых опасных вирусов.

Работа выполнена при поддержке проектов БРФФИ: «Математическое моделирование передачи и распространения COVID-19 инфекции на основе систем дифференциальных уравнений и алгоритмов обработки данных с применением технологии машинного обучения» Ф21МН-001, № ГР 20213518 от 27.09.2021; «Ретроспективный анализ клинического и иммунологического статуса групп COVID-19 пациентов с сопутствующим туберкулезом и ВИЧ инфекцией по данным РНПЦ Пульмонологии и фтизиатрии г. Минска», № ГР 20210456 от 31.03.2021; «Разработка и скрининг мукозной вакцины против COVID-19 на основе векторной платформы кишечного аденовируса», № ГР 20210889 от 26.04.2021.

Список литературы

- [1] Coronavirus genomics and bioinformatics analysis / P.C. Woo [et al.] // Viruses. – 2010. – V. 2, № 8. – P. 1804-20.
- [2] Bioinformatics and evolutionary insight on the spike glycoprotein gene of QX-like and Massachusetts strains of infectious bronchitis virus / S.H. Abro [et al.] // Virol J. – 2012. – V. 9. – P. 211.
- [3] Bioinformatics analysis of the factors controlling type I IFN gene expression in autoimmune disease and virus-induced immunity / D. Feng [et al.] // Front Immunol. – 2013. – V. 4. – P. 291.
- [4] Proficiency Testing of Virus Diagnostics Based on Bioinformatics Analysis of Simulated In Silico High-Throughput Sequencing Data Sets / A. Brinkmann [et al.] // J Clin Microbiol. – 2019. – V. 57, № 8. – P. e00466-19. doi: 10.1128/JCM.00466-19. Print 2019 Aug.
- [5] Virus bioinformatics: databases and recent applications / P. Kellam [et al.] // Appl Bioinformatics. – 2002. – V. 1, № 1. – P. 37-42.
- [6] [Sequence analysis for genes encoding nucleoprotein and envelope protein of a new human coronavirus NL63 identified from a pediatric patient in Beijing by bioinformatics] / J.F. Xing [et al.] // Bing Du Xue Bao. – 2007. – V. 23, № 4. – P. 245-51.
- [7] Stata time-series : reference manual : release 10 / StataCorp LP. // – Stata Press, 2007. – 343 p.

- [8] JMP statistical discovery software / B. Jones [et al.] // Wiley Interdisciplinary Reviews: Computational Statistics. – 2011. – V. 3, № 3. – P. 188-194.
- [9] R programming for bioinformatics / R. Gentleman // – CRC, 2009. – 326 p.
- [10] Python data analysis / I. Idris // – Place Published: Packt Publishing Ltd, 2014. – 348 p.
- [11] Assignment of epidemiological lineages in an emerging pandemic using the pangolin tool / Á. O’Toole [et al.] // Virus Evolution. – 2021. – V. 7, № 2. – P. veab064.
- [12] Assessing uncertainty in the rooting of the SARS-CoV-2 phylogeny / L. Pipes [et al.] // Molecular biology and evolution. – 2021. – V. 38, № 4. – P. 1537-1543.
- [13] Nextclade: clade assignment, mutation calling and quality control for viral genomes / I. Aksamentov [et al.] // Journal of Open SourceSoftware. – 2021. – V. 6, № 65. – P. 3773.
- [14] Алгоритмы обработки геномов коронавируса для целей и задач современной иммуноинформатики, вакциномики и вирусологии / М.В. Спринджук, Владыко, А.С., Титов, Л.П., Чжочжуан, Лу, Берник, В.И. // Цифровая трансформация. – 2022. – V. 22, № 8. – P. 71-81.

BIOINFORMATION ANALYSIS FOR DANGEROUS CORONAVIRUS INFECTION MONITORING

Sprindzuk M.V.

PhD of Technical Sciences

MD (medical doctor)

Senior research scientist

Computer science lecturer

Working on position of associate professor

Department of Economical Informatics

Faculty of Computer Engineering

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

Institute of Mathematics of the NASB

United Institute of Informatics Problems of the NASB

E-mail: stepanenkomatvei@yandex.ru

Abstract. The analysis of bioinformatic data of Mongolia and Belarus has been performed. Waves of mortality and recovery in Mongolia are visualized according to the time series epidemiological data. The profile of genomic variations of coronavirus in Belarus and Mongolia has been investigated.

Keywords: medical cybernetics, coronavirus, epidemiology, medical systems, applied mathematics, monitoring of dangerous infections

УДК 004.942, 51-7, 615.8

МОДЕЛИРОВАНИЕ ИМИТАЦИОННОЙ МОДЕЛИ АППАРАТА ИСКУССТВЕННОЙ ВЕНТИЛЯЦИИ ЛЕГКИХ



Н.И. Куприянов
Магистрант кафедры
электронной техники и
технологии БГУИР
n.kupriyanov@bsuir.by



П.В. Камлач
Заместитель декана ФКП;
кандидат технических наук,
доцент, доцент кафедры
электронной техники и технологии
БГУИР
kamlachpv@bsuir.by



А.В. Слижэва
Студентка кафедры
электронной техники и
технологии БГУИР
slizevaa08@gmail.com

Н.И. Куприянов

Окончил Белорусский государственный университет информатики и радиоэлектроники. Область научных интересов связана с разработкой и моделированием аппаратов искусственной вентиляции лёгких, разработкой приложений на языке Java.

П.В. Камлач

Окончил Белорусский государственный университет информатики и радиоэлектроники. Область научных интересов связана с медицинской электроникой и информационными технологиями в медицине, организацией учебного и научно-исследовательского процессов в техническом университете.

А.В. Слижэва

Обучается в Белорусском государственном университете информатики и радиоэлектроники. Область научных интересов связана с проектированием и производством программно-управляемых электронных средств.

Аннотация. Разработана модель течения газовой смеси в дыхательном контуре аппарата искусственной вентиляции легких. Данная модель может использоваться для расчетов оптимальных значений при проведении искусственной вентиляции легких. Также её можно применять при моделировании различных ситуаций, которые могут возникнуть при лечении пациента.

В данной работе рассматривались случаи, когда трубки прямые и, когда на них оказывалось физическое давление, а также, когда трубки были подвергнуты изменению геометрии (согнуты). Рассчитаны значения избыточного давления в дыхательном контуре аппарата искусственной вентиляции легких. Даны рекомендации по использованию модели.

Ключевые слова: Аппарат искусственной вентиляции легких, моделирование, имитационная модель.

Введение.

Аппараты искусственной вентиляции легких (ИВЛ) – это технические устройства, осуществляющие воздухообмен в дыхательных путях организма. Аппарат ИВЛ формирует требуемый объем газовой смеси, при необходимости обогащает воздушную смесь анестетиком, и подает ее в дыхательный контур, где с помощью клапанов вдоха и выдоха производится однонаправленное движение воздушной смеси [1].

В практике ИВЛ особое значение приобрел метод поддержания положительного давления конца выдоха (ПДКВ). Метод поддержания ПДКВ заключается в том, что в конце выдоха давление в дыхательных путях не снижается до нулевого уровня, а остается выше атмосферного на определенную величину, установленную врачом или медсестрой. Применение умеренного уровня ПДКВ показано всем пациентам, которым проводится ИВЛ, даже при отсутствии явной патологии легких, поскольку это позволяет предупредить нарушение газообмена в легких и улучшить распределение подаваемого газа по легочным путям [2]. Основной задачей

пользователей аппаратов ИВЛ, является не допустить изменения параметров давления в дыхательном контуре. Поэтому требуется проводить различные испытания аппаратов ИВЛ, которые помогут предотвратить внештатные ситуации [3].

Благодаря интенсивному развитию информатики и компьютерных технологий стало возможно решать такие задачи, как моделирование течения газовой смеси в дыхательном контуре аппарата ИВЛ, не используя больших временных и финансовых затрат. Упростить решение такой задачи возможно с использованием моделирования. Одним из наиболее распространенных и удобных способов моделирования сложных систем является имитационное компьютерное моделирование объектов и процессов реального мира [4].

Актуальность.

Так как целью исследования является разработка модели течения газовой смеси в дыхательном контуре аппарата ИВЛ, то предполагается, что данные исследования помогут улучшить дыхательный контур существующих аппаратов искусственной вентиляции лёгких и помогут в разработке новых.

Усовершенствование существующих аппаратов ИВЛ является востребованным, потому что в наше время случается множество катастроф и аварий, в следствии которых люди попадают в больницу в очень тяжелом состоянии, не способные самостоятельно дышать. И в определенные сезоны года, когда происходят вспышки гриппа и различных вирусов, люди часто попадают в больницу, где им необходимо оказать эффективную помощь.

Методика проведения эксперимента.

Моделирование течения газовой смеси в дыхательном контуре аппарата ИВЛ основывается на понятии потока. Поток – это скорость изменения объёма. В респираторной механике поток измеряют в литрах в минуту. Поток создаётся градиентом давлений. Когда поток наталкивается на препятствие, давление увеличивается.

Поток дыхательной смеси в легких должен преодолеть не только эластическое сопротивление самой ткани, но и резистивное сопротивление дыхательных путей. Поскольку трахеобронхиальное дерево представляет собой систему трубок различной длины и ширины, то сопротивление газотоку в легких можно определить по известным физическим законам. В целом, сопротивление потоку зависит от градиента давлений в начале и в конце трубки, а также от величины самого потока.

Уравнение Бернулли, которым описывается поток, имеет некоторые ограничения:

- устойчивый поток;
- несжимаемый поток (что также означает, что плотность постоянна);
- поток без трения;
- поток по обтекаемой линии.

В практических ситуациях проблемы могут быть проанализированы с использованием расширенного уравнения Бернулли [5]:.

$$p_1 + \frac{1}{2} \cdot \rho \cdot v_1^2 + \rho \cdot g \cdot y_1 = p_2 + \frac{1}{2} \cdot \rho \cdot v_2^2 + \rho \cdot g \cdot y_2 + H_L, \quad (1)$$

где H_L – потери напора из-за трения или вязкости;

y_1 и y_2 – высоты входа и выхода;

p_1 и p_2 – давления на входе и выходе;

ρ – плотность среды;

v_1 и v_2 – скорости на входе и выходе;

g – ускорение свободного падения.

Для оценки течения газовой смеси в дыхательном контуре аппарата искусственной вентиляции лёгких была разработана геометрическая модель, состоящая из маски,

теплообменник (влагообменник), разделителя и трубок (прямой (рисунок 1) и изогнутой (рисунок 2)).

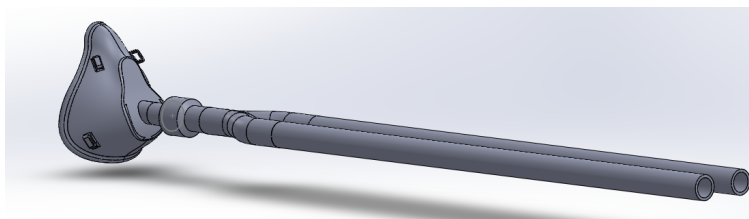


Рисунок 1. Геометрическая модель дыхательного контура аппарата ИВЛ с прямыми трубками

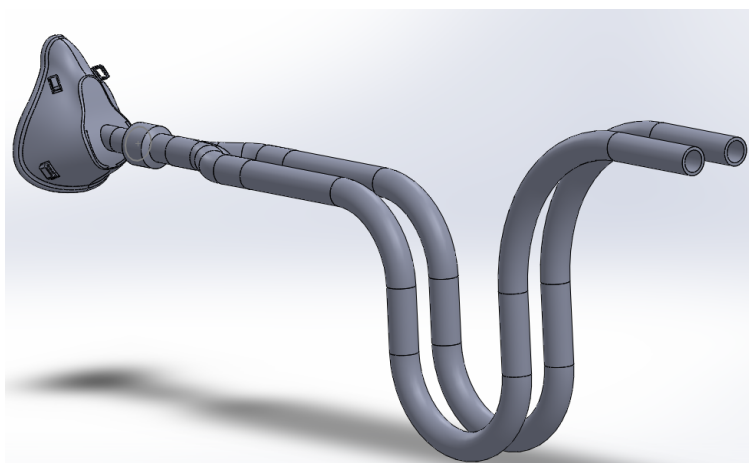


Рисунок 2. Геометрическая модель дыхательного контура аппарата ИВЛ с изогнутыми трубками

Для моделирования процесса деформации трубок, которые могут происходить на практике в результате случайного сдавливания или перегиба, использовалась система уравнений для расчета изменения геометрических параметров трубок:

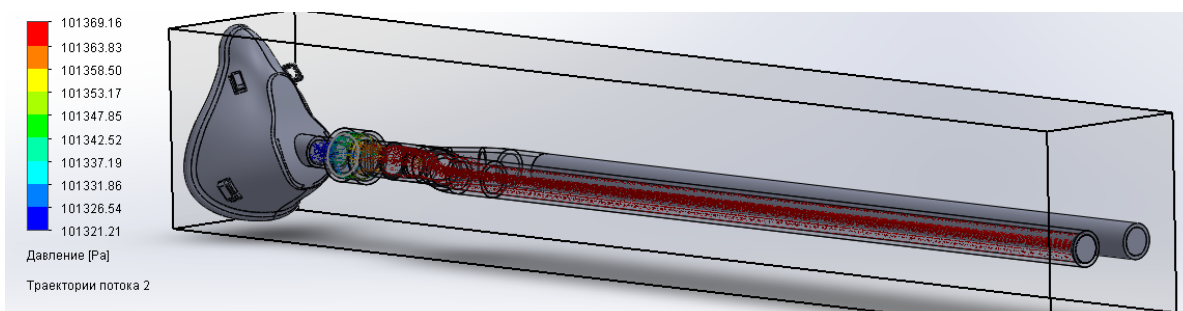
$$\Delta x = \frac{m \cdot g}{k}, \quad (2)$$

$$\Delta l = \frac{\Delta x}{2}, \quad (3)$$

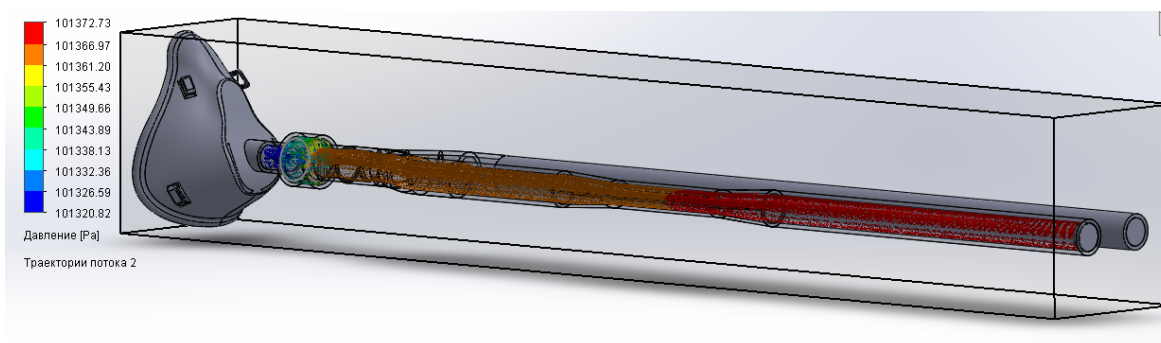
где Δx – изменение ширины трубки,
 Δl – изменение высоты трубки,
 m – масса;
 g – ускорение свободного падения;
 k – модуль упругости полиэтилена.

Результаты и их обсуждение.

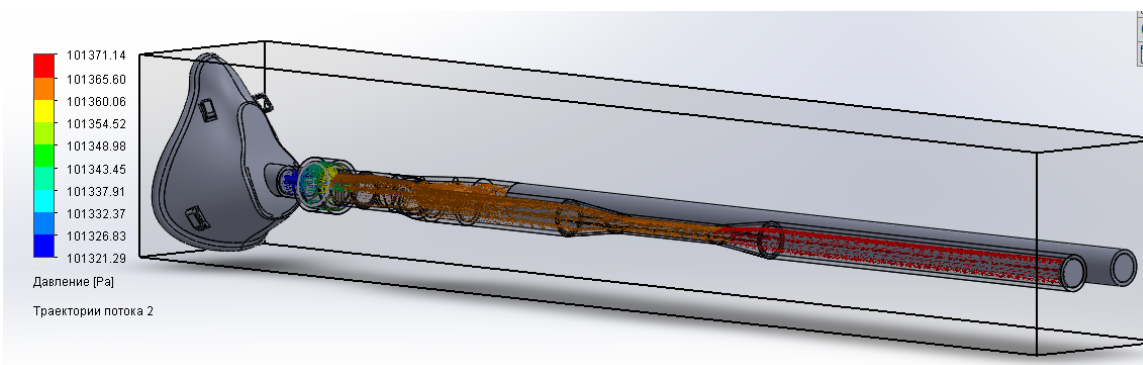
При моделировании задавались граничные условия: давление окружающей среды на входе маски 101325 Па, начальная скорость потока 60 л/мин, объемный расход на входе будет равен 0,001 м³/с. Для полноты результатов моделирования расчеты проводились, также для скоростей потока – 30 л/мин, объемный расход на входе будет равен 0,0005 м³/с, и 120 л/мин, объемный расход на входе будет равен 0,002 м³/с. На рисунке 3 представлены результаты моделирования течения газовой смеси в дыхательном контуре аппарата ИВЛ с прямыми трубками в *SolidWorks*.



a



b



c

Рисунок 3. Результаты компьютерного моделирования течения газовой смеси в дыхательном контуре с прямыми трубками со скоростью 30 л/мин.: a – без давления на трубку; b – с давлением 17 кг; c – с давлением 33,5 кг.

Были рассчитаны параметры (давление и скорость) течения газовой смеси в дыхательном контуре аппарата ИВЛ при различных начальных скоростях потока (30, 60 и 120 л/мин). Результаты моделирования избыточного давления на месте надавливания на трубку дыхательного контура отображены в таблице 1.

Таблица 1. Результаты расчёта избыточного давления в трубке при надавливании

Заданная скорость потока, л/мин	Расчетная величина избыточного давления в трубке, Па		
	На трубку не надавливали	На трубку надавили 17 кг	На трубку надавили 33,5 кг
30	44,16 Па	41,97 Па	40,60 Па
60	155,43 Па	152,58 Па	129,94 Па
120	581,48 Па	585,53 Па	498,16 Па

На рисунке 4 представлен график зависимости избыточного давления от массы надавливания.

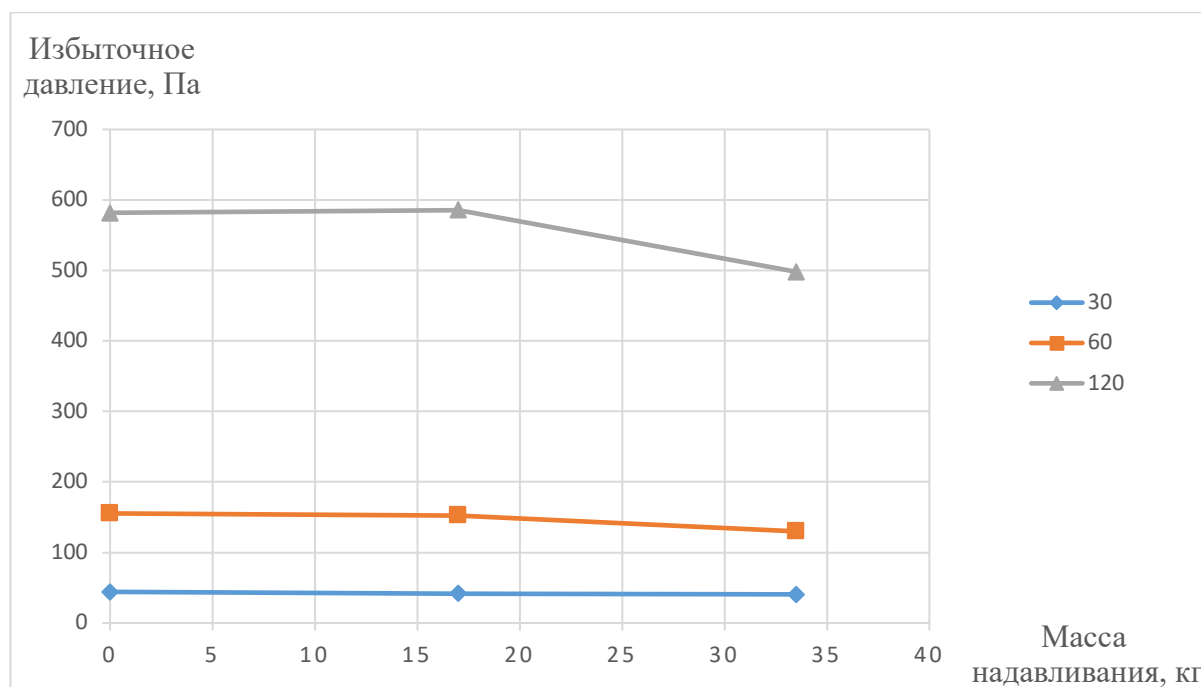
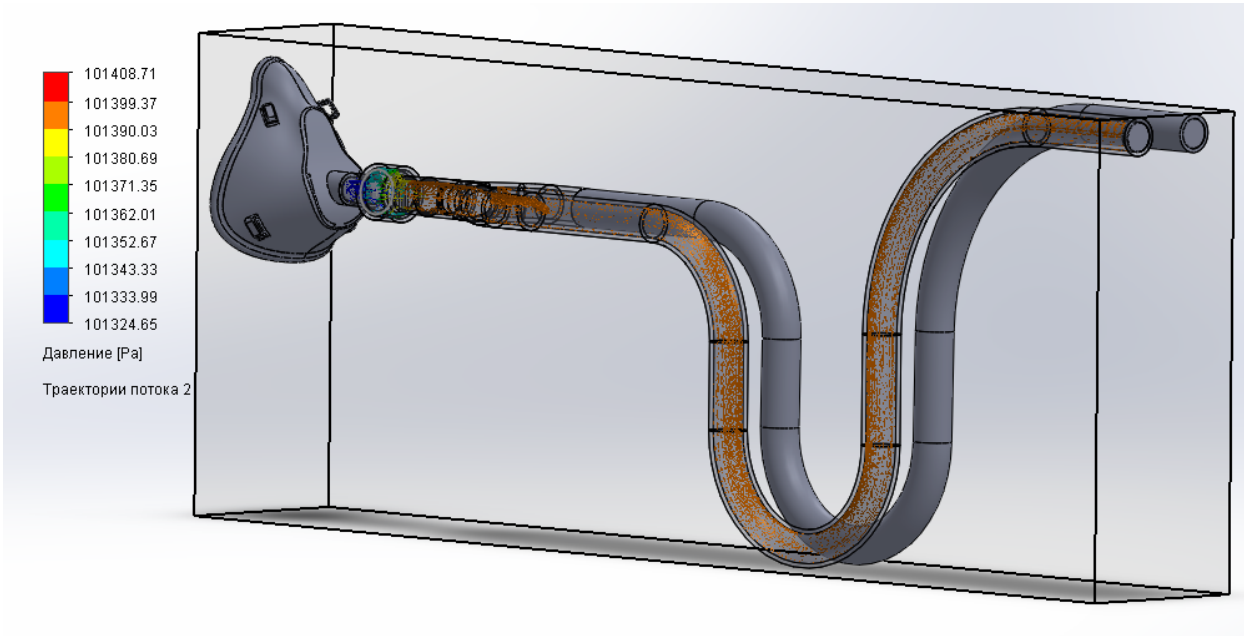
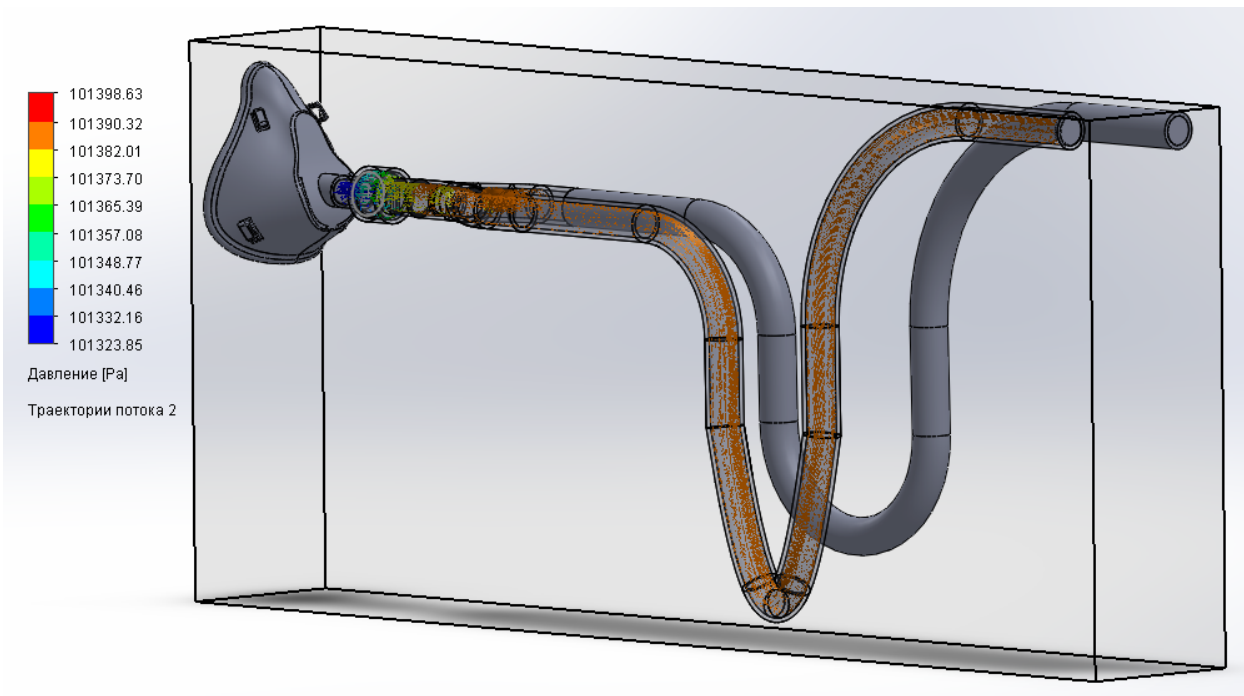


Рисунок 4. Зависимость избыточного давления от массы надавливания

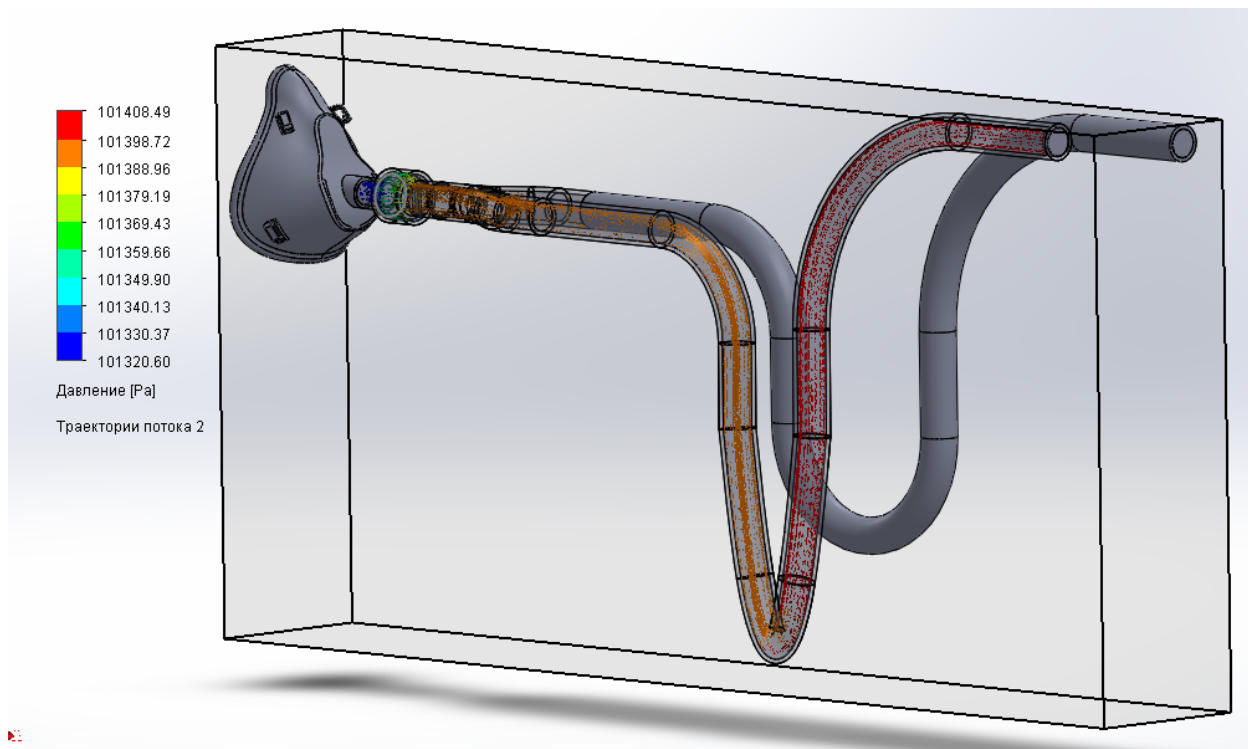
На рисунке 5 представлены результаты моделирования течения газовой смеси в дыхательном контуре аппарата ИВЛ с изогнутыми трубками в *SolidWorks*.



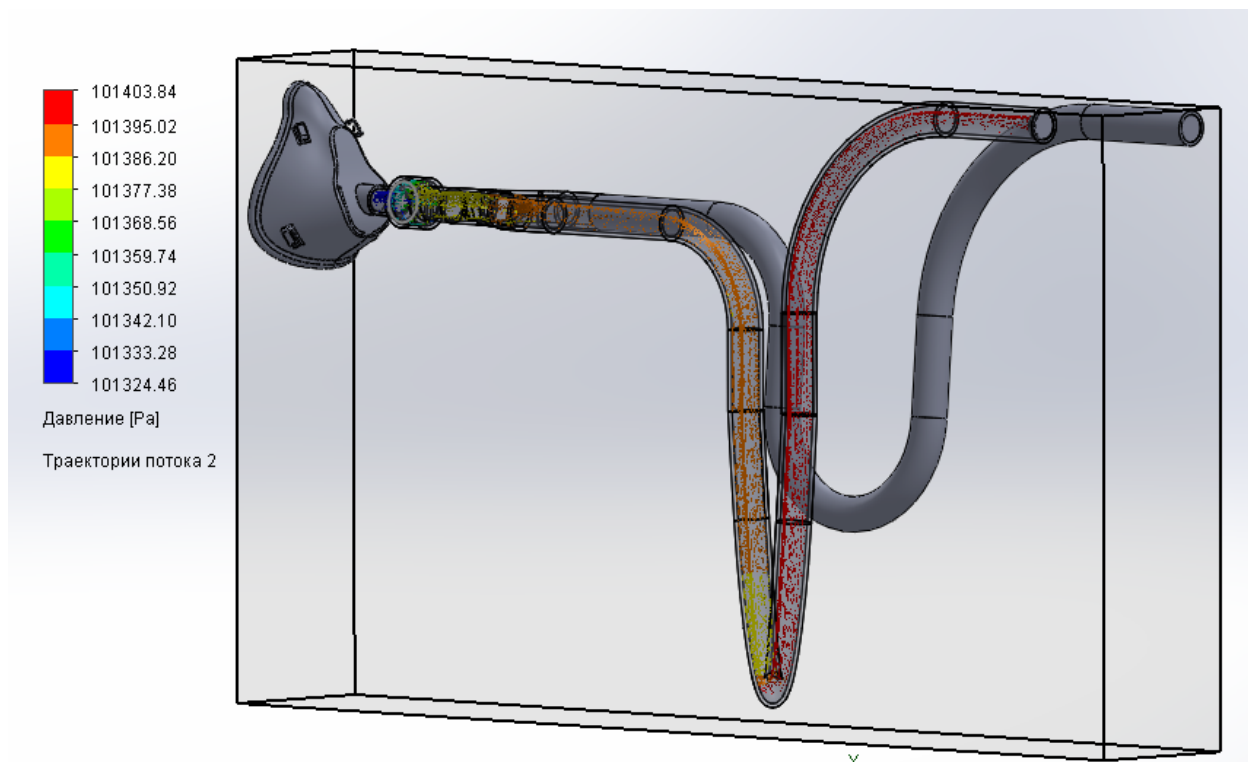
a



b



c



d

Рисунок 5. Результаты компьютерного моделирования течения газовой смеси в дыхательном контуре с изогнутыми трубками со скоростью 30 л/мин.: a – трубку согнули на 1,85 мм; b – трубку согнули на 4,69 мм; b – трубку согнули на 9,48 мм; c – трубку согнули на 75 мм
Результаты моделирования избыточного давления на месте изгиба трубки дыхательного

контура отображены в таблице 2.

Таблица 2. Результаты расчёта избыточного давления в трубке при надавливании

Заданная скорость потока, л/мин	Расчетная величина избыточного давления в трубке, Па			
	Трубку согнули на 1,85 мм	Трубку согнули на 4,69 мм	Трубку согнули на 9,48 мм	Трубку согнули на 75 мм
30	61,20 Па	73,72 Па	65,32 Па	74,37 Па
60	228,26 Па	275,18 Па	236,01 Па	283,30 Па
120	895,79 Па	1074,47 Па	915,65 Па	1108,57 Па

На рисунке 6 представлен график зависимости избыточного давления от изгиба.

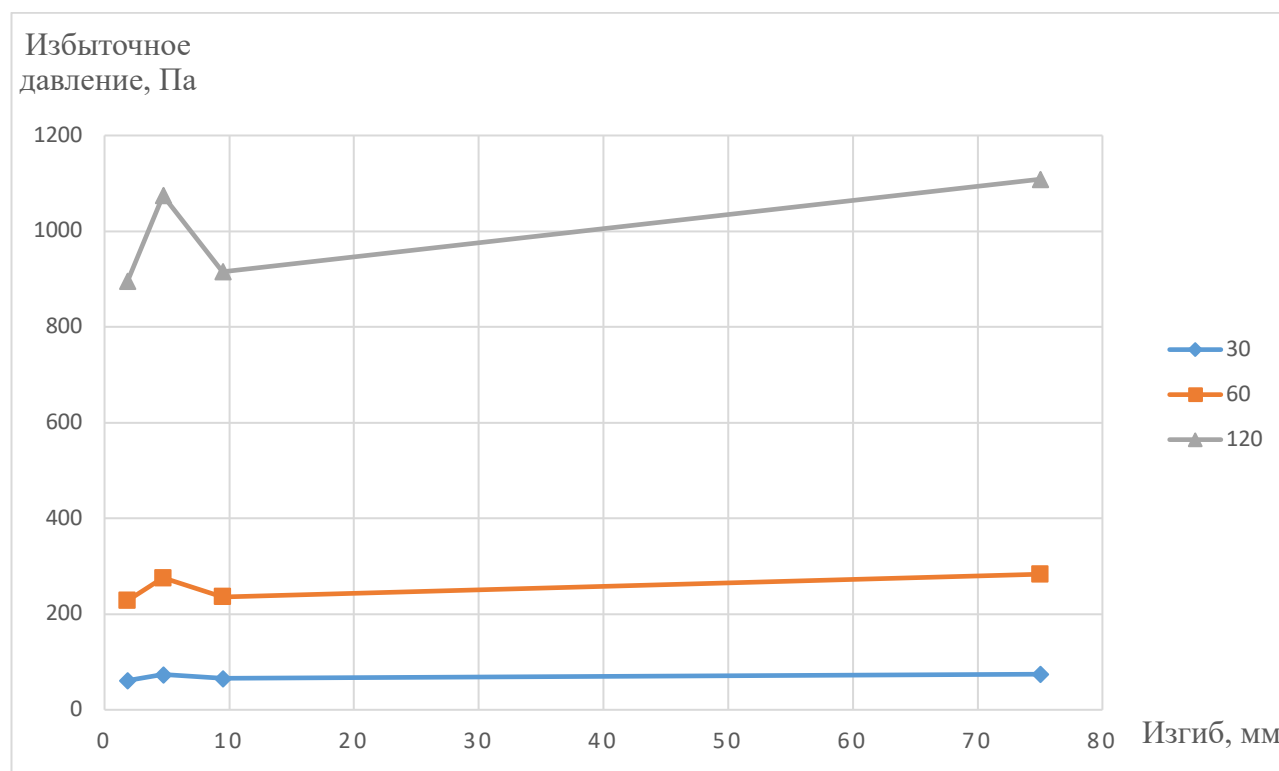


Рисунок 6. Зависимость избыточного давления от изгиба

По результатам расчётов видно, что изменение давления в трубке происходит неравномерно и не пропорционально. Это обусловлено разными геометрическими параметрами при различном давлении на трубку и различных изгибах.

Заключение.

Разработана модель течения газовой смеси в дыхательном контуре аппарата ИВЛ. Данная модель может использоваться для расчетов оптимальных значений при ИВЛ. Также её можно применять при моделировании различных ситуаций, которые могут возникнуть при лечении

пациента. В данной работе рассматривались случаи, когда трубки прямые и на них надавили чем-то тяжёлым, а также, когда трубки согнуты. Для недопущения таких ситуаций применять специальные крепежи к аппаратам ИВЛ, которые будут держать трубки в подвешенном состоянии или использовать особые конструкции трубок.

Список литературы

- [1] Брыгин П.А. Методы и режимы современной искусственной вентиляции легких. Ульяновск: Ульяновский государственный университет; 2000.
- [2] Бурлаков Р.И., Гальперин Ю.Ш., Юревич В.М. Искусственная вентиляция легких: принципы, методы, аппаратура. М.: Медицина; 1986.
- [3] Имитационная модель аппарата искусственной вентиляции лёгких = Simulation Model of the Artificial Lung Ventilation Apparatus / П. В. Камлач [и др.] // Доклады БГУИР. – 2023. – Т. 21, № 1. – С. 104-108.
- [4] Эльберг М.С., Цыганков Н.С. Имитационное моделирование: учеб. пособие. Красноярск : Сиб. федер. ун-т; 2017.
- [5] Jonuskaite A. Flow Simulation with SolidWorks. Helsinki : Arcada University of Applied Sciences; 2017.

SIMULATION MODELING OF THE VENTILATOR ARTIFICIAL LUNG VENTILATION

M.I. Kuprianau

*Master student of the Department of
Electronic Engineering and
Technology, Belarusian State
University of Informatics and
Radioelectronics.*

P.V. Kamlach

*PhD, Associate Professor, Associate
Professor of the Department of Electronic
Engineering and Technology of the
Belarusian State University of Informatics
and Radioelectronics*

A.V. Slizhova

*Student of the Department of
Electronic Engineering and
Technology, BSUIR*

*Department of Electronic Engineering and Technology
Faculty of Computer Engineering
Belarusian State University of computer science and Radio Electronics, Republic of Belarus
E-mail: n.kupriianov@bsuir.by*

Abstract. A model for the flow of a gas mixture in the respiratory circuit of an artificial lung ventilation apparatus has been developed. This model can be used to calculate the optimal values for artificial lung ventilation. It can also be used to simulate various situations that may arise during the treatment of a patient.

In this paper, the cases were considered when the tubes were straight and when they were subjected to physical pressure, as well as when the tubes were subjected to a change in geometry (bent). The values of excess pressure in the respiratory circuit of the artificial lung ventilation apparatus were calculated. Recommendations on the use of the model are given.

Keywords: Ventilator, modeling, simulation model.

УДК 004.67

УПРАВЛЕНИЕ ПЛАНИРОВАНИЕМ И ПРОГНОЗИРОВАНИЕМ ФИНАНСОВ В АВТОМАТИЗИРОВАННОЙ ПРОГРАММЕ 1С ERP С ИСПОЛЬЗОВАНИЕМ BIG DATA



Н.В. Новицкая

Магистрант кафедры проектирования
информационно-компьютерных систем
БГУИР n.novitskaia@bsuir.by



С.В. Матийко

Студент кафедры информатики БГУИР
cvity6692@gmail.com

Н.В. Новицкая

Окончила Белорусский государственный университет информатики и радиоэлектроники. Область научных интересов связана с разработкой автоматизированных программ управления, организацией рабочих процессов в техническом университете.

Матийко С. В.

Студент 4 курса специальности «Программное обеспечение информационных технологий» БГУИР

Аннотация: Рассмотрены возможности интеграции больших данных в программу 1С ERP для улучшения управления финансами компаний. В работе рассмотрены такие темы как: управление бюджетом, управление кредитным риском, управление инвестициями, управление рисками и управление налогами. Рассмотрены конкретные примеры использования программы 1С ERP с интеграцией Big Data для улучшения управления финансами компаний. Работа также обращает внимание на необходимость наличия соответствующих ресурсов для успешной интеграции и использования больших данных в управлении финансами компаний.

Ключевые слова: 1С ERP, BIG DATA, данные, управление, система, финансы, анализ, средства, прогнозирование, эффективность.

Введение.

В условиях современной экономики, где бизнес-процессы становятся все более сложными и быстро меняются, необходимо иметь мощный инструмент для управления финансами компании. Один из таких инструментов – автоматизированная программа управления предприятием (ERP), которая позволяет интегрировать все бизнес-процессы в единую систему. Однако, с появлением больших данных, возможности ERP значительно расширяются. Ниже будет рассмотрено управление планированием и прогнозированием финансов в автоматизированной программе 1С ERP в условиях больших данных.

Программа 1С ERP с интеграцией Big Data предоставляет множество возможностей для сбора, хранения и анализа данных, необходимых для прогнозирования и планирования финансовой деятельности предприятия. Такие данные могут включать в себя финансовые отчеты, данные по продажам, данные по затратам и другие. Программа 1С ERP позволяет создавать автоматизированные бюджетные планы, которые основаны на Big Data аналитике. Бюджетные планы могут быть настроены на любой период времени, и они могут быть использованы для прогнозирования будущих доходов и затрат предприятия. Программа 1С ERP с интеграцией Big Data предоставляет возможность анализировать данные в режиме реального времени, что позволяет быстро реагировать на изменения в экономической ситуации. Это позволяет принимать более обоснованные решения по управлению финансами.

1С ERP позволяет создавать отчеты и анализировать данные с помощью графиков и диаграмм, что делает данные более понятными и доступными для менеджеров и других

сотрудников предприятия [1]. 1С ERP с интеграцией Big Data предоставляет возможность автоматической классификации данных и применения алгоритмов машинного обучения для прогнозирования финансовых показателей. Это уменьшает риски ошибок и повышает точность прогнозов.

Программа 1С ERP с интеграцией Big Data является мощным инструментом для управления финансами предприятия. Она позволяет управлять рисками, оптимизировать затраты и принимать более обоснованные решения на основе анализа данных. Такой подход к управлению финансами предприятия помогает повысить эффективность бизнеса и увеличить прибыль [2].

1С ERP также позволяет автоматизировать процесс управления запасами и производственными ресурсами. Это помогает оптимизировать производственные процессы и уменьшить затраты на складирование и переработку материалов.

Предоставляется возможность для мониторинга и анализа финансовых операций, что позволяет снизить риски финансовых потерь. Другой модуль - «Управление финансовым риском» - позволяет проводить анализ финансового риска компании и оценку ее платежеспособности, а также прогнозировать вероятность невозврата задолженности (рисунок 1).



Рисунок 1. Пример информационной панели показателей

1С ERP обеспечивает более высокую степень автоматизации и точности управления финансами, чем традиционные методы. Она позволяет сэкономить время, затраченное на ручной ввод данных, и повысить качество принимаемых решений. В целом, управление финансами с использованием программы 1С ERP и Big Data аналитики позволяет предприятию улучшить свою финансовую производительность и повысить эффективность своих операций.

При использовании программы 1С ERP с интеграцией Big Data, данные можно собирать и анализировать из разных источников, включая внешние данные, такие как социальные медиа и демографические данные [3]. Это позволяет предприятию получить более широкий обзор рынка и узнать о новых тенденциях, которые могут влиять на бизнес. Анализ больших данных позволяет выявлять тенденции и определять эффективность различных стратегий и тактик, что помогает оптимизировать бизнес-процессы и улучшить результаты.

1С ERP с интеграцией Big Data аналитики может быть особенно полезным для крупных предприятий, которые имеют множество отделов и подразделений. Это помогает координировать

действия всех подразделений, упрощать процессы и обеспечивать более гладкую работу всей компании.

Одной из основных задач планирования и прогнозирования финансов является определение бюджета на следующий год. С помощью программы 1С ERP с интеграцией Big Data, компания может анализировать данные из различных источников, включая исторические данные, данные о рынке, экономические данные и данные о производственных мощностях [5]. Это помогает предприятию более точно прогнозировать будущие финансовые результаты и разрабатывать бюджет на следующий год [7].

Еще одним важным аспектом управления финансами является управление капиталом. С помощью программы 1С ERP, предприятие может анализировать данные о результатах своей деятельности и определять, какие проекты или направления бизнеса требуют большего капитала. Также программа позволяет предприятию анализировать данные о конкурентах и рынке, что помогает принимать решения по управлению капиталом [4].

Для управления результативностью компании необходимо отслеживать ход реализации проектов, выполнение задач, поручений и инициатив, сформированных в рамках стратегических целей компании. Программный продукт позволяет управлять всеми оперативными поручениями и задачами, направленными на улучшение конкретных целевых показателей. Долгосрочные задачи могут быть внесены в систему как проекты, с детализацией по этапам и представлены в виде календарного графика - диаграммы Ганта (рисунок 2). За каждый этап или веху проекта могут быть назначены ответственные сотрудники.

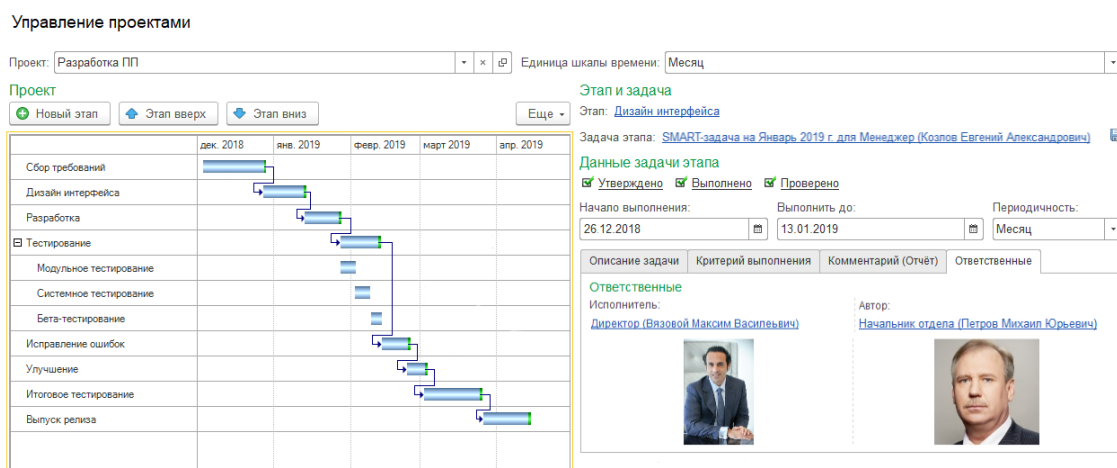


Рисунок 2. Пример информационной панели показателей

Программа 1С ERP с интеграцией Big Data позволяет компаниям управлять налоговыми расходами и создавать отчетность в соответствии с налоговыми требованиями. Компании могут использовать данные о налоговых изменениях, чтобы принимать более эффективные решения и снижать налоговые риски. Важным аспектом управления финансами является также управление рисками. С помощью программы 1С ERP, предприятие может анализировать данные о рисках, связанных с финансовой деятельностью, и принимать меры для их снижения. Это позволяет предотвратить возможные финансовые потери и улучшить общую финансовую стабильность предприятия.

Заключение. Таким образом, использование программы 1С ERP с интеграцией Big Data позволяет компаниям более эффективно управлять своими финансами, планировать свою деятельность, достигать поставленных целей и улучшить свои результаты за счет более точного прогнозирования будущих финансовых результатов, управления капиталом, анализа показателей эффективности бизнеса, управления рисками и автоматизации процессов. Это помогает

предприятиям повышать свою финансовую стабильность, эффективность и получить более точное представление о том, какой будет ее финансовое состояние в будущем, и принимать на основе этой информации решения о распределении ресурсов и управлении финансовыми потоками. Однако, для того чтобы эффективно использовать эти возможности, компании должны иметь соответствующие ресурсы, такие как обученный персонал и подходящую инфраструктуру для сбора, обработки и анализа больших данных.

Важным аспектом управления финансами является возможность более точно определить потребность в финансовых ресурсах в будущем и предпринимать меры для обеспечения этих ресурсов. Это позволяет компании укреплять свою финансовую позицию, обеспечивая устойчивый рост и развитие. Наконец, управление прогнозированием финансов позволяет компаниям получать более точные данные о своей финансовой деятельности и использовать эти данные для принятия управленческих решений. Анализ и прогнозирование финансовых показателей позволяет компаниям быстрее и более точно оценивать эффективность своих инвестиций, управлять своими расходами и доходами, а также определять оптимальную стратегию развития.

Список литературы

- [1] 1С:ERP Управление предприятием [Электронный ресурс]. – Режим доступа: <https://1c.ru/erp/>. – Дата доступа: 01.04.2023.
- [2] Черный, Д.В. Управление финансами: учебник / Д.В. Черный, Л.А. Карасева. – СПб. : Питер, 2019. – 336 с.
- [3] Big Data in Finance: From Predictive Analytics to Risk Management. - Wiley, 2021. - 448 с.
- [4] Попова И.Ю. Управленческий учет в 1С:ERP/ Попова И.Ю., Старкова Н.М. - М.: Издательский дом "НОРМА-ИНФРА", 2019. - 256 с.
- [5] Андреева О.А. Финансовый анализ в 1С:ERP/ Андреева О.А., Беляева Н.А. - М.: Издательство "Эксмо", 2020. - 224 с.
- [6] Айрапетян В. В. Пути совершенствования оперативного планирования на предприятии //Современные научные исследования и разработки, 2019. - No 1 (30). - С. 76 - 80.
- [7] Ибрагимов Р. И. Совершенствование финансового планирования в организации // Наукосфера, 2021. - No 4 -2. - С. 222 - 225.

MANAGEMENT OF PLANNING AND FINANCIAL FORECASTING IN THE AUTOMATED 1C ERP PROGRAM USING BIG DATA

N.V. Navitskaya

master student, Department of Information Computer Systems Design BSUIR

S. V. Matiyko

Student of Belarusian State University of Informatics and Radioelectronics

*Department of Information and Computer Systems Design
Faculty of Computer Engineering*

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

E-mail: alexvikt.minsk@gmail.com

Abstract: This thesis is devoted to the topic of financial planning and forecasting management in the 1C ERP automated program using Big Data. The possibilities of integrating big data into the 1C ERP program to improve financial management of companies are considered. The paper covers such topics as: budget management, credit risk management, investment management, risk management and tax management. Specific examples of using the 1C ERP program with Big Data integration to improve financial management of companies are considered. The work also draws attention to the need for appropriate resources for the successful integration and use of big data in corporate financial management.

Keywords: 1C ERP, BIG DATA, data, management, system, finance, analysis, tools, forecasting, efficiency.

УДК 621.3.049.77 : 621.373.8

МОДЕЛИРОВАНИЕ ЛАЗЕРНОГО НАГРЕВА ПРИ ДВИЖЕНИИ ЛАЗЕРА ПО ЗАДАННОЙ ТРАЕКТОРИИ



А.И. Лапко

Старший преподаватель кафедры ИТАС
lapko@bsuir.by



В.Л. Ланин

Профессор кафедры электронной техники и технологии БГУИР, доктор технических наук

А.И. Лапко

Окончил Белорусский государственный университет информатики и радиоэлектроники. Область научных интересов связана с исследованием технологических направлений реализации многокристальных модулей.

В.Л. Ланин

Окончил Минский радиотехнический институт. Профессор кафедры электронной техники и технологии. Автор 10 монографий, имеет 45 летний опыт работы в области технологии сборки электронной аппаратуры.

Аннотация. Выполнено моделирование перемещения сфокусированного лазерного луча по двум вариантам траектории схемы для получения отверстия в кремниевой пластине, необходимой для формирования контактных соединений между компонентами 3D-модуля. Произведен выбор оптимальной траектории движения, обеспечивающий минимальную затрату времени и исключающий перегрев изделия

Ключевые слова: моделирование, лазерный луч, траектория движения.

Введение.

Среди перспективных технологических направлений реализации многокристальных модулей является сборка 3D-модулей. Особенностью данной технологии является расположение компонентов сборки не только на одной плоскости, но и по вертикали. Формирование контактных соединений между компонентами 3D-модуля может осуществляться с помощью проволочного монтажа, а также по технологии поверхностного монтажа. Для технологии поверхностного монтажа необходимо формирование переходных отверстий в кремниевой пластине.

Технология формирования отверстий в кремниевой подложке является сравнительно новой, но быстроразвивающейся и эффективной. Применение лазерного излучения для формирования таких отверстий дает возможность уменьшения диаметра до десятков микрон, что повышает плотность элементов в 3D электрических модулях [1]. Преимуществами использования лазерного нагрева является возможность точного контроля технологических режимов и объема подводимой энергии, бесконтактное воздействие. Высокая плотность энергии лазерного излучения позволяет за небольшие промежутки времени достигать высоких температур нагрева, в частности температуры плавления различных материалов, что потенциально позволяет на одном технологическом оборудовании выполнять широкий спектр технологических операций.

Основной проблемой формирования отверстий лазерным излучением является их конусообразность, которая зависит как от длины лазерного луча, так и от режимов обработки.

Моделирование лазерного нагрева при сверлении отверстий.

Для сверления микроотверстий используется способ перемещения сфокусированного лазерного луча по заданной траектории при точным перемещении стола с пластиной на низкой скорости относительно неподвижного лазерного луча. Для моделирования траектории движения лазера предложено два варианта.

Вариант 1 согласно формуле 1.1 с движением луча по гипотрохоиде, визуализация которого представлена на рисунке 1а,

$$\begin{cases} x=(r-m \times r) \times \cos(m \times t)+h \times \cos(t-m \times t) \\ y=(r-m \times r) \times \sin(m \times t)-h \times \sin(t-m \times t), \end{cases} \quad (1.1)$$

где: t – время; x, y – координаты; r, m – коэффициенты.

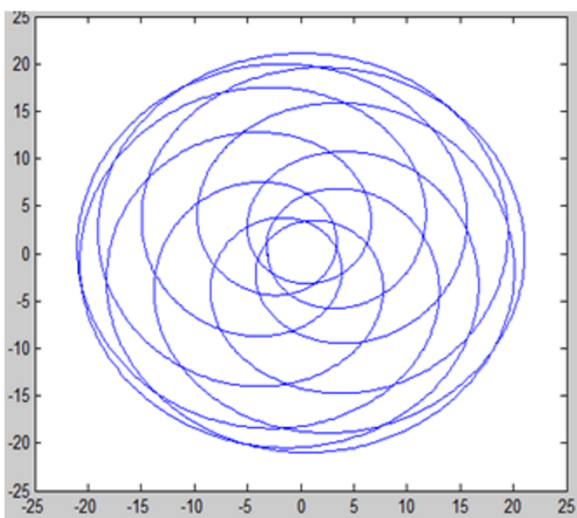
Вариант 2 согласно формуле 1.2 с движением по спирали, визуализация движения представлена на рисунке 1.1, б,

$$\begin{cases} x=c+a \times 2t+\pi \\ y=v+a \times t \times \sin(t+q), \end{cases} \quad (1.2)$$

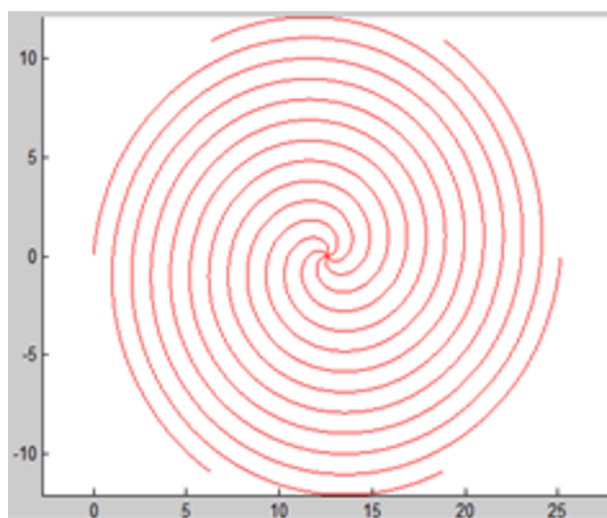
где: c, v, q – коэффициенты.

Коэффициенты c и v задаются уравнением:

$$c=4 \pi \times \cos \theta., \quad v=4 \pi \times \sin \theta. \quad (1.3)$$



а



б

Рисунок 1. Визуализация движения лазера при формировании отверстия (а – вариант 1 с движением луча по гипотрохоиде; б – вариант 2 с движением по спирали)

Для моделирования процесса лазерного нагрева применялся программный комплекс COMSOL Multiphysics 53а. Перед началом моделирования необходимо задать глобальные переменные, которые будут участвовать в расчетах. Использовалась кремниевая пластина диаметром 40 мм и толщиной 275 мкм. Лазер мощностью 10 Вт, длительность импульса – 1,2 нс, частота – 10–300 кГц. Коэффициент излучения поверхности пластины составляет примерно 0,8. Для действующей длины волны лазера предполагается, что поглощающая способность равна излучающей способности. Следовательно, все лазерное излучение будет уходить на нагрев пластины. Диаметр формируемого отверстия 0,1 мм. Расчет лазерного излучения при формировании микроотверстий приведен в источнике [2].

В результате моделирования получены температурно-временные зависимости (рисунок 2) и тепловые поля на поверхности кремниевой пластины (рисунок 3).

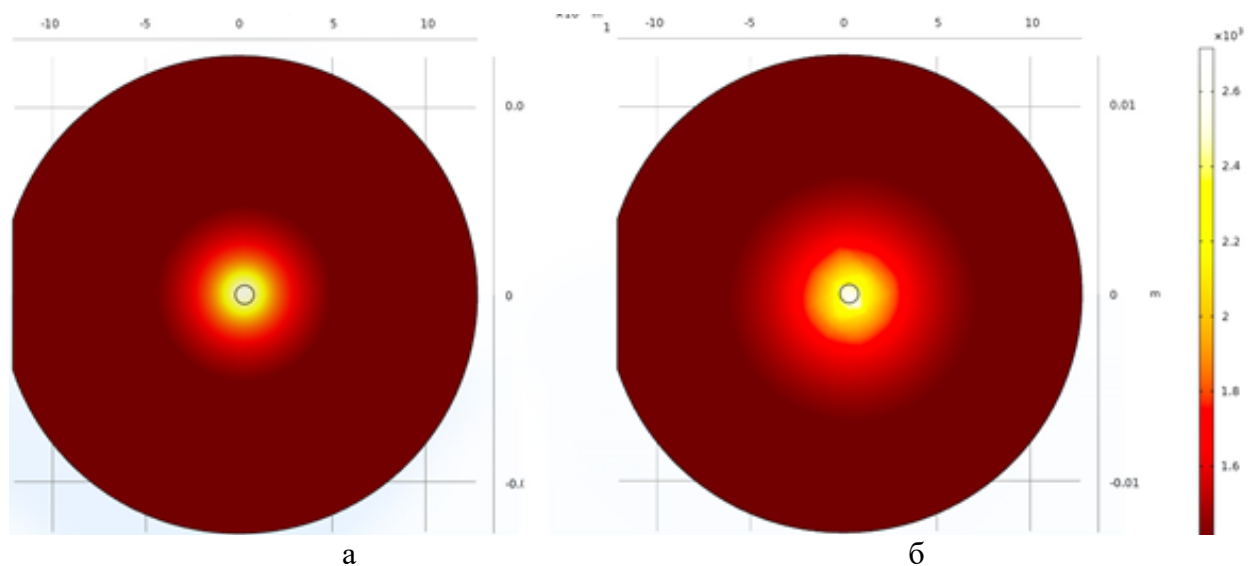


Рисунок 2. Распределение тепловых полей при формировании отверстий (а – вариант 1 с движением луча по гипотрохоиде; б – вариант 2 с движением по спирали)

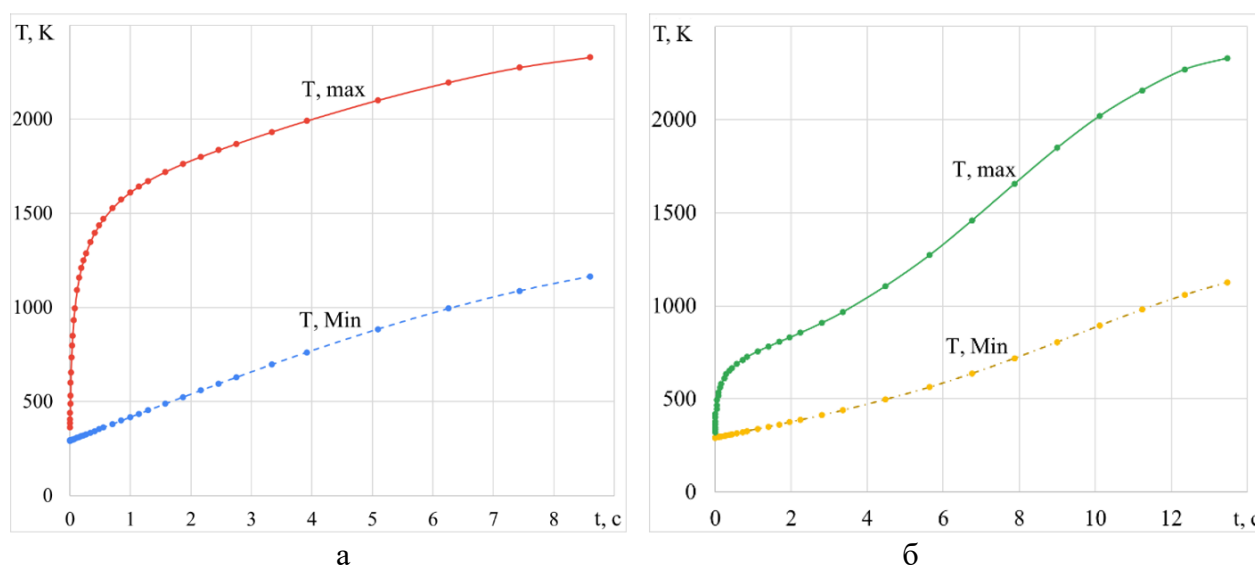


Рисунок 3. Зависимости температуры от времени при формировании отверстий (а – вариант 1 с движением луча по гипотрохоиде; б – вариант 2 с движением по спирали)

В результате моделирования лазерного нагрева при формировании отверстий получены следующие результаты: при движении луча по гипотрохоиде (вариант 1) продолжительность операции составила 8,6 с, для движения по спирали (вариант 2) – 13,5 с. Минимальная температура нагрева пластины для обоих вариантов составила 1200°C.

По распределению тепловых полей при формировании отверстий проведена оценка площади локального перегрева пластины. В случае с движением луча лазера по спирали площадь перегрева пластины получилась в 1,8 раза больше, чем для варианта движения по гипотрохоиде.

Исходя из сравнения форм графиков зависимости температуры от времени при формировании отверстий в исследуемых вариантах движения можно судить о сравнительно низкой эффективности использования траектории движения луча по спирали.

Заключение.

Применение перемещения сфокусированного лазерного луча по траектории вызывает проблему выбора оптимального пути, который обеспечит получение отверстий оптимального качества с минимальной затратой времени и исключением перегрева изделия. На данный момент лучшие показатели достигнуты при движении луча по гипотрохоиде (вариант 1), что не исключает необходимости дальнейших исследований в этой области, уточнении моделей и экспериментов для практического подтверждения.

Список литературы

[1]. Ланин, В.Л. Формирование отверстий в кремниевой подложке 3D электронного модуля лазерным излучением / В.Л. Ланин, В.Т. Фам, А.И. Лаппо // Доклады БГУИР, 2021, №3. –С. 58–65.

[2]. Лаппо, А.И. Моделирование процесса лазерной прошивки отверстий в кремнии при формировании 3D структур / А.И. Лаппо, Т.С. Боброва, О.В. Кузнецова / Материалы международной научной конференции «ITS-2019» Минск: БГУИР, 2019. – С. 232-233.

QUALITY EVALUATION OF INFORMATION TRANSFER IN A DISPATCHING SYSTEM BASED ON MQTT ARCHITECTURE

A.I. Lappo

Assistant Professor of ITAS Department

V.L. Lanin

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

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

E-mail: lappo@bsuir.by

Abstract. The simulation of the movement of a focused laser beam along two variants of the circuit trajectory to obtain a hole in a silicon wafer necessary for the formation of contact connections between the components of a 3D module is performed. The choice of the optimal trajectory of movement was made, which ensures the minimum time consumption and excludes overheating of the product.

Keywords: modeling, laser beam, motion trajectory.

УДК 004.438:004.85

АКТУАЛЬНОСТЬ ИСПОЛЬЗОВАНИЯ ЯЗЫКА JAVASCRIPT ДЛЯ РАЗРАБОТКИ МОДЕЛЕЙ МАШИННОГО ОБУЧЕНИЯ



Г.А. Пискун

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



В.Ф. Алексеев

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



Т.М. Воронко

Инженер-программист Центра информатизации и инновационных разработок БГУИР, магистрант БГУИР
voronko232001@gmail.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.

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.

Т.М. Воронко

Магистрант 1 курса специальности «Электронные системы и технологии» кафедры проектирования информационно-компьютерных систем. Инженер-программист в отделе инновационных разработок в сфере образования Центра информатизации и инновационных разработок Белорусского государственного университета информатики и радиоэлектроники.

Аннотация. Выполнен анализ применимости и актуальности использования языка программирования *JavaScript* для разработки моделей машинного обучения. Рассмотрены основные преимущества и недостатки данного метода работы с искусственным интеллектом в сравнении с аналогами, такими как *Python*, а также описаны особенности *JavaScript*-библиотеки *TensorFlow.js*, с предоставлением примера программы для построения модели, предсказывающей координаты точки по заданным параметрам. Установлено, что использование *JavaScript* может быть полезно для легких и нетребовательных к ресурсам кроссплатформенных веб-приложений, использующих модели машинного обучения.

Ключевые слова: машинное обучение, библиотека, искусственный интеллект.

Введение.

На сегодняшний день, большинство разработчиков для создания моделей машинного обучения отдадут предпочтение таким языкам, как *Python*, *R* и *C++*, в виду зрелости их экосистем по работе с искусственным интеллектом (ИИ). Однако последние разработки в области *JavaScript* делают этот язык все более привлекательным для разворачивания моделей машинного обучения. Вполне возможно, что через несколько лет мы увидим значительный рост популярности использования данного языка для разработки подобных моделей, особенно по мере того, как ноутбуки и мобильные устройства становятся все более мощными, а популярность самого *JavaScript* растет [1].

Так как *JavaScript* – это в первую очередь язык программирования, который даёт возможность реализовывать сложное поведение веб-страницы, как, к примеру, динамически изменяющийся *UI* или управление бизнес-логикой, внедрение машинного обучения в проекты,

написанные на данном языке, может быть очень полезно, положительно влияя на скорость и простоту разработки, а также на дальнейшее управление этими проектами [2].

Рассмотрим основные недостатки и преимущества использования данного языка для разработки моделей машинного обучения.

Недостатки и преимущества использования *JavaScript*.

В качестве основных недостатков использования *JavaScript* для машинного обучения, препятствующих увеличению его популярности в данной сфере, на сегодняшний день можно обозначить следующие характеристики:

1. **Низкая скорость работы с данными.** *JavaScript* – это однопоточный язык, а это означает, что модель машинного обучения может получать только одну часть информации за раз, что делает процесс обучения для этих моделей более длительным по сравнению с аналогами, написанными на других языках [3].

Однопоточность также может вызывать задержки или зависания системы при возникновении трудностей с выполнением определенного кода. Это одна из основных причин, по которой многие разработчики выбирают *Python*, который может похвастаться ускоренным процессом обучения и, как правило, работает быстрее, что ускоряет разработку приложений [3].

На рисунке 1 представлен график, отображающий время, которое тратят модели, написанные на *JavaScript* и *Python*, на обучение и прогнозирование [4], из которого можно сделать вывод, что модели на *Python* работают в разы быстрее, чем на *JavaScript*.

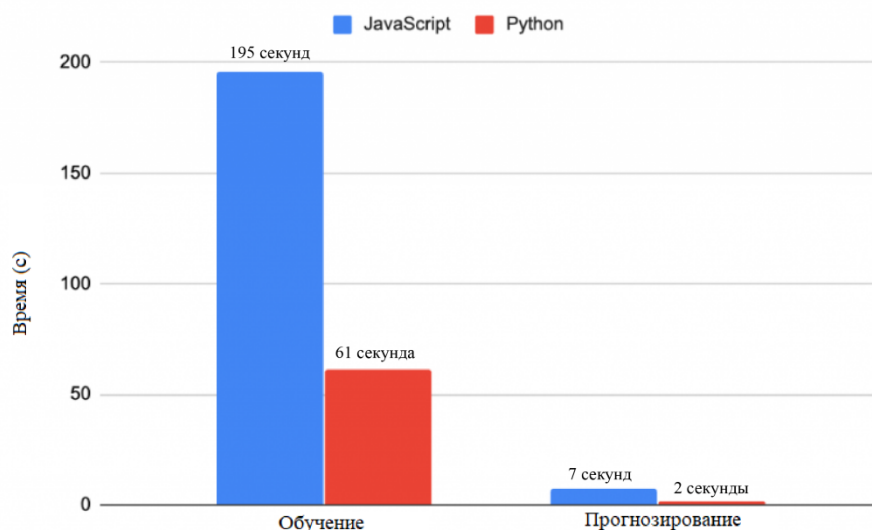


Рисунок 1. Время, затраченное на обучение и прогнозирование моделями на *JavaScript* и *Python*

2. **Малая коллекция библиотек.** Несмотря на то, что в *JavaScript* быстро растет число библиотек для машинного обучения, многие языки обладают гораздо более обширной коллекцией встроенных библиотек, которые прошли проверку временем и хорошо протестированы [3].

3. **Работа только с малым объемом данных.** Еще одним недостатком *JavaScript* является то, что существуют ограничения на объем данных, к которым он может обращаться, хранить и использовать. *JavaScript*-код не имеет доступа по умолчанию к файловой системе основного браузера, что также может привести к ограничениям в размерах файлов [3].

Говоря о преимуществах использования данного языка программирования в машинном обучении, необходимо выделить следующие показатели:

1. Популярность и доступность. *JavaScript* является одним из самых популярных и широко используемых языков программирования в мире, что повышает его доступность в сфере разработки приложений для машинного обучения. В то время как большинство других языков программирования по умолчанию не поддерживаются рядом устройств, почти каждый современный телефон и ноутбук поддерживает *JavaScript* без необходимости установки дополнительного программного обеспечения. Из-за его широкой популярности сообщество поддержки *JavaScript* также очень обширно и активно [3].

Поскольку многие люди уже используют *JavaScript*, и так как его легко освоить даже тем, у кого нет опыта, он может быть менее пугающим для новичков в машинном обучении, чем другие, менее знакомые языки. Машинное обучение для *JavaScript* также очень быстро и легко освоить, если вы уже являетесь разработчиком *JavaScript*. В конечном счете, команде разработчиков, уже использующей *JavaScript*, гораздо проще изучить алгоритмы машинного обучения на этом языке, чем изучать совершенно другой, что в долгосрочной перспективе может сэкономить время и ресурсы [3].

2. Приватность данных. Кибербезопасность является одной из главных забот предприятий и разработчиков программного обеспечения, поэтому компании проводят оценки безопасности разработки, чтобы снизить угрозу утечки данных, связанной с кодом. Как правило, куда более безопасно запускать машинное обучение на устройстве пользователя, а не на сервере [3].

Поскольку большинство приложений машинного обучения основаны на архитектуре клиент-сервер и передают информацию обратно на сервер, личные сообщения, изображения и электронные письма могут попасть в нежелательные руки на сервере, на котором работает модель машинного обучения. *JavaScript*, с другой стороны, не зависит от сервера для возможностей машинного обучения и может работать без отправки вашей информации третьей стороне [3].

3. Кастомизация машинного обучения. *JavaScript* также упрощает настройку моделей машинного обучения для отдельных пользователей по сравнению с приложениями, разработанными на других языках, что делает его идеальным для проектов, предполагающих широкое развертывание и множество разных пользователей. Доказано, что персонализированная целевая реклама способствует развитию бизнеса за счет повышения лояльности клиентов, увеличения продаж и предоставления вам данных о ваших клиентах, которые помогут вам узнать о вашем рынке и о том, что они ищут [3].

Для серверного приложения машинного обучения настройка будет означать хранение на сервере разных моделей машинного обучения для каждого пользователя, занимающих много места. С помощью *JavaScript* у вас есть возможность разместить копию модели машинного обучения на устройстве каждого пользователя. Отсюда вы можете использовать обучающие библиотеки для настройки модели на каждом уникальном устройстве. Это делает приложение более доступным для пользователей и снижает вероятность того, что вам придется покупать несколько дорогих серверов [3].

На рисунке 2 представлена схема распределения моделей на устройствах конечных пользователей.



Рисунок 2. Схема распределения моделей на устройствах конечных пользователей

4. Простая интеграция машинного обучения в веб- и мобильных приложениях. Для JavaScript уже существует богатый набор кроссплатформенных инструментов разработки мобильных приложений, таких как *Cordova* и *Ionic* [4].

Эти инструменты стали очень популярными, поскольку позволяют один раз написать код и развернуть его для устройств *iOS* и *Android*. Чтобы сделать код совместимым с различными операционными системами, средства кроссплатформенной разработки запускают «веб-представление» — объект браузера, который может запускать код *JavaScript* и может быть встроен в собственное приложение целевой операционной системы. Эти объекты браузера поддерживают библиотеки машинного обучения *JavaScript* [4].

Единственным исключением является *React Native*, популярная кроссплатформенная среда разработки мобильных приложений, которая не использует веб-просмотр для запуска приложений. Однако, учитывая популярность мобильных приложений машинного обучения, *Google* выпустила специальную версию *TensorFlow.js* для *React Native* [4].

В готовое мобильное приложение можно добавить встроенный объект браузера (например, *WKWebView* в *iOS*) для интеграции с кодом машинного обучения, а готовую версию приложения машинного обучения для браузера, можно легко перенести в мобильное приложение с небольшими изменениями или без них [4].

На сегодняшний день самым распространенным инструментом для разработки моделей машинного обучения на *JavaScript* является библиотека *TensorFlow.js*.

Библиотека для машинного обучения на JavaScript – TensorFlow.js

Данная библиотека с открытым исходным кодом, разработанная компанией *Google*, предназначена для разворачивания машинного обучения в среде браузера и является основным проводником в мир ИИ для *JavaScript*-разработчика. Она максимально использует вышеперечисленные достоинства языка *JavaScript* в сфере машинного обучения, что открывает уникальные и ранее недоступные возможности для разработчиков.

Библиотека позволяет создавать новые модели машинного обучения с нуля, также как управлять и модифицировать уже существующие. Модель, созданная с помощью *TensorFlow* на *Python*, может быть конвертирована для работы в браузере [5].

На рисунке 3 представлены примеры проектов с официального сайта библиотеки, созданных с помощью нее [6].

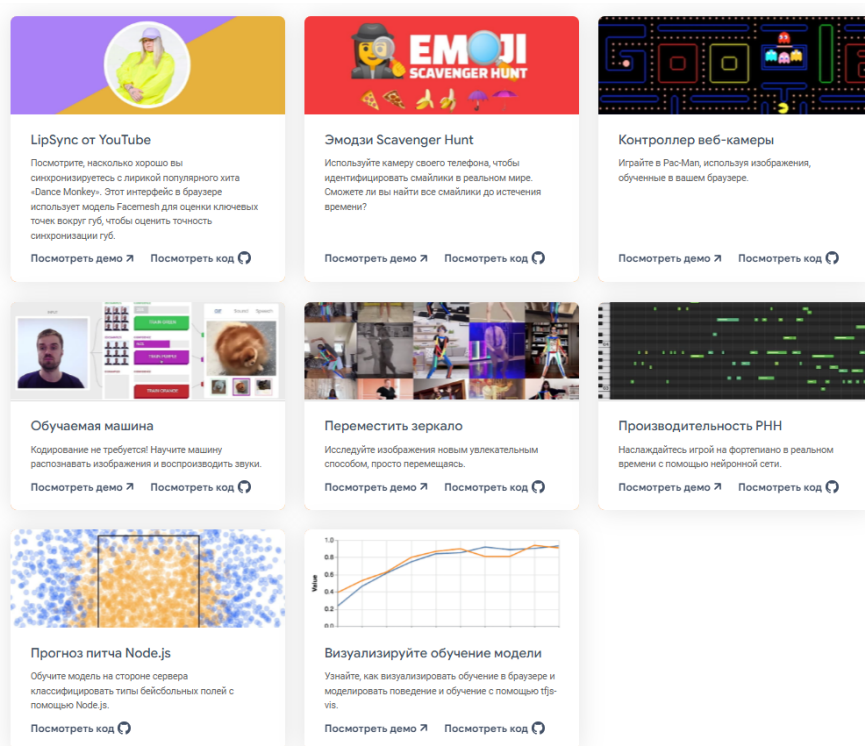


Рисунок 3. Примеры проектов, созданных с использованием библиотеки *TensorFlow.js*

Для обучения модели обычно как разработчику, так и конечному потребителю моделей машинного обучения, требуется установка соответствующих библиотек. Однако в *TensorFlow.js* интерфейсом является браузер, поэтому при подключении скриптового файла исчезает необходимость в установке библиотек и отслеживании зависимостей, все «доставляется на дом» [7].

Строение библиотеки можно представить в виде оболочечной структуры – ядерного *API* и покрывающих его высокоуровневых слоев [7]:

1. *Ops API* (сокращенно от *operations* – низкоуровневые операции). Синтаксически эта составляющая близка к классическому *TensorFlow* с *Python*-интерфейсом.

2. *Layers API* (высокоуровневые слои). Это *API* аналогично библиотеке *Keras*, сводящей действия в обучении к минималистичным наборам наиболее распространенных общих команд.

На основе высокоуровневого *API* можно построить собственный интерфейс для взаимодействий пользователя с моделью. Это открывает новые возможности для быстрого создания веб-приложений с захватом данных с сенсоров переносимых устройств (акселерометра, гироскопа, камеры, *GPS* и т.д.) и обучения моделей для подбора релевантного контента, исходя из поведения пользователя [7].

При этом автоматически обеспечивается безопасность, так как данные создаются и хранятся на стороне клиента. Кроме того, *TensorFlow.js* позволяет использовать предобученные модели, что крайне удобно для трудоемких задач, например, таких как распознавание образов [7].

В настоящее время разработчики все чаще и чаще отдают предпочтение фронтенд приложениям перед бэкенд-серверами в качестве сред для машинного обучения. Благодаря *TensorFlow.js*, управление моделями машинного обучения стало возможным в статических *HTML*-документах без настройки сервера или базы данных, позволяя этим сервисам существовать полностью на стороне клиента [5].

Ниже представлен листинг кода для построения модели, предсказывающей координаты точки по заданным параметрам [7].

```
<html>
  <head>
    <!-- Загрузка TensorFlow.js -->
    <script src="https://cdn.jsdelivr.net/npm/@tensorflow/tfjs@0.10.0">
</script>
    <script>
      // Определим модель линейной регрессии. Создадим простейшую нейронную сеть
из одного слоя.
      const model = tf.sequential();
      model.add(tf.layers.dense({units: 1, inputShape: [1]}));

      // Подготовим модель для обучения: обозначим функцию стоимости и
оптимизатор.
      model.compile({loss: 'meanSquaredError', optimizer: 'sgd'});
      // Создадим некоторые данные для обучения.
      const xs = tf.tensor2d([1, 2, 3, 4], [4, 1]);
      const ys = tf.tensor2d([1, 3, 5, 7], [4, 1]);
      // Обучаем модель по данным.
      model.fit(xs, ys).then(() => {
        // Используем модель для предсказания координаты точки.
        // Откройте этот файл в браузере и зайдите в консоль разработчика.
        model.predict(tf.tensor2d([5], [1, 1])).print();
      });
    </script>
  </head>
  <body>
</body>
</html>
```

Заключение.

На сегодняшний день, можно смело сказать, что хоть *JavaScript* и не претендует на лавры *Python* в качестве самого оптимального языка для машинного обучения, его актуальность в сфере разработки легковесных и нетребовательных к ресурсам веб-приложений растет в очень крупных масштабах. Его доступность и простота в использовании на фоне роста технологичности и производительности портативных устройств, а также тенденцией к разработке кроссплатформенных приложений очень выгодно выделяются на фоне аналогов. Такие библиотеки машинного обучения, как *TensorFlow.js*, могут заметно преобразить разработку и саму работу веб-приложений, наделяя их полезным и уникальным функционалом.

Список литературы

- [1] Why use JavaScript for machine learning? [Электронный ресурс]. URL: <https://hub.packtpub.com/should-you-use-javascript-for-machine-learning-and-how-do-you-get-started/> (Дата обращения: 28.03.2023).
- [2] Шолле, Франсуа JavaScript для глубокого обучения: TensorFlow.js / Франсуа Шолле, Эрик Нильсон, Стэн Байлесчи, Шэнкуинг Цэй – СПб.: Питер, 2021. – 576 с.
- [3] JavaScript for Machine Learning? Weighing The Pros and Cons [Электронный ресурс]. URL: <https://javascript.plainenglish.io/javascript-for-machine-learning-weighing-the-pros-and-cons-e7577d0e3f48> (Дата обращения: 28.03.2023).
- [4] Performance Comparison: JavaScript vs. Python for Machine Learning [Электронный ресурс]. URL: <https://dlabs.ai/blog/performance-comparison-javascript-vs-python-for-machine-learning/> (Дата обращения: 28.03.2023).
- [5] The Pros and Cons of Using JavaScript for Machine Learning [Электронный ресурс]. URL: <https://dlabs.ai/blog/the-pros-and-cons-of-using-javascript-for-machine-learning/> (Дата обращения: 28.03.2023).
- [6] TensorFlow.js — библиотека для машинного обучения на JavaScript. [Электронный ресурс]. URL: <https://www.tensorflow.org/js?hl=ru> (Дата обращения: 28.03.2023).

[7] TensorFlow.js: машинное обучение на JavaScript с доставкой в браузер [Электронный ресурс]. URL: <https://proglib.io/p/tensorflow-js> (Дата обращения: 28.03.2023).

RELEVANCE OF USING JAVASCRIPT LANGUAGE FOR DEVELOPING MACHINE LEARNING MODELS

G.A. Piskun

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

V.F. Alekseev

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

T.M. Voronko

Software Engineer of the Center of Informatization and Innovative Developments of BSUIR, master student of BSUIR

Faculty of Computer Engineering

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

E-mail: voronko232001@gmail.com

Abstract. The applicability and relevance of using the JavaScript programming language for developing machine learning models is analyzed. The main advantages and disadvantages of this method of working with artificial intelligence in comparison with analogues such as Python are considered, and the features of the TensorFlow.js JavaScript library are described, providing an example program for building a model that predicts the coordinates of a point according to given parameters.

As a result of the analysis, it was found that the use of JavaScript can be useful for lightweight and undemanding cross-platform web applications that use machine learning models.

Keywords: machine learning, library, artificial intelligence.

УДК 378.17

НЕЙРОСЕТЬ КАК ФАКТОР ОПТИМИЗАЦИИ ТАЙМ-МЕНЕДЖМЕНТА СПЕЦИАЛИСТА ПРИ РЕШЕНИИ ТЕКУЩИХ ПРОИЗВОДСТВЕННЫХ ЗАДАЧ



А.В. Санец
Магистрант
специальности «Охрана
труда и эргономика»,
a.sanets@bsuir.by



О.С. Медведев
Старший преподаватель кафедры
ИПиЭ, магистр техники и
технологии, аспирант
o.med@bsuir.by



Т.Ю. Шлыкova
Доцент кафедры ИпиЭ БГУИР,
кандидат психологических наук,
доцент
t.shlykova@bsuir.by

А.В. Санец

Окончил Белорусский государственный университет информатики и радиоэлектроники по специальности 1-58 01 01 «Инженерно-психологическое обеспечение информационных технологий». Область научных интересов связана с исследованием проблем тайм-менеджмента, изучение нейросетей и их применение в различных областях.

О.С. Медведев

В 2013 году окончил БГУИР, специальность «Инженерно-психологическое обеспечение информационных технологий». Квалификация инженер-системотехник. В 2019 году окончил магистратуру по специальности «Управление безопасностью производственных процессов» с присвоением академической степени магистра техники и технологии. Ведет курсы «CISCO».

Т. Ю. Шлыкova

В 1999 году окончила БГПУ им. Максима Танка, дефектология (логопедия). Окончила БГПУ им. Максима Танка, аспирантура по специальности 19.00.07. «Педагогическая психология». Доцент кафедры психологии БГПУ. Заведующая кафедрой гуманитарных дисциплин ЧУО «Институт предпринимательской деятельности». Заведующая кафедрой психологии и педагогики БГАТУ. Доцент кафедры инженерной психологии и педагогики БГУИР.

Аннотация. В данной работе рассматриваются основные характеристики нейронных сетей, которые определяют оптимизацию современного тайм-менеджмента. Проведен анализ нейронной сети в условиях решения текущих задач современного специалиста с целью управления временем и увеличения производительности. Излагается описание различных моделей нейросетей и их применение в решении различных задач, а также методы использования нейросетей для анализа данных о работе и распределении времени. Представлен тайм-менеджмент как техника организации контроля и распределения времени, которая позволяет вовремя достигать поставленных целей и задач, повышать эффективность и результативность.

Ключевые слова: Нейронные сети, тайм-менеджмент, машинное обучение, производительность, эффективность.

Введение.

На сегодняшний день нейронные сети все больше внедряются в профессиональную жизнь не только специалистов в сегменте информационных технологий, но специалистов профессиональных областей, где требуется следование временному регламенту при решении текущих производственных задач— военнослужащие, педагоги, спортсмены, руководители и все, без исключения, работники режимных предприятий.

Нейронные сети как информационная технология имеет неисчерпаемый ресурс управления временем и может выступать в качестве фактора оптимизации тайм-менеджмента современного специалиста в ходе решения текущих производственных задач. Выработка производственного режима времени и его наполнение алгоритмом работы—важнейший тактический прием

современного управления, автоматизация которого обеспечит эффективность решения текущих задач.

Нейронные сети являются мощным инструментом машинного обучения, который позволяет компьютерам обучаться и выполнять задачи, которые традиционно требовали человеческого интеллекта. Сегодня нейронные сети могут применяться практически для любой задачи.

Области применения нейронных сетей:

1. В области медицинских исследований и помощи в обработке данных полученных в ходе обследования пациента. Известны примеры применения при диагностике рака, лицевой хирургии, обработке данных полученных с помощью компьютерной томографии.

2. В области организации маркетинга. Прогнозирование продаж, а соответственно сокращение риска связанного с перепроизводством продукции, издержек, связанных с поставками и хранением товаров. Среди наиболее полезных выделяют ChatGPT, DALL-E 2 и Anyword [1].

3. В области обеспечения безопасности и противодействия преступности. Идентификации лиц в системах безопасности, фиксация различных опасных объектов в поле зрения видеофиксаторов. Среди крупных нейросетей в этой области выделяют VisionLabs и NTechLab [2].

4. В области лингвистики. Усовершенствование машинного перевода, анализ текстовых данных, попытки более точной передачи смысловой нагрузки. Самым известным среди переводчиков работающим на нейронных сетях является DeepL [3]

5. В области мультимедийной обработки данных. Создание изображений на основе текстовых запросов, редактирование уже существующих изображений. Автоматизированное создание и обработка музыкальных произведений. Среди известных и крупных проектов выделяют такие нейросети как Midjourney и Stable Diffusion [4,5].

Существуют различные архитектуры нейросетей:

- нейросети прямого распространения;
- сверточные нейросети;
- рекуррентные нейросети.

Нейросети прямого распространения. Модель этого искусственного интеллекта, используют прямые связи между слоями нейронных ячеек для предсказания выходных данных. В них используется один или несколько скрытых слоев для преобразования входных данных в выходные данные. Это позволяет нейросети прямого распространения подстраиваться под обновления данных и использования их для предсказания будущих результатов [6].

Сверточные нейросети. Данные модели нейронных сети используют сверточные слои для выявления пространственных зависимостей в данных, таких как изображения. Они подходят для классификации и сегментации изображений [6].

Рекуррентные нейросети. В рекуррентной модели нейросети используют рекуррентные соединения для обработки последовательных данных, таких как текст и видео. Они хорошо подходят для задач обработки естественного языка и прогнозирования серий. Представителями таких нейросетей являются LSTM и GRU [6]

На данный момент существуют два примера успешных нейросетей:

- Chatbot GPT от OpenAI
- Midjourney - генеративная модель, созданная Anthropic, PBC.

Chatbot GPT – это 175-миллиардный модельный язык, который может генерировать текст, отвечать на вопросы и даже писать код [7].

Midjourney – это другая популярная система на базе искусственного интеллекта, которая создает изображения из подсказок пользователя [4].

Техническое развитие нейронных сетей способствует их использованию при создании сервисов, связанных с планированием и организацией последовательности задач, которые должен выполнить современный специалист.

Нейросети в структуре тайм-менеджмента.

Применение нейросетей в решении задач управления временем специалистов открывает множество возможностей, связанных с организацией осознанного контроля и распределения времени, помощи с эффективным распределением времени для достижения поставленных целей, позволит повысить результативность и скорость выполнения работы, снизить стресс от сомнений, связанных с выбором последовательности выполнения задач.

Способы внедрения нейросетей в системы обеспечения тайм менеджмента:

1) Одним из способов является использование нейросетей для анализа данных о вашей работе и распределения времени. Например, использование нейросети для анализа рабочего дня и определения, когда специалист наиболее продуктивен и эффективен. Эта информация может помочь в создании рекомендаций, позволяющих распределить свое время более эффективно, а также оптимизировать время работы (создание рабочего графика для работы в наиболее продуктивные часы).

2) Нейросети могут использоваться для прогнозирования времени, необходимого для выполнения задач. Можно использовать нейросеть для анализа прошлых проектов и определения, сколько времени потрачено на выполнение задач, и последующего использования этой информации для прогнозирования времени, необходимого для выполнения будущих задач.

3) Также возможен обратный процесс: использование тайм-менеджмента для управления временем, которое вы тратите на обучение нейросетей. Проблематика этого направления тайм-менеджмента заключается в том, чтобы ограничивать объемы обрабатываемых данных нейронных сетей во избежание затягивания.

4) Менеджмент обучения. Можно использовать нейросети в тайм-менеджменте на этапе планирования и контроля времени, которое необходимо на изучение новых технологий и определения, какие знания более эффективно получить.

Заключение.

На современном этапе развития управленческой деятельности внедрение нейросети в тайм-менеджмент является уместным и целесообразным. Возможности, которые предоставляет нейросеть могут существенно улучшить работу специалистов за счет оптимизации управления временем при решении текущих производственных задач. Повышение показателей эффективности и продуктивности работы является убедительным аргументом внедрения нейросети в деятельность современных специалистов.

Актуальность изложенных вопросов не исчерпана. Дальнейшее изучение проблемы внедрения нейросети в деятельность специалистов связано с операционализацией задач нейросети с учетом потребностей, особенностей личности специалистов, а также специфики текущих производственных задач. Эти и другие важные вопросы заявленной проблемы будут изложены нами в последующих публикациях.

Список литературы

[8]. Нейросети в маркетинге: как применять и какие есть [Электронный ресурс]. – Режим доступа: https://blog.click.ru/growthhacking/nejroseti-v-marketinge/#h2_3 (Дата обращения: 21.03.2022).

[9]. Как работает распознавание лиц и можно ли обмануть эту систему [Электронный ресурс]. – Режим доступа: <https://trends.rbc.ru/trends/industry/6050ac809a794712e5ef39b7> (Дата обращения: 21.03.2022).

[10]. DeepL Translate [Электронный ресурс]. – Режим доступа: <https://www.deepl.com/translator> (Дата обращения: 21.03.2022).

[11]. Как пользоваться Midjourney: команды, настройки и примеры [Электронный ресурс]. – Режим доступа: <https://trends.rbc.ru/trends/industry/63e0d08f9a7947342f156b9b> (Дата обращения: 21.03.2022).

[12]. Stable Diffusion. Самая впечатляющая нейросеть. [Электронный ресурс]. – Режим доступа: <https://vc.ru/future/506283-stable-diffusion-samaya-vpechatlyayushchaya-neyroset-kak-polzovatsya-novym-instrumentom-ii> (Дата обращения: 21.03.2022).

[13]. Введение в архитектуры нейронных сетей [Электронный ресурс]. – Режим доступа: <https://habr.com/ru/companies/oleg-bunin/articles/340184/> (Дата обращения: 21.03.2022).

[14]. GPT-3 [Электронный ресурс]. – Режим доступа: <https://openai.com/research/gpt-3> (Дата обращения: 21.03.2022).

NEURAL NETWORK AS A FACTOR OF OPTIMIZING TIME MANAGEMENT OF A SPECIALIST WHEN SOLVING CURRENT PROBLEMS PRODUCTION TASKS

A.V. Sanets

*Master student of the education
«Labor protection and
ergonomics»*

O.S. Medvedev

*Senior Lecturer of the
Department of EPE,
Master of Engineering and
Technology*

T.Y. Shlykova

*Candidate of Psychological
Sciences, Associate
Professor of the department
EPE BSUIR*

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

E-mail: a.sanets@bsuir.by, o.med@bsuir.by, t.shlykova@bsuir.by

Annotation. This paper discusses the main characteristics of neural networks that determine the optimization of modern time management. The neural network was analyzed in terms of solving the current problems of a modern specialist in order to manage time and increase productivity. A description of various models of neural networks and their application in solving various problems, as well as methods for using neural networks to analyze data on work and time distribution, is presented. Time management is presented as a technique for organizing control and distribution of time, which allows you to achieve your goals and objectives on time, increase efficiency and effectiveness.

Keywords: neural networks, time management, machine learning, productivity, efficiency.

УДК 159.9.07:629.41:331.44

МЕТОДИКА ВЫРАБОТКИ НАВЫКА РЕЛАКСАЦИИ ДЛЯ РАБОТНИКОВ ЛОКОМОТИВНЫХ БРИГАД



Н.В. Щербина

*Старший преподаватель кафедры инженерной психологии и эргономики
УО «Белорусский государственный университет информатики и
радиоэлектроники», магистр технических наук
shcherbina@bsuir.by*

Н.В. Щербина

Старший преподаватель кафедры инженерной психологии и эргономики УО «Белорусский государственный университет информатики и радиоэлектроники», магистр технических наук, «исследователь» в области технических наук. Проводит научные исследования в областях промышленной безопасности, эргономики, психологии и безопасности труда.

Аннотация: Предложена методика выработки навыка релаксации по психофизиологическим, личностным показателям машинистов, помощников машинистов локомотивных бригад Белорусской железной дороги с учетом их психофизиологических, личностных показателей. Для реализации данной методики используются аппаратно-программные комплексы NeuroDog и УПДК-МК (универсальный психодиагностический комплекс) АО «Нейроком».

Ключевые слова: биологическая обратная связь, саморегуляция, выработка навыка релаксации, функциональное состояние, электродермальность активности кожи, локомотивная бригада, психофизиологические и личностные показатели.

Введение.

С каждым годом растет интерес к обучению работников управлению своим психоэмоциональным состоянием, поскольку в системе «человек–машина–среда» нередко именно ошибка оператора приводит к нарушению работоспособности системы или ее элементов (техники и человека), или отказу системы, или человеческим жертвам.

Воздействие на человека-оператора психоэмоционального напряжения в сочетании с достаточно монотонными условиями труда способствует возникновению психосоматических (гипертония, язвенная болезнь и др.), сердечно-сосудистых заболеваний и психических расстройств. Ежегодный отпуск и возможность оздоровления в санаторно-курортных учреждениях чаще всего помогают восстановлению организма, поскольку, как отмечает А.Н. Долецкий: «...релаксация является наиболее естественной формой саморегуляции» [1]. Однако, качественное расслабление, способствующее восстановлению организма человека, не должно ограничиваться условиями стационара и полным отсутствием рабочих нагрузок. Правильно используемый межсменный (межрейсовый) отдых также способен повысить качество жизни и увеличить трудовое долголетие человека-оператора (машиниста, помощника машиниста локомотивной бригады).

Информационный поиск [2, 3] выявил, что существует достаточное количество различных систем, позволяющих обучать человека управлению своим психоэмоциональным состоянием, большинство из которых используют принципы биологической обратной связи (БОС). Обучение навыку саморегуляции осуществляется с использованием соответствующих аппаратно-

программных систем (большинство из которых используют БОС) и часто под контролем профильного специалиста, но в последующем испытуемый может использовать выработанный навык трансформации своего функционального состояния произвольно, в нужный для него момент времени, без использования аппаратно-программных средств. Методики исследований физиологических показателей [4], применяемых для организации биологической обратной связи являются неинвазивными, немедикаментозными, не имеют строгих противопоказаний и допускаются к применению не только в лечебно-профилактических учреждениях, но и в условиях производства под наблюдением специалиста или самостоятельно.

Материалы и методы.

В диссертационном исследовании для развития навыка релаксации у машинистов, помощников машинистов локомотивных бригад Белорусской железной дороги выбран метод БОС с использованием аппаратно-программного комплекса (АПК) NeuroDog (АО «Нейроком»). В исследовании [5-7] приняли участие 106 машинистов локомотивных бригад Моторвагонного депо г. Минска в возрасте от 18 до 65 лет, мужчины, средний возраст по выборке составил $32 \pm 11,4$ года. Стаж работы участников эксперимента в диапазоне от 1 до 36 полных лет.

Результаты исследования показали, что выработка навыка релаксации зависит от психофизиологических и личностных показателей машинистов, помощников машинистов локомотивных бригад. А именно, чем больше времени уделять на тренировки по выработке навыка релаксации, тем дольше работник локомотивной бригады способен находится в состоянии релаксации, но при этом увеличивается время от начала сеанса до «погружения» в состояние релаксации. Чем старше и опытнее машинист, тем больше времени ему необходимо для перехода из состояния оперативного покоя в состояние покоя. Чем выше степень ситуативной тревожности, степень неприспособленности к условиям ситуации, которая не позволяет испытуемому выполнять требования и достигать желаемой цели ($r_s=0,23$); чем выше степень неуверенности в себе, тревоги, чувства вины, суетливости ($r_s=0,21$); чем больше количество допущенных ошибок в простой двигательной реакции, т.е. при оценке способности машиниста быстро реагировать на появление значимого сигнала ($r_s=0,21$); чем выше уровень стремления личности к установлению дружелюбных отношений и сотрудничеству с окружающими ($r_s=0,20$); чем меньше степень ответственности, уравновешенности, стойкости и упорства ($r_s=-0,29$); чем меньше степень тревожности, усталости, эмоционального возбуждения ($r_s=-0,26$); чем меньше степень общительности, смелости, чувствительности, импульсивности ($r_s=-0,24$); чем меньше допущено ошибок при выполнении быстрых и точных действиях вменяющейся обстановке ($r_s=-0,21$); чем меньше степень подозрительности, замкнутости, скрытности ($r_s=0,20$), тем больше времени необходимо испытуемому для перехода из состояния оперативного покоя в состояние релаксации. Чем больше времени испытуемый тратит на сеансы ($r_s=0,64$); чем дольше он находится в состоянии релаксации ($r_s=0,63$) и ($r_s=0,62$); чем меньше степень тревожности, эмоционального возбуждения, усталости ($r_s=-0,25$); чем меньше степень самоконтроля и целенаправленности ($r_s=-0,21$); чем меньше степень эмоциональной стабильности: выдержанности, спокойствия, флегматичности ($r_s=-0,20$); чем меньше степень гибкости мышления, быстроты принятия решения, склонности к экспериментам ($r_s=-0,19$); чем меньше уровень субъективного контроля над эмоционально положительными событиями и ситуациями ($r_s=-0,19$), тем больше времени необходимо испытуемому для перехода из состояния спокойного бодрствования в состояние релаксации.

По результатам БОС-тренинга на выработку навыка релаксации испытуемых распределили в три группы: «успешная», «менее успешная», «неуспешная» [5, 6]. Составлены личностные профили для каждой группы машинистов, помощников машинистов локомотивных бригад по методике Т. Лири (рисунок 1), по методике Р.Б. Кеттелла (рисунок 2), по методике ММРІ (рисунок 3).

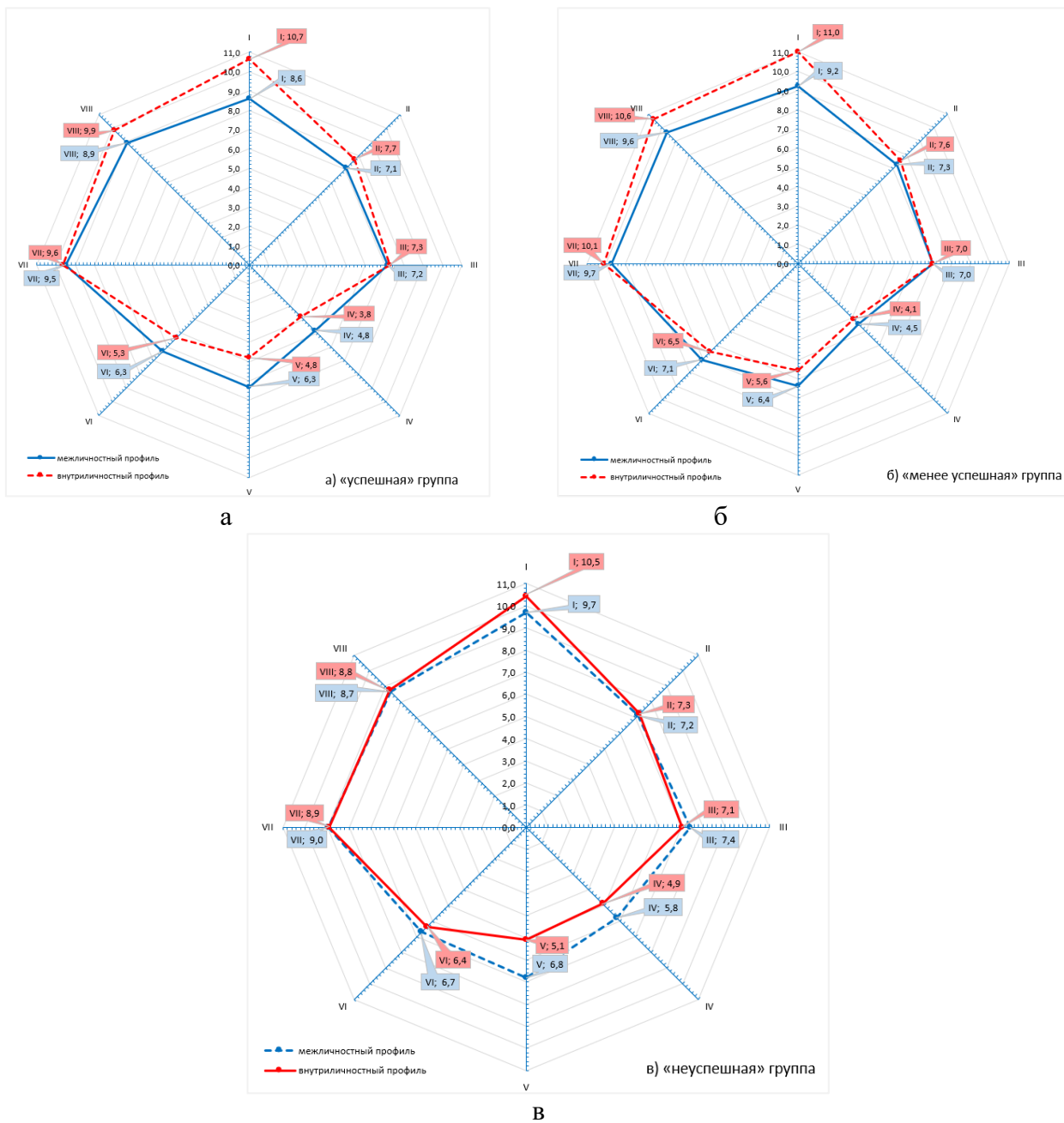
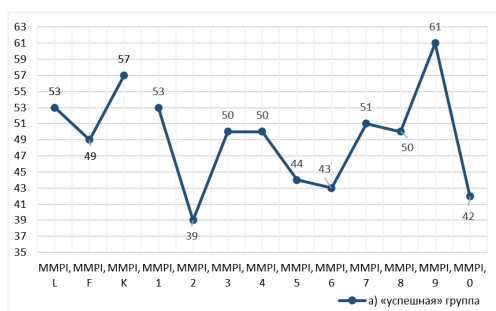


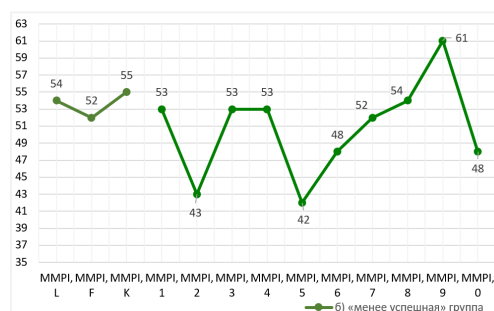
Рисунок 1. Личностный профиль групп машинистов и помощников машинистов локомотивных бригад по методике Т.Лири: а) «успешная»; б) «менее успешная»; в) «неуспешная»

		«успешная» группа					«менее успешная» группа					«неуспешная» группа											
Отрицательный полюс		1	2	3	4	5	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5	Положительный полюс	
Эмоциональные качества / Эмоциональный интеллект																							
фактор С	Эмоциональная неустойчивость																					Эмоциональная устойчивость	
фактор L	Подозрительность																					Доверчивость	
фактор O	Тревожность																					Спокойствие	
фактор Q4	Напряженность																					Расслабленность	
Волевые качества																							
фактор E	Подчиненность																					Властность	
фактор G	Робость																					Смелость	
фактор H	Небрежность																					Ответственность	
фактор Q3	Низкий самоконтроль																					Высокий самоконтроль	
Коммуникативные качества																							
фактор A	Отчужденность																					Общительность	
фактор F	Пессимизм																					Оптимизм	
фактор I	Жесткость																					Мягкосердие	
фактор Q2	Зависимость от группы																					Самостоятельность	
Интеллектуальные качества																							
фактор B	Низкий интеллект																					Высокий интеллект	
фактор M	Практичность																					Мечтательность	
фактор N	Прямолинейность																					Дипломатичность	
фактор Q1	Консерватизм																					Гибкость	
Вторичные интегральные факторы																							
фактор F1	Низкая тревожность																					Высокая тревожность	
фактор F2	Интроверсия																					Экстраверсия	
фактор F3	Чувствительность																					Уравновешенность	
фактор F4	Конформность																					Независимость	

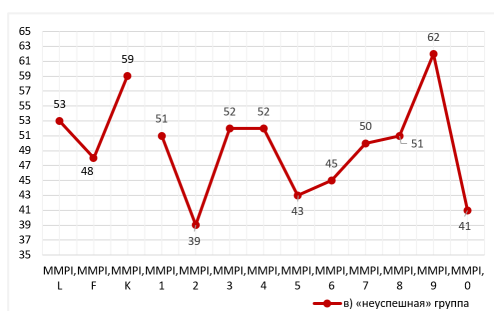
Рисунок 2. Личностный профиль групп машинистов и помощников машинистов локомотивных бригад по методике Р.Б.Кеттелла: а) «успешная»; б) «менее успешная»; в) «неуспешная»



а



б



в

Рисунок 3. Личностный профиль групп машинистов и помощников машинистов локомотивных бригад по методике MMPI: а) «успешная»; б) «менее успешная»; в) «неуспешная»

Сравнительный анализ групп [8] показал различия по уровню бдительности, скорости переключения внимания, эмоциональной устойчивости, объему внимания. Расчеты Н-критерия Краскела-Уоллиса показали статистически значимые различия в указанных группах по

следующим показателям: время выполнения задания на эмоциональную устойчивость ($H(2, 106) = 5,88; p = 0,0530$); количество ошибок, допущенных при выполнении задания на концентрацию внимания ($H(2, 106) = 10,38; p = 0,0056$); теплинг/2 ($H(2, 106) = 5,76; p = 0,0560$); внутриличностный октант 6 (Лири) ($H(2, 106) = 6,42; p = 0,0403$); внутриличностный октант 8 (Лири) ($H(2, 106) = 5,77; p = 0,0558$); доминирование в межличностных отношениях (Лири) ($H(2, 106) = 5,65; p = 0,0592$); фактор G (низкая-высокая нормативность поведения) ($H(2, 106) = 6,81; p = 0,0333$) и интернальность в области достижений ($H(2, 106) = 5,71; p = 0,0576$) [9]. Получены математические модели классификации машинистов, помощников машинистов локомотивных бригад в группы «успешная», «менее успешная», «неуспешная» по их психофизиологическим, личностным показателям [10]; модели прогнозирования «показателей» релаксации (КГР-150, КГРср-60; суммы КГР-150, КГРср-60, общее время сеанса) для машинистов, помощников машинистов локомотивных бригад по их психофизиологическим, личностным показателям [11]. Предложена информационная система для оценки и прогнозирования выработки навыка релаксации на основе БОС у машинистов, помощников машинистов локомотивных бригад по их психофизиологическим, личностным показателям [12].

Результаты.

Учитывая полученные результаты исследования, а также предложенную в научной работе [13] обобщенную методику выработки навыка на саморегуляцию функционального состояния с использованием метода биологической обратной связи по психофизиологическим или физиологическим показателям человека, автором предложена скорректированная методика выработки навыка релаксации с использованием аппаратно-программных комплексов NeuroDog и УПДК-МК (универсальный психодиагностический комплекс) (АО «Нейроком») для работников локомотивных бригад с учетом их психофизиологических и личностных показателей.

Этапы методики:

Этап 1. Психолог вносит данные периодических осмотров (данные АПК УПДК-МК) в информационную систему [12] «Профиль бригады машинистов локомотивного депо». Программа рассчитывает прогнозируемые показатели релаксации (КГР-150, КГРср-60; суммы КГР-150, КГРср-60, общее время сеанса [11]) для конкретного машиниста, помощника машиниста локомотивной бригады. Психолог планирует индивидуальную программу по выработке навыка релаксации для каждого работника, определяя необходимое количество БОС-сеансов.

Этап 2. Психолог объясняет машинисту назначение электродов (датчиков) и в общем виде сущность проводимого БОС-сеанса. В течение 5-10 мин машинист привыкает к условиям обстановки (при последующих БОС-сеансах длительность *этапа 2* сводится к минимуму).

Этап 3. Машинист с помощью психолога определяет базовый уровень управляемого показателя. На основании измерений (регистрации) кожно-гальванической реакции в этом периоде машинист определяет критерий управления своим психофизиологическим состоянием.

Этап 4. Период покоя. Психолог знакомит машиниста с инструкцией на выработку навыка релаксации, в которой также содержится мотивирующее сообщение, побуждающее машиниста стараться выполнить задание (уложить «щенка» спать).

Инструкция для машиниста. Вам предлагается выполнить БОС-сеанс с использованием АПК NeuroDog по параметрам электродермальной активности кожи. Длительность сеанса 15 мин. Вам необходимо надеть кольцо, браслет-датчик, измеряющие ЭДА как показатель спокойствия (условный уровень бодрствования). Данные от датчика передаются в программу, где обрабатываются и представляются в виде активности виртуального щенка, соответствующей вашей собственной активности. (Машинисту предъявляют изображения двух состояний виртуального щенка: состояние активного бодрствования и состояние полной релаксации (состояние спокойного бодрствования)). Попробуйте расслабиться, наблюдайте, что происходит с изображением на экране монитора. Вовремя БОС-сеанса вам понадобится

максимально расслабиться, чтобы на экране монитора отобразилось изображение спящего щенка (состояние спокойного бодрствования). Запомните свое внутреннее состояние в те моменты, когда «щенок» укладывался и засыпал. Вам предлагается научиться воспроизводить это состояние и стараться удерживать его в течение всего БОС-сеанса. Для инициирования состояния сосредоточения вам предлагается сфокусировать внимание на какой-либо точке вблизи монитора, но так чтобы при этом видеть сигнал обратной связи (состояние виртуального щенка). Вы самостоятельно выбираете предпочтительную для вас стратегию саморегуляции (например, мышечное расслабление, изменение паттерна дыхания и др.).

Этап 5. Период саморегуляции. Машинисту предъявляется управляющее изображение («динамика поведения щенка»), состоящее из текущего значения психофизиологического параметра («щенок» в состоянии активного бодрствования) и критерия управления («щенок» в состоянии полной релаксации (состоянии спокойного бодрствования)), полученного на *этапе 3*.

Этап 6. Психолог контролирует адекватность выработанного навыка на релаксацию у машиниста, помощника машиниста локомотивной бригады по методу БОС с целью исключения «плацебо» эффекта.

Заключение.

Автором предложена методика выработки навыка релаксации по психофизиологическим, личностным показателям машинистов, помощников машинистов локомотивных бригад Белорусской железной дороги с учетом их психофизиологических, личностных показателей. Для реализации данной методики используется АПК NeuroDog (АО «Нейроком») в БОС-тренинге для выработки навыка релаксации и АПК УПДК-МК (АО «Нейроком»). Данные, полученные с использованием последнего, используются в информационной системе «Профиль бригады машинистов локомотивного депо» для расчета прогнозируемых показателей релаксации и прогнозирования группы по выработке навыка релаксации, которые в дальнейшем использует психолог для разработки индивидуальной программы выработки навыка релаксации у машиниста, помощника машиниста локомотивной бригады.

Список литературы

- [1] Долецкий А.Н. Нейрофизиологические механизмы адаптивного биоуправления и пути повышения его эффективности: автореферат дис. на соиск. учен. степ. докт. мед. наук: 03.03.01/ Долецкий Алексей Николаевич; ГБОУ ВПО «Волгоградский государственный медицинский университет» Министерства здравоохранения РФ. – Волгоград, 2013. – 44 с.
- [2] Гедранович Ю. А. Обзор и сравнительный анализ методов и систем для развития навыков релаксации. / Ю. А. Гедранович, В. В. Савченко, К. Д. Яшин, Н. В. Щербина // Журнал «Человеческий фактор: проблемы психологии и эргономики», 2016, № 1 (77), С. 62 – 69.
- [3] Гедранович Ю. А. Обзор и сравнительный анализ методов и систем для развития навыков релаксации. / Ю.А. Гедранович, В.В. Савченко, К. Д. Яшин, Н.В. Щербина // Журнал «Человеческий фактор: проблемы психологии и эргономики», 2016, № 2 (78), С. 44 – 50.
- [4] Щербина, Н. В. Обзор методов исследования физиологических показателей, используемых в системах с биологической обратной связью / Н. В. Щербина // Эргодизайн. – 2023. – № 1(19). – С. 81-89. – DOI 10.30987/2658-4026-2023-1-81-89.
- [5] Щербина, Н. В. Исследование метода выработки навыка на релаксацию с биологической обратной связью по параметрам электродермальной активности / Н.В. Щербина, В.В. Савченко, К.Д. Яшин // Новости медико-биологических наук. – 2019. – № 1/2019. – Том 19. – С. 65-73.
- [6] Щербина, Н.В. Первичная обработка и анализ данных для оценки состояния релаксации машинистов железнодорожного транспорта / Н. В. Щербина, В. В. Савченко, К. Д. Яшин // BIG DATA and Advanced Analytics = BIG DATA и анализ высокого уровня : сборник материалов V Международной научно-практической конференции, Минск, 13–14 марта 2019 г. В 2 ч. Ч. 2 / Белорусский государственный университет информатики и радиоэлектроники; редкол. : В. А. Богуш [и др.]. – Минск, 2019. – С. 121 – 125.
- [7] Щербина, Н. В. Оценка состояния релаксации машинистов железнодорожного транспорта / Н. В. Щербина // Электронные системы и технологии: сборник тезисов докладов 55 юбилейной научной конференции аспирантов, магистрантов и студентов (Минск, 22–26 апреля 2019 г.) / отв. ред. Радненок А. Л. – Минск, БГУИР, 2019. – С. 607-608.
- [8] Щербина, Н. В. Анализ различий между группами машинистов локомотивных бригад по

профессионально важным психофизиологическим и личностным показателям / Н. В. Щербина // BIG DATA and Advanced Analytics = BIG DATA и анализ высокого уровня : сборник научных статей VIII Международной научно-практической конференции, Минск, 11-12 мая 2022 года / Белорусский государственный университет информатики и радиоэлектроники ; редкол.: В. А. Богуш [и др.]. – Минск, 2022. – С. 265–272.

[9] Щербина Н.В. Анализ различий между группами машинистов локомотивных бригад с разной степенью способности к произвольной регуляции их функционального состояния. / Н.В. Щербина // Доклады БГУИР. 2022;20(2):21-29. <https://dx.doi.org/10.35596/1729-7648-2022-20-2-21-29>.

[10] Щербина Н.В. Дифференциальная диагностика способности к выработке навыка релаксации у машинистов локомотивных бригад. / Н.В. Щербина // Доклады БГУИР. 2022;20(4):96-103. <https://dx.doi.org/10.35596/1729-7648-2022-20-4-96-103>.

[11] Щербина Н.В. Прогнозирование способности к выработке навыка релаксации у машинистов локомотивных бригад. / Н.В. Щербина // Доклады БГУИР. 2022;20(7):95-101. <https://dx.doi.org/10.35596/1729-7648-2022-20-7-95-101>.

[12] Щербина, Н.В. Визуализация психофизиологических и личностных данных машинистов локомотивных бригад в Tableau / Н.В. Щербина // BIG DATA and Advanced Analytics = BIG DATA и анализ высокого уровня : сборник материалов VIII Международной научно-практической конференции, Минск, 11-12 мая 2022 г. В 2 ч. Ч. 2 / Белорусский государственный университет информатики и радиоэлектроники; редкол. : В. А. Богуш [и др.]. – Минск, 2022. – С. 215 – 218.

[13] Бондарь А.И., Савченко В.В., Планида Е.В., Филипович Л.В., Микуло Е.В. Методика выработки навыка на адаптивное преобразование функционального состояния спортсменов // Научные труды НИИ физической культуры и спорта Респ. Беларусь: сб. науч. тр. – Минск: Издательский центр БГУ, 2008. – Вып.8. – С.54-59.

RELAXATION SKILL DEVELOPMENT METHOD FOR WORKERS OF LOCOMOTIVE BRIGADES

N.V. Shcherbina

Master of Technical Science

Senior Lecturer, Department of Engineering Psychology and Ergonomics

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

E-mail: shcherbina@bsuir.by

Abstract. A technique for developing a relaxation skill based on psychophysiological, personal indicators of drivers, assistant drivers of locomotive crews of the Belarusian Railway is proposed, taking into account their psychophysiological, personal indicators. To implement this technique, the hardware-software complexes NeuroDog and UPDC-MK (universal psychodiagnostic complex) of Neurocom JSC are used.

Keywords: biofeedback, self-regulation, relaxation skill development, functional state, skin electrodermal activity, locomotive crew, psychophysiological and personal indicators.

УДК 339.138:004.9

ВНЕДРЕНИЕ CRM И ОПТИМИЗАЦИЯ БИЗНЕС-ПРОЦЕССОВ В СФЕРЕ ТУРИЗМА



Д.А. Фролова

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



П. Д. Булахова

Студентка инженерно-экономического
факультета БГУИР
polinabulahov@gmail.com

Д.А. Фролова

Окончила БГУИР. Область научных интересов связана с влиянием маркетинговых возможностей на формирование и развитие конкурентного бизнеса.

Ю.А. Янукович

Обучается в БГУИР. Область научных интересов связана с современными инструментами интернет-маркетинга.

Аннотация. В данной статье рассмотрено использование CRM-систем для оптимизации бизнес-процессов. Проанализирована динамика использования CRM-систем белорусскими компаниями. Приведены основные этапы внедрения CRM-систем в турфирму.

Ключевые слова: CRM-систем, оптимизация бизнес-процессов, планирование работ.

Введение.

Настоящее положение сферы туризма в Беларуси говорит о том, что данная сфера не играет важную роль в экономике страны. Хотя могла бы как это происходит в других странах, например, Франции или Италия, и другие. У нашей страны большой потенциал: выгодное географическое положение, наличие культурного наследия, однако страна занимает одно из последних мест в рейтинге мирового туристского рынка.

В последние несколько лет в Беларуси сократилось число турагентств. В условиях конкуренции на рынке и падения спроса на услуги туристических агентств, агентствам следует обратить внимание на необходимость увеличения конверсии и продажи для того, чтобы удержать уровень своей прибыли и укрепиться на хороших позициях на рынке. Оптимизация бизнес-процессов, увеличение прибыли, а также сокращение затрат – это то, что позволяет развиваться компании в любой ситуации.

Повышение продуктивности работы и развития бизнеса в сфере туризма (и не только туризма) связано с совершенствованием и оптимизацией бизнес-процессов. Одним из инструментов оптимизации бизнес-процессов является внедрение CRM-систем. CRM-система – программа для автоматизации и контроля взаимодействия компании с клиентами. Она хранит и структурирует информацию о заказах, покупателях, помогает оптимизировать маркетинг, повысить продажи и уровень обслуживания. Т.е. они позволяют видеть весь бизнес в ключе конкретных процессов и цифр. Используя данный инструмент можно легко измерять и сравнивать эффективность работы сотрудников, проводить аналитику и аудит компании, находить слабые места и планировать дальнейшее развитие.

Основная часть.

На данный момент организации в Беларуси только начинают внедрять CRM-системы. По данным belretail.by около 70% компаний в Беларуси используют цифровые технологии

для автоматизации бизнеса. И только около 16% компаний пользуются CRM-системами. Рынок CRM стремительно развивается.

Большая часть компаний Беларуси внедряют CRM-системы для использования следующих ее функций:

- планирование работы;
- ведение коммуникаций с клиентами;
- получение информации и аналитика;
- автоматический обмен документами;
- автоматизация различных отчетов.

Фирмы, что используют CRM-систему, достаточно быстро увеличивают свои продажи. Объем продаж зависит от количества лидов, конверсии и т.п. Для повышения продаж, этим нужно управлять, не стоит это бросать на самотек. CRM – это контроль лидов, проработка сделки, управление продажами, работа с постоянными клиентами и продажи для них. По статистике offlinescrm.ru увеличивают благодаря использованию CRM-систем компании увеличивают свои продажи до 29 %. Сотрудники отдела продаж и остальной персонал повышает свою продуктивность на 30 % после того, как получают доступ к CRM-системе из своих мобильных устройств.

За последние 3 года ситуация изменилась: CRM-системы стало использовать на 6% компаний больше, доля использования учетных программ для контроля бизнеса снизилась с 55% до 45%, а также увеличилось использование самописного софта для контроля за бизнес-процессами с 6% до 9%.

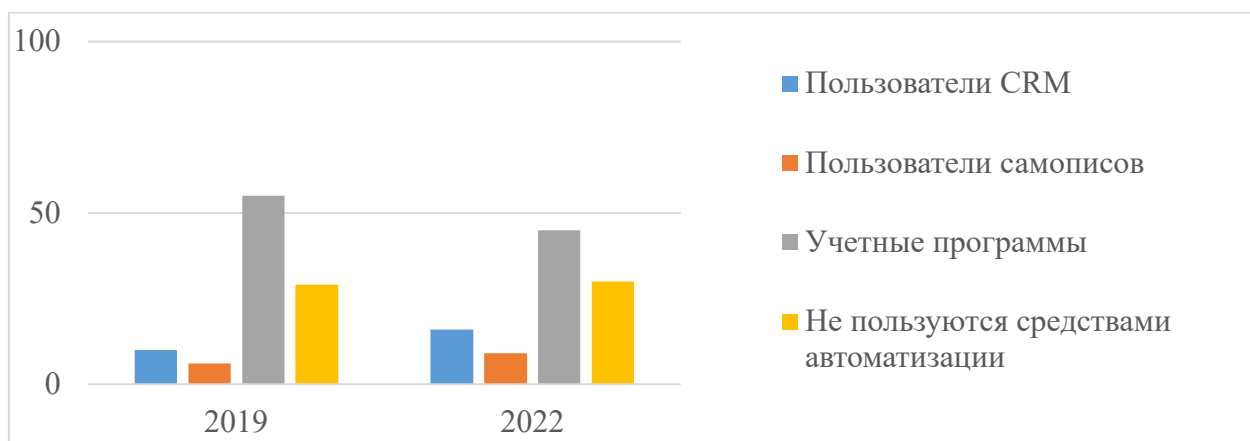


Рисунок 1. Использование разных способов автоматизации работы с клиентами в разные годы, %

С помощью данных технологий можно коммуницировать с клиентами, налаживать с ними взаимоотношения и собирать данные о существующих и будущих клиентах. Такие платформы собирают численные данные, проводят анализ и предоставляют результаты в виде различных графиков и отчетов.

CRM-системы помогают оптимизировать ключевые бизнес-процессы, особенно продажи. Автоматизация процессов маркетинга и продаж – основная часть всех CRM-систем.

CRM маркетинг – бизнес-подход, который подразумевает анализ, контроль и улучшение коммуникации с клиентами и построения с ними взаимоотношений.

Этапы по введению CRM-маркетинга в турфирму

1. Внедрить CRM-систему (офисофф, [tourcontrol](http://tourcontrol.com), u-on.travel лучше всего подойдут для сферы туризма). Работа CRM должна отслеживать все бизнес-процессы, связанные с

коммуникациями с клиентами и продажами. Использовать CRM следует при любом взаимодействии с клиентами.

2. Провести анализ коммуникационных каналов. Понять происхождение лидов, отследить путь клиента, когда и как совершаются покупки.

3. Необходимо подключить сквозную аналитику для отслеживания любого действия клиента.

4. Далее следует настроить сбор статистических данных, например: возраст клиентов, их геолокация, продолжительность пребывания на сайте, клики, глубина просмотра и другое. Выбрать те показатели, которые будут фиксироваться в CRM-системе.

5. Собрать данные и проанализировать. Изучить отчёты, выявить закономерности. Понять сильные и слабые стороны бизнес-процессов компании.

6. Разработать стратегию.

7. С помощью оценки результатов, совершенствовать стратегию.

Благодаря собранному данным, искать способы для улучшения взаимоотношений с клиентами, устранять недостатки и достигать все поставленные бизнес-цели.

CRM-маркетинг помогает осознать нужды аудитории, позволяет удерживать клиентов, повторно привлекать их к покупке и сподвигать потребителей к повторной покупке, чтобы увеличивать число постоянных клиентов.

CRM-маркетинг лучше всего внедрять при условиях:

- многочисленность клиентов;
- высокая стоимость услуг;
- высокая вероятность сбора достоверных данных о потребителях.

Универсальная CRM-система, которая бы подходила к любому бизнесу и любому размеру компании – редкость. Поэтому каждая отдельная организация должна решать для себя, какой вид CRM-системы внедрять, если это целесообразно и необходимо. Существуют следующие виды: операционные CRM-системы (выполняют повседневные процессы), аналитические (базы данных с информацией о клиентах и бизнес-процессах) и коллективные (повышают эффективность работы между различными отделами фирмы). Все они выполняют разные задачи, но цель одна – оптимизировать бизнес-процессы.

Не каждой фирме выгодно использовать CRM-систему, поэтому прежде чем задумываться об оптимизации бизнес-процессов с помощью данного инструмента, следует рассчитать эффективность от ее внедрения, т.е. прежде, чем внедрять данный инструмент на полной основе, его следует протестировать небольшой отрезок времени, а затем понять: окупаются ли затраты на внедрение CRM-системы.

Заключение.

Эффективность от внедрения CRM-системы можно рассчитать с помощью следующих показателей, если имеется модель бизнеса, цели и числовые данные до внедрения инструмента:

- количество предпродаж (upsell);
- продолжительность отдельного этапа воронки продаж;
- продолжительность цикла продаж;
- пожизненная ценность клиента (customer lifetime value);
- стоимость привлечения клиента (customer acquisition cost);
- индекс потребительской лояльности (net promoter score).

Представлен примерный список показателей. Требуемые показатели зависят от цели.

Необходимо помнить, что все плюсы от внедрения CRM-систем проявляются только при правильном использовании инструмента. Внедрение системы не гарантирует решение всех проблем фирме, которые связаны с плохой работой торгового персонала,

незаполненной базы данных и т.п. Важно осознавать цели внедрения CRM-системы и каких результатов хочет добиться компания.

Список литературы

[1] Сайт «Belretail» [Электронный ресурс]. – Электронные данные. – Режим доступа: <https://belretail.by>

[2] Сайт «Offlinecrm» [Электронный ресурс]. – Электронные данные. – Режим доступа: <https://offlinecrm.ru/>

[3] Unisender «CRM-система» [Электронный ресурс]. – Электронные данные. – Режим доступа: <https://www.unisender.com/ru/glossary/>

[4] «CRM-systems» [Электронный ресурс]. – Электронные данные. – Режим доступа: <https://crm-systems.info/crm-dlya-turizma/>

[5] «SendPluse» [Электронный ресурс]. – Электронные данные. – Режим доступа: <https://sendpulse.by/support/glossary/crm>

IMPLEMENTATION OF CRM AND OPTIMIZATION OF BUSINESS PROCESSES IN THE SPHERE OF TOURISM

D.A. Frolova,

Lecturer

BSUIR Economics,

Master of Economic Sciences

P.D. Bulahova

Student of BSUIR

Belarusian State University of Informatics and Radioelectronics

Minsk, Republic of Belarus

E-mail: frolova.profstud@gmail.com, polinabulahov@gmail.com

Abstract. This article discusses the use of CRM systems to optimize business processes. The dynamics of the use of CRM-systems by Belarusian companies is analyzed. The main stages of the implementation of CRM-systems in a travel agency are given.

Keywords: CRM systems, business process optimization, work planning.

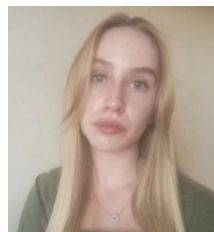
УДК 004.658+339.37

ИСПОЛЬЗОВАНИЕ BIG DATA AND ADVANCED ANALYTICS, А ТАКЖЕ DATA SCIENCE ДЛЯ ОПТИМИЗАЦИИ ПРОИЗВОДСТВЕННЫХ И БИЗНЕС-РЕШЕНИЙ В СФЕРЕ РОЗНИЧНОЙ ТОРГОВЛИ



О.Н. Шкор

Старший преподаватель, магистр экономических наук, доктор философии в области экономики
shkor@bsuir.by



Э.В. Котович

Студентка 3 курса инженерноэкономического факультета специальности "Электронный маркетинг" БГУИР
kotovicherika@gmail.com

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.

Э.В. Котович

Родилась в 2003 году в Минске. В 2020 году закончила ГУО «Средняя школа №147 г. Минска». В этом же году поступила в УО «БГУИР», была зачислена на платную форму обучения по специальности «Электронный маркетинг» инженерно-экономического факультета.

Аннотация. Данная статья содержит в себе комплексное описание применения технологий BIG DATA and Advanced Analytics, а также Data Science для достижения определённых целей в сфере розничной торговли. Big Data способна оптимизировать многие бизнес-процессы, в том числе, способствовать решению маркетинговых вопросов, связанных с изучением поведения потребителей, проведением сегментации клиентов, организацией визуализации данных в удобной форме, построение прогнозов и оценка рыночной ситуации, повышение лояльности потребителей и уменьшение их оттока. В статье также приведены примеры успешного применения данных инструментов крупнейшими ритейл-компаниями.

Ключевые слова: Big Data, розничная торговля, розничная аналитика, маркетинг.

Введение.

«Большие данные» (Big Data) произвели революцию, которая оказала большое влияние на бизнес в целом. Анализируя информацию о поведении потребителей, продавцы получают возможность точно определять тренды, выявлять предпочтения, прогнозировать спрос и оптимизировать предложение.

Big Data и её аналитика позволяет выявить и представить современные тенденции, возникающие закономерности и корреляционные связи, присутствующие в больших объёмах необработанных данных. В дальнейшем, полученные результаты позволяют принимать рациональные решения для различных сфер бизнеса.

Работа с Big Data включает в себя 3 последовательных шага: сбор самих данных, моделирование и применение полученных результатов на практике.

Сбор данных включает в себя использование внутренних и внешних данных, чтобы дать более широкое представление о том, что происходит с операциями внутри компании или с клиентом. Моделирование – это использование этих данных для получения работоспособных

моделей, которые помогут составить более детализированный прогноз, либо позволит провести оптимизацию конкретной отрасли бизнеса. И третий фактор успеха – это трансформация компании для использования этих данных в моделях [1].

Актуальность.

Эволюция Big Data и её популярность в современных реалиях не обошла стороной сферу розничной торговли. Как и для других сфер бизнеса, основная цель «больших данных» в ритейле – монетизирование данных с последующим увеличением прибыли торгового объекта. А потому, изучение использования данных инструментов может принести много пользы как владельцам розничных сетей, так и их потребителям, а, следовательно, является перспективным в современных реалиях.

Розничная аналитика – это процесс использования Big Data для управления процессами ценообразования, движениями цепочек поставок (SCM) и повышения лояльности клиентов. Big Data используются для выявления закономерностей, т

Для сферы ритейла Big Data способствуют более глубокому пониманию покупательских привычек и методов привлечения новых клиентов. Анализ «больших данных» в розничной торговле позволяет компаниям создавать рекомендации для клиентов на основе их истории покупок, что и создаёт так называемый «персонализированный покупательский опыт», а, также, позволяет улучшать уровень обслуживания клиентов.

Такие массивы данных также помогают прогнозировать будущие тенденции и принимать стратегические решения на основе анализа рынка. Более подробно рассмотрим потенциальные возможности использования Big Data и Advanced Analytics для оптимизации бизнес-процессов розничной торговли.

Основные пути использования Big Data и Advanced Analytics для сферы розничной торговли.

1. Создание персонализированного ассортимента, соответствующего потребностям потребителей.

Задача торговых представителей способствовать поставке требуемого ассортимента и формировать грамотную выкладку товаров на торговых площадках. Для достижения этих целей необходимо исследовать такие виды данных, как покупательская способность потребителей, факторы поведения и восприятие, демографическая ситуация в конкретном регионе.

Так, анализ предпочтений потребителей к разным товарам поможет целесообразно использовать финансовые ресурсы компании и грамотно распределять торговое пространство, под соответствующие товарные позиции. А с помощью инструментов видеоаналитики и компьютерного зрения (CV) менеджеры магазина будут получать уведомления о наличии определённого товара, его месторасположении и т.д.

2. Генерация персонифицированного предложения.

Многочисленные маркетинговые исследования подтвердили, что персонализация для клиентов является важным фактором совершения покупки. На этом этапе может использоваться такой инструмент, как программы лояльности. В наши дни он также собирается с помощью транзакций по кредитным картам, IP-адресов, входов пользователей в систему и многого другого. Стоит также отметить и алгоритмы коллаборативной фильтрации (collaborative filtering). Данная фильтрация позволяет генерировать целевые рекомендации на основе ранее полученных данных об интересах пользователей.

Всё это помогает собирать дополнительную информацию для предприятия розничной торговли, которая далее использует её для анализа динамики покупок и расходов потребителей в прошлом, чтобы прогнозировать будущие расходы и предлагать персонализированные рекомендации.

3. Персонализированная доставка.

Скорость—также является одним из составляющих факторов, на который обращает внимание потребитель. Однако, удивить одним лишь этим показателем—недостаточно, важно, также, учитывать индивидуальный подход. Большинство компаний при транспортировке своего товара используют RFID-метки, которые позволяют получить такую информацию: геолокация, габариты груза, метеорологические условия, трафик дорог, а также поведение водителя.

Основная цель использования такого инструмента аналитики—обеспечить ясный процесс транспортировки для клиента и составление экономичного маршрута для предприятия [2].

4. Применение Customer Journey Analytics.

Данный метод представляет собой мониторинг взаимодействия потребителя с брендом на протяжении всего его пути. Инструменты CJA позволяют собирать данные из каждой точки «соприкосновения» бренда и клиента и отслеживать процессы взаимодействия.

CJA может помочь ритейлерам ответить на такие вопросы, как: где клиенты на самом деле ищут информацию о продукте? Где мы их теряем? Каковы наиболее эффективные способы связи с ними и как заставить их купить наш товар?

Основными точками анализа CJA могут быть: изучение отзывов о товарах, просмотр рекламы, подписка и отслеживание новостей о бренде в соц-сетях.

Результатом проведения Customer Journey Analytics является визуализация данных, то есть, составление Customer Journey Map. CJM может представлять собой инфографику, таблицу, схему и другую форму презентации. CJM позволяет увидеть поведение потребителей (их чувства и эмоции), их проблему и боли, а также возможное разрешение данной ситуации. С помощью такой карты, предоставляется возможным ликвидирование ненужных точек «соприкосновения», выявление барьеров при приобретении продуктов и их стадию, устранение данных барьеров и дальнейшее повышение конверсии и улучшение применяемой маркетинговой стратегии [3].

5. Управление торговыми точками с разным местоположением.

Розничные сети и предприятия розничной торговли могут использовать аналитику, чтобы понять разницу в спросе на их продукцию в разных географических точках. Используя аналитику потребительских расходов, продавцы могут использовать эти данные для лучшего обслуживания клиентов в определенных регионах, а также для более эффективного хранения продуктов [4].

6. Стимулирование продаж оптимальными ценами.

Компании могут использовать расширенную аналитику, чтобы оценить, как клиенты реагируют на различные уровни цен на продукты, и предсказать, как изменения цен повлияют на продажи. Это позволит ритейлерам устанавливать конкурентоспособные цены, которые максимизируют рентабельность продаж, и предлагать скидки или наборы продуктов для менее популярных товаров, компенсируя упущенные продажи и увеличивая оборачиваемость запасов.

7. Предупреждение финансовых рисков и планирование путей их решения.

Компании могут использовать аналитику розничных данных, чтобы внимательно отслеживать разрушительные события по мере их развития, оценивать их потенциальное влияние на бизнес и принимать стратегические решения о том, как смягчить последствия сбоя.

Розничные продавцы могут даже применять аналитику к данным в режиме реального времени из внешних источников, чтобы предвидеть сбой еще до того, как они произойдут, и соответствующим образом планировать [5].

Применение технологий Big Data и Advanced Analytics крупнейшими ритейлерами мира.

1. Один из самых известных игроков рынка розничной торговли «Walmart» разрабатывает крупнейшее в мире частное облако с алгоритмами, созданными для отслеживания данных о запасах, транзакциях и активности конкурентов. Это позволяет им практически мгновенно реагировать на изменения рынка.

2. Некоторые компании используют принципы совместного действия. В результате, казалось бы, запутанного сотрудничества между «Pantene», «The Weather Channel» и сетью

гипермаркетов «Walgreens», «Pantene» увидела стремительный рост продаж более чем на 10% в магазинах «Walgreens» благодаря своему проекту «haircast», основанному на данных.

С помощью данных прогноза от «The Weather Channel» розничные продавцы могли продавать выбранные продукты с учетом сезонных изменений и прогноза погоды на неделю, что способствовало увеличению продаж.

3. Глобальный гигант «Amazon» также усовершенствовал искусство сбора и применения данных в своих пользовательских рекомендациях. «Amazon» настолько успешно использует тактику маркетинга и продаж на основе больших данных, что 35% ее продаж приходится на алгоритм рекомендаций клиентов [6].

4. «Starbucks» – этот колоссальный бренд кофе не нуждается в особом представлении – и есть причина, по которой этот бренд из Сиэтла не только выжил, но и процветал на протяжении многих десятилетий своего существования. В двух словах, эта причина – инновации. «Starbucks» обладает сверхъестественной способностью открывать несколько филиалов в одном квартале и получать приличную прибыль от каждого из них. Используя аналитику больших данных в своих интересах, «Starbucks» может прогнозировать потенциал роста каждого нового магазина, анализируя такие показатели, как местоположение, посещаемость, демография района и поведение клиентов. Более того, «Starbucks» собирал информацию о своих более чем 90 миллионах транзакций в неделю и использовала эти данные, чтобы предоставить своим клиентам персонализированный опыт, стимулируя инновации, включая его индивидуальную схему цифровых вознаграждений, которая становится более интуитивно понятной, чем больше данных она собирает о покупательских привычках клиентов и история покупки.

5. Обнаружение мошенничества – это серьезная проблема, направленная на предотвращение потерь и поддержание доверия клиентов. Наиболее часто наблюдаемые виды мошенничества включают мошеннический возврат приобретенных продуктов и кражу информации о кредитной или дебетовой карте.

Манипуляторы всегда изобретают новые инструменты и технологии для мошенничества, и розничные продавцы должны использовать розничную аналитику, чтобы выявлять мошеннические действия и предотвращать их до того, как они произойдут. Благодаря множеству технологий больших данных, таких как «Hadoop», «MapReduce» и «Spark», можно выполнить анализ более 50 петабайт данных, чтобы точно предсказать риски, что ранее было невозможно.

«Amazon» имеет интенсивную программу по обнаружению и предотвращению мошенничества с кредитными картами, которая привела к сокращению мошенничества на 50% в течение первых 6 месяцев. «Amazon» разработала инструменты обнаружения мошенничества, которые используют скоринговый подход в прогнозном анализе.

Эта розничная аналитика зависит от огромных наборов данных, которые содержат не только финансовую информацию о транзакциях, но также отслеживают информацию о браузере, IP-адреса пользователей и любые другие связанные технические данные, которые могут помочь «Amazon» усовершенствовать свои аналитические модели для обнаружения и предотвращения мошенничества [7].

Заключение.

Таким образом, Big Data и Advanced Analytics могут не только повысить операционную маржу компаний более чем на 50%, но и произвести революцию во всех сферах розничной торговли.

Применение этих инструментов позволит выявить целесообразный путь управления запасами и логистикой, а также даст подробные сведения о привычках клиентов.

Они используются для стимулирования продаж, оптимизации процесса продаж с помощью рекомендаций по продуктам и удобных вариантов оплаты, а также для улучшения обслуживания клиентов по всем направлениям.

Роль «больших данных» в розничной торговле также заключается в выявлении потенциальных узких мест и поиске обходных решений до того, как они перерастут в более

серьезные проблемы, что позволит ритейлерам сэкономить на затратах, связанных с простоями и сбоями.

Список литературы

[1] Putting big data and advanced analytics to work: [Электронный ресурс]. URL: <https://www.mckinsey.com/capabilities/growth-marketing-and-sales/our-insights/putting-big-data-and-advanced-analytics-to-work>. (Дата обращения: 20.03.2023).

[2] Big Data на службе розничной торговли [Электронный ресурс]. URL: <https://trends.rbc.ru/trends/innovation/5d66c23c9a79476e7d9aff95> (Дата обращения: 18.03.2023).

[3] Customer journey analytics & Customer Journey Map: что это и зачем применять [Электронный ресурс]. URL: <http://surl.li/fwwfv> (Дата обращения: 19.03.2023).

[4] The Power of Retail Analytics [Электронный ресурс]. URL: <https://www.yodlee.com/data-analytics/big-data-retail-analytics> (Дата обращения: 17.03.2023).

[5] Retail Revolution: The Role of Big Data Analytics in Retail [Электронный ресурс]. URL: <https://global.hitachi-solutions.com/blog/retail-analytics/> (Дата обращения: 19.03.2023).

[6] The importance of big data in retail [Электронный ресурс]. URL: <https://global.hitachi-solutions.com/blog/retail-analytics/> (Дата обращения: 21.03.2023).

[7] 5 Big Data and Hadoop Use Cases in Retail Analytics [Электронный ресурс]. URL: <https://www.projectpro.io/article/5-big-data-and-hadoop-use-cases-in-retail-analytics/91#toc-4> (Дата обращения: 21.03.2023).

POWERING BIG DATA AND ADVANCED ANALYTICS, AND DATA SCIENCE TO OPTIMIZE MANUFACTURING AND BUSINESS DECISIONS IN RETAIL

O.N. Shkor

*Senior Lecturer at the Department of
Economics of BSUIR*

E.V. Kotovich

Student of BSUIR

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

E-mail: shkor@bsuir.by

Abstract. This article contains a comprehensive description of the use of BIG DATA and Advanced Analytics technologies, as well as Data Science, to achieve certain goals in the retail industry. BIG DATA is able to optimize many business processes, including helping to solve marketing issues related to studying consumer behavior, customer segmentation, organizing data visualization in a convenient form, building forecasts and assessing the market situation, increasing consumer loyalty and reducing their churn. The article also provides examples of the successful use of these tools by the largest retail companies.

Keywords: Big Data, retail, retail analytics, marketing.

УДК 004.021:004.75

ПРИМЕНЕНИЕ BIG DATA В СФЕРЕ ОБРАЗОВАНИЯ



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



В.Е. Буюков
Студент БГУИР
vitya.buyukov@mail.ru



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

В.Е. Буюков

Студент кафедры информатики Белорусского государственного университета информатики и радиоэлектроники.

К.И. Котельников

Студент кафедры информатики Белорусского государственного университета информатики и радиоэлектроники.

А.Н. Марков

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

Аннотация. Статья рассматривает роль и применение Big Data в образовательной сфере. В статье анализируются основные преимущества и вызовы использования Big Data в образовании, включая возможности сбора, хранения, анализа и использования больших объемов данных для оптимизации образовательных процессов и улучшения результатов обучения. Рассматриваются различные технологии и инструменты Big Data, такие как Tableau, Power BI и Google Analytics for Education, и примеры их применения в университетской среде.

Ключевые слова: Big Data, образовании, технологии, Tableau.

Введение.

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

Актуальность.

Актуальность использования Big Data в образовании выражена в нескольких ключевых аспектах:

Улучшение образовательных процессов: Анализ больших объемов данных позволяет выявлять тенденции и паттерны в обучении, оптимизировать учебные программы, методики и подходы к обучению. На основе данных можно определить эффективность определенных учебных методик, выявить слабые места в обучении студентов и разработать индивидуальные подходы к обучению, а также предлагать рекомендации по улучшению учебного процесса.

Персонализация обучения: Использование Big Data позволяет создавать персонализированные образовательные программы, адаптированные к индивидуальным потребностям и стилю обучения студентов. Анализ данных о предпочтениях, интересах, уровне знаний и успехах студентов позволяет создавать индивидуальные образовательные планы,

предлагать дополнительные материалы и активности, а также предсказывать возможные трудности в обучении и предлагать соответствующую поддержку.

Принятие обоснованных решений: Использование Big Data позволяет администраторам образовательных учреждений принимать обоснованные решения на основе фактических данных. Анализ данных о показателях успеваемости студентов, эффективности преподавателей, использования ресурсов учебного заведения и других аспектов позволяет оптимизировать распределение ресурсов, планировать бюджет, определять приоритеты и проводить мониторинг реализации образовательных программ.

Раннее выявление рисков и поддержка студентов: Анализ данных о поведенческих и академических показателях студентов позволяет раннее выявлять риски отставания, отсутствия мотивации или возможных трудностей в обучении. Это позволяет предоставлять своевременную поддержку, такую как дополнительные занятия, наставничество или консультации, и предотвращать досрочное прекращение обучения.

Определение Big Data.

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

История использования Big Data

Использование Big Data в образовании имеет довольно долгую историю, начиная с ранних попыток сбора, анализа и использования данных в образовательных организациях. Ниже приведены некоторые ключевые моменты и этапы истории использования Big Data в образовании:

Ранние фазы: В начале 2000-х годов использование Big Data в образовании было ограничено ограниченными и несвязанными усилиями, такими как использование статистических данных и оценок студентов для анализа успеваемости, а также использование простых систем управления обучением (LMS) для сбора и анализа данных о студентах.

Развитие систем управления обучением (LMS): В последующие годы системы управления обучением (LMS), такие как Blackboard, Moodle, Canvas и другие, стали более распространенными в университетах и других образовательных организациях. Эти системы начали собирать и хранить все больше данных о студентах, таких как активность на курсе, оценки, взаимодействие с контентом и другие, что открыло новые возможности для анализа этих данных и использования Big Data в образовании.

Рост доступности и масштабирования технологий Big Data: В середине 2010-х годов с появлением новых технологий Big Data, таких как Apache Hadoop, Apache Spark, и других, стало возможным собирать, хранить и анализировать большие объемы данных в реальном времени. Это позволило более эффективно использовать данные в образовании, такие как данные о студентах, преподавателях, курсах и других аспектах образовательного процесса.

Развитие аналитических платформ: В последние годы появились специализированные аналитические платформы, такие как Blackboard Analytics, Learning Analytics and Knowledge (LAK), Educational Data Mining (EDM) и другие, которые предоставляют инструменты и функции для анализа данных в образовании. Эти платформы позволяют проводить различные виды анализа, такие как анализ успеваемости студентов, анализ эффективности преподавателей, анализ эффективности курсов и других аспектов образовательного процесса на основе данных, собранных из различных источников, включая системы управления обучением, электронные журналы, социальные сети и другие.

Рост интереса к области Learning Analytics: В последние годы Learning Analytics (аналитика обучения) стала одной из наиболее активно развивающихся областей использования Big Data в

образовании. Learning Analytics включает анализ данных обучения, таких как данные о взаимодействии студентов с контентом, времени, проведенном на заданиях, прогрессе в обучении и других факторах, чтобы предоставлять студентам, преподавателям и администраторам образовательных организаций дополнительные инсайты и рекомендации для более эффективного обучения и улучшения образовательного процесса.

Развитие анализа данных для принятия решений: В последние годы анализ данных стал важным инструментом для принятия решений в образовании. Он помогает университетам и другим образовательным организациям оптимизировать процессы обучения и улучшать результаты студентов. С использованием Big Data и аналитических инструментов становится возможным прогнозировать успех студентов, предлагать персонализированные образовательные программы, оптимизировать распределение ресурсов и решать другие стратегические задачи. [2]

Технологии Big Data.

Существует множество технологий Big Data, которые могут быть применены в университетах для улучшения образовательных процессов и решения различных задач. Вот некоторые из таких технологий:

Learning Analytics: Это область, основанная на анализе данных обучения, которая позволяет собирать, анализировать и интерпретировать данные о взаимодействии студентов с образовательными ресурсами, их успехах и прогрессе в обучении. Learning Analytics может использоваться для выявления образовательных трендов, предоставления персонализированных рекомендаций, оценки эффективности образовательных программ и т.д. Примеры популярных инструментов Learning Analytics включают Blackboard Analytics, Brightspace Analytics, и другие.

Data Mining и машинное обучение: Эти технологии позволяют анализировать большие объемы данных для выявления паттернов, тенденций и взаимосвязей. В образовании они могут быть использованы для анализа данных о успеваемости студентов, их активности в образовательных системах, предсказания успеха студентов, определения факторов, влияющих на успех обучения и других целей.

Интернет вещей (IoT): Технологии IoT могут использоваться в университетах для сбора данных о различных аспектах студенческой жизни, таких как посещаемость лекций, активность на кампусе, использование ресурсов и технического оборудования и т.д. Эти данные могут быть анализированы для оптимизации ресурсов, улучшения условий обучения и повышения общего опыта студентов.

Анализ социальных сетей: Социальные сети могут быть использованы для сбора данных о взаимодействии студентов, их активности и социальной динамике в университетской среде. Анализ этих данных может помочь в выявлении тенденций, связанных с социальной интеграцией студентов, предсказании их успеха, выявлении лидеров и определении влияния социального окружения на учебные результаты.

Автоматизированная оценка и анализ работы студентов: существуют технологии, такие как автоматизированная оценка и анализ работы студентов, которые используют алгоритмы машинного обучения и анализ данных для автоматической оценки работ студентов, определения уровня их профессионализма, выявления ошибок и предоставления обратной связи. Это может помочь университетам оптимизировать процесс оценки, улучшить качество обратной связи студентам и облегчить работу преподавателей.

Большие данные в научных исследованиях: В университетах также широко применяются технологии Big Data в научных исследованиях, таких как анализ больших объемов научных данных, моделирование, прогнозирование и анализ результатов экспериментов. Это может помочь ученым в создании новых знаний, определении тенденций и выявлении паттернов в научных данных.

Анализ социально-экономических данных: В университетах также могут применяться технологии Big Data для анализа социально-экономических данных, таких как данные о

занятости выпускников, доходах студентов, рынке труда и других аспектах социально-экономической ситуации студентов и выпускников. Эти данные могут быть использованы для разработки стратегий подготовки студентов к рынку труда, оптимизации карьерных путей и улучшения результатов трудоустройства выпускников.

Анализ данных о ресурсах и инфраструктуре университетов: Технологии Big Data также могут использоваться для анализа данных о ресурсах и инфраструктуре университетов, таких как использование аудиторий, библиотек, спортивных сооружений, транспортных средств и других ресурсов. Это может помочь в оптимизации распределения ресурсов, улучшении эффективности и экономии ресурсов. [3]

Технологии, применяемые в университетах

Tableau: Tableau – это популярный инструмент для визуализации данных и аналитики, который может быть использован в университетах для анализа данных о студентах, преподавателях, курсах и других аспектах академической деятельности. Он предлагает мощные возможности визуализации данных, создания интерактивных дашбордов и отчетов, а также инструменты для анализа данных различных источников. Применяется в Гарвардский университет (Harvard University), Стэнфордский университет (Stanford University), Мичиганский университет (University of Michigan), Университет Оклахомы (University of Oklahoma).

PowerBI: PowerBI – это еще один популярный инструмент для визуализации данных и аналитики, разработанный компанией Microsoft. Он также может быть использован в университетах для анализа данных, связанных с академическим процессом, студентами, преподавателями и другими аспектами образования. PowerBI предлагает интуитивный интерфейс, возможности визуализации данных и мощные инструменты анализа данных. Применяется в Массачусетский технологический институт (Massachusetts Institute of Technology, MIT), Университет Мичигана (University of Michigan), Калифорнийский университет, Беркли (University of California, Berkeley).

Google Analytics for Education: Это решение, предложенное Google, специально разработанное для образовательных учреждений, которое позволяет анализировать данные о посещениях веб-сайтов, использовании онлайн-платформ и других аспектах цифровой деятельности в образовании. Он предлагает инструменты для анализа данных о посетителях, взаимодействии с веб-ресурсами и эффективности онлайн-курсов.

Learning Management System (LMS) Analytics: Множество Learning Management System (LMS), таких как Moodle, Canvas, D2L Brightspace и другие, также предлагают встроенные инструменты аналитики данных. Они могут быть использованы для анализа данных о студентах, преподавателях, курсах, взаимодействии со студентами и других аспектах обучения и образовательной деятельности.

Все эти технологии применяются Массачусетский технологический институт (Massachusetts Institute of Technology, MIT), Университет Мичигана (University of Michigan), Калифорнийский университет, Беркли (University of California, Berkeley), Гарвардский университет (Harvard University), Стэнфордский университет (Stanford University), Университет Оклахомы (University of Oklahoma)[4]

Принцип работы Tableau.

Подключение к данным: Tableau позволяет подключаться к различным источникам данных, таким как базы данных, электронные таблицы, веб-сервисы, файлы CSV и другие. Tableau поддерживает множество источников данных и позволяет визуализировать данные из них без необходимости предварительной обработки.

Создание визуализаций: С использованием Tableau Desktop, пользователи могут создавать визуализации данных путем перетаскивания и настройки полей данных на полотно рабочего пространства. Tableau предоставляет широкий набор графических элементов, таких как столбчатые диаграммы, круговые диаграммы, линейные графики, карты, таблицы и другие, которые могут быть использованы для создания разнообразных визуализаций.

Взаимодействие с визуализациями: Созданные визуализации в Tableau являются интерактивными, что позволяет пользователям взаимодействовать с данными на визуализациях. Пользователи могут выполнять различные действия, такие как фильтрация данных, сортировка, детализация, агрегация и другие, чтобы исследовать данные и получить инсайты.

Создание дашбордов: Tableau также позволяет создавать дашборды – композиции из нескольких визуализаций, которые предоставляют совокупный обзор данных. Пользователи могут создавать дашборды, устанавливать фильтры и взаимодействовать с визуализациями на дашборде, чтобы получить комплексное представление данных.

Публикация и обмен визуализациями: С использованием Tableau Server или Tableau Online, пользователи могут публиковать созданные визуализации и дашборды в интернете и делиться ими с другими пользователями через веб-браузер. Tableau Server и Tableau Online также обеспечивают возможность управления правами доступа, планирования автоматических обновлений данных и масштабирования для обслуживания больших пользовательских сообществ.

Мобильный доступ: Tableau также предоставляет мобильные приложения, такие как Tableau Mobile, которые позволяют пользователям доступ к визуализациям и дашбордам через мобильные устройства, такие как смартфоны и планшеты. Это позволяет пользователям получать доступ к визуализациям и анализировать данные даже вне офиса или на ходу.

Автоматизация и интеграция: Tableau также предоставляет возможности автоматизации процессов, таких как автоматическое обновление данных, создание расписаний публикации, создание сценариев и автоматическое выполнение действий на основе данных. Tableau также имеет мощные возможности интеграции с другими инструментами и системами, такими как базы данных, системы хранения данных, CRM, ERP и другие.

Развитие и расширение функциональности: Tableau активно развивается и предлагает регулярные обновления и новые функциональности, такие как расширенные возможности визуализации, аналитики, геоаналитики, машинного обучения и другие. Это позволяет пользователям использовать Tableau для более сложных и продвинутых аналитических задач.[5]

Показатели эффективности.

Давайте рассмотрим несколько примеров повышения показателей эффективности Tableau на конкретных университетах:

Университет штата Мичиган (США). Сокращение времени на анализ данных: Внедрение Tableau позволило сотрудникам университета сократить время на анализ данных на 50%. Ранее процесс анализа данных занимал несколько дней, теперь эту работу можно выполнить за несколько часов. Увеличение числа пользователей: Внедрение Tableau также способствовало росту числа пользователей. Вначале использовали Tableau около 100 пользователей, теперь более 1000 сотрудников университета активно используют Tableau для анализа данных и создания визуализаций.

Университет Шеффилда (Великобритания). Улучшение качества принимаемых решений: Внедрение Tableau в университете Шеффилда позволило сотрудникам более быстро и точно анализировать данные о поступлении студентов, академических результатах и исследованиях. Это способствовало улучшению качества принимаемых решений и оптимизации учебных программ.

Университет Уотерлоо (Канада). Оптимизация процессов финансового анализа: Tableau был успешно внедрен в университете Уотерлоо для оптимизации процессов финансового анализа, включая бюджетирование, планирование и анализ затрат. Это позволило университету более эффективно управлять финансовыми ресурсами и принимать более обоснованные решения.

Университет Мельбурна (Австралия). Увеличение прозрачности и доступности данных: Tableau был внедрен в университете Мельбурна с целью увеличения прозрачности и доступности данных. Ранее множество данных было разбросано в различных системах, и

доступ к ним был ограничен. Внедрение Tableau позволило создать единый источник данных и предоставить более широкий доступ к аналитическим визуализациям, что способствует эффективности на 15%.

Заключение.

Big Data имеет огромный потенциал для трансформации образования, предоставляя возможности для более эффективного анализа данных, оптимизации процессов, принятия более информированных решений и улучшения качества образования. Внедрение технологий Big Data в университеты и другие образовательные учреждения может привести к улучшению успеваемости студентов, оптимизации ресурсов, повышению эффективности учебных программ и управления образовательными процессами.

Список литературы

- [1] Gartner, Inc. "Big Data." Gartner IT Glossary [Электронный ресурс], Адрес: <https://www.gartner.com/en/information-technology/glossary/big-data>.
- [2] Long, P., & Siemens, G. (2011). Penetrating the fog: Analytics in learning and education. *EDUCAUSE review*, 46(5), 30-40.
- [3] Ferguson, R. (2012). Learning analytics: Drivers, developments and challenges. *International Journal of Technology Enhanced Learning*, 4(5/6), 304-317.
- [4] Bolliger, D. U., & Halupa, C. (2018). Analytics and Data-Driven Decision Making in Higher Education: Benefits, Challenges, and Future Directions. *Online Learning Journal*, 22(3), 115-136.
- [5] Официальный сайт Tableau [Электронный ресурс] – Адрес: <https://www.tableau.com/>

APPLICATION OF BIG DATA IN THE FIELD OF EDUCATION

K.I.Katselnikau
Student of the BSUIR

V.E. Buyukov
Student of the BSUIR

A.N.MARKOV
Senior lecturer of the department,
Deputy head of the Center for
Informatization and Innovative
Developments

Department of Computer Science
Faculty of Computer Systems and Networks
Belarusian State University of computer science and Radio Electronics, Republic of Belarus
E-mail: vitya.buyukov@mail.ru

Abstract. The article examines the role and application of Big Data in the educational sphere. The article analyzes the main advantages and challenges of using Big Data in education, including the possibilities of collecting, storing, analyzing and using large amounts of data to optimize educational processes and improve learning outcomes. Various Big Data technologies and tools, such as Tableau, Power BI and Google Analytics for Education, and examples of their application in the university environment are considered.

Keywords: Big Data, Education, Technology, Tableau.

УДК 004.42:004.65

ОБЗОР ПРОГРАММНОЙ ПЛАТФОРМЫ APACHE HADOOP ДЛЯ ОБРАБОТКИ И ХРАНЕНИЯ БОЛЬШИХ ДАННЫХ



Г.А. Пискун
Доцент кафедры
проектирования
информационно-компьютерных
систем БГУИР, кандидат
технических наук, доцент
piskunbsuir@gmail.com



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



Т.М. Воронко
Инженер-программист
Центра
информатизации и
инновационных
разработок БГУИР,
магистрант БГУИР
voronko232001@gmail.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.

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.

Т.М. Воронко

Магистрант 1 курса специальности «Электронные системы и технологии» кафедры проектирования информационно-компьютерных систем. Инженер-программист в отделе инновационных разработок в сфере образования Центра информатизации и инновационных разработок Белорусского государственного университета информатики и радиоэлектроники.

Аннотация. Выполнен обзор программной платформы для хранения и обработки больших данных на примере *Apache Hadoop*, которая представляет собой свободно распространяемый набор утилит, библиотек и фреймворк для разработки и выполнения распределённых программ, работающих на кластерах из сотен и тысяч узлов. Рассмотрена концептуальная архитектура платформы через описание основных модулей, входящих в нее: *Hadoop Common*, *Hadoop Distributed File System*, *Hadoop YARN* и *Hadoop MapReduce*. Проанализирована область применения данной технологии. Установлено, что такие платформы для обработки и хранения больших данных как *Apache Hadoop*, являются одними из самых важных инструментов для работы с данными в современном мире, обеспечивая безопасность инфраструктуры и оптимизируя бизнес-процессы.

Ключевые слова: *Big Data*, архитектура, инфраструктура.

Введение.

Платформа для работы с большими данными (*Big Data*) выполняет множество функций: «очищение» массива данных от лишней информации; обработка и структурирование массива данных; анализ массива данных; защита данных; обеспечение доступа ко всему объёму постоянно изменяемых данных и так далее. Из всех вышеперечисленных функций приоритетное значение имеет анализ постоянно обновляемых данных. В современных условиях результаты такого анализа будут иметь решающее значение для компаний и предприятий при создании новых персонализированных товаров и услуг, позволят спрогнозировать дальнейшее направление развития [1].

В настоящее время существует множество проектов для работы с *big data*, которые отличаются разными моделями, структурой, особенностями анализа, разнообразными программными комплексами и техническими требованиями. Одной из самых распространенных и эффективных архитектур *big data* является *Apache Hadoop (Hadoop)* [1].

Несмотря на то, что платформа *Hadoop* старше и уступает в скорости во многих задачах своим аналогам, таким как *Apache Spark*, она все еще пользуется огромной популярностью среди разработчиков. Платформа надежна и хорошо протестирована, может быть развернута на большей части UNIX-подобных операционных систем, не будучи сильно требовательной к ресурсам. Благодаря предоставлению файловой системы и распределению рабочей нагрузки, использование платформы эффективно с финансовой точки зрения. Также *Hadoop* поддерживается многими облачными сервисами [2].

Рассмотрим более подробно концептуальную архитектуру платформы.

Концептуальная архитектура платформы *Hadoop*.

Hadoop – это основополагающая технология хранения и обработки больших данных, являющаяся проектом верхнего уровня фонда *Apache Software Foundation* [3].

Проект состоит из 4-х основных модулей [3]: *Hadoop Common*, *Hadoop Distributed File System*, *Hadoop YARN*, *Hadoop MapReduce*. Схематично концептуальная архитектура платформы *Hadoop* представлена на рисунке 1, где основные модули обведены красным цветом.

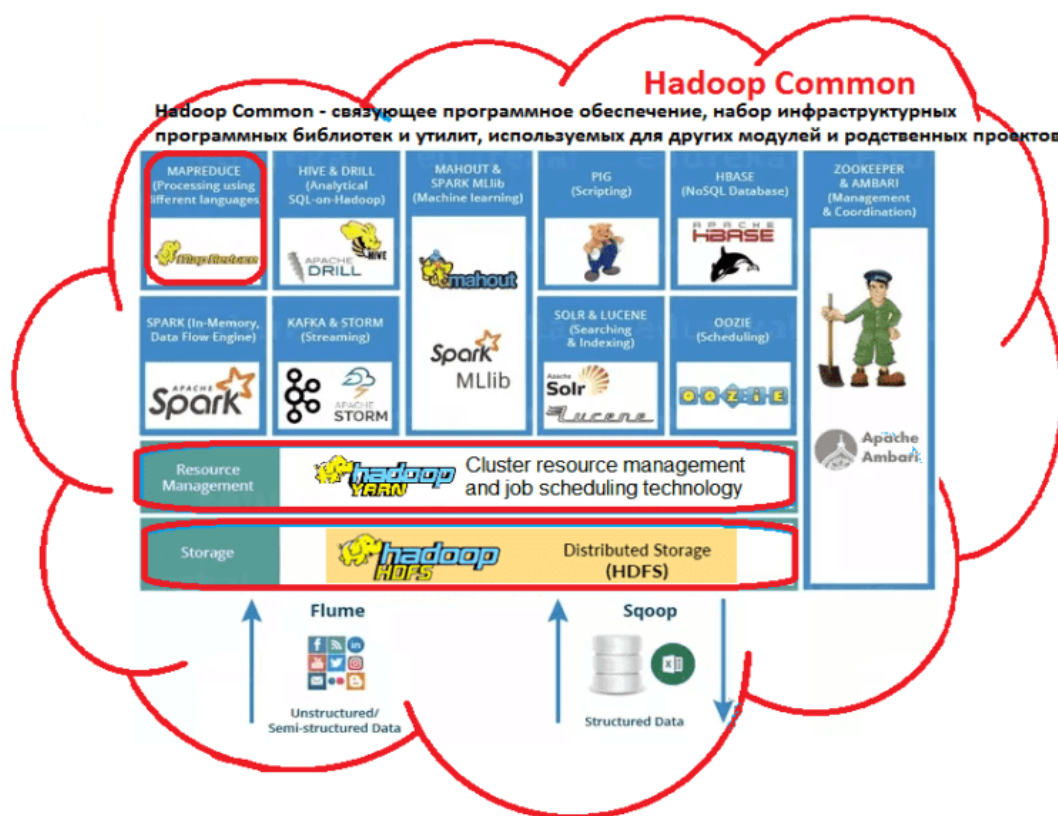


Рисунок 1. Концептуальная архитектура платформы *Hadoop*

Рассмотрим данные модули более подробно:

1. ***Hadoop Common*** – набор инфраструктурных программных библиотек и утилит, которые используются в других решениях и родственных проектах, в частности, для управления распределенными файлами и создания необходимой инфраструктуры [3].

В *Hadoop Common* входят библиотеки управления файловыми системами, поддерживаемыми *Hadoop*, и сценарии создания необходимой инфраструктуры и управления распределённой обработкой, для удобства выполнения которых создан специализированный упрощённый интерпретатор командной строки (*FS shell, filesystem shell*), запускаемый из оболочки операционной системы командой вида: *hdfs dfs -command URI*, где *command* – команда интерпретатора, а *URI* – список ресурсов с префиксами, указывающими тип поддерживаемой файловой системы, например *hdfs://example.com/file1* или *file:///tmp/local/file2*. Большая часть команд интерпретатора реализована по аналогии с соответствующими командами *Unix* (например, *cat, chmod, chown, chgrp, cp, du, ls, mkdir, mv, rm, tail*). При этом поддерживаются некоторые ключи аналогичных *Unix*-команд, например, ключ рекурсивности *-R* для *chmod, chown, chgrp*, есть команды специфические для *Hadoop* (например, *count* подсчитывает количество каталогов, файлов и байтов по заданному пути, *expunge* очищает корзину, а *setrep* модифицирует коэффициент репликации для заданного ресурса) [4].

2. ***Hadoop Distributed File System (HDFS)*** – распределённая файловая система *Hadoop* для хранения файлов больших размеров с возможностью потокового доступа к информации, поблочно распределённой по узлам вычислительного кластера, который может состоять из произвольного аппаратного обеспечения. *HDFS*, как и любая файловая система – это иерархия каталогов с вложенными в них подкаталогами и файлами [5].

На рисунке 2 отображено схематическое представление *HDFS* [6].

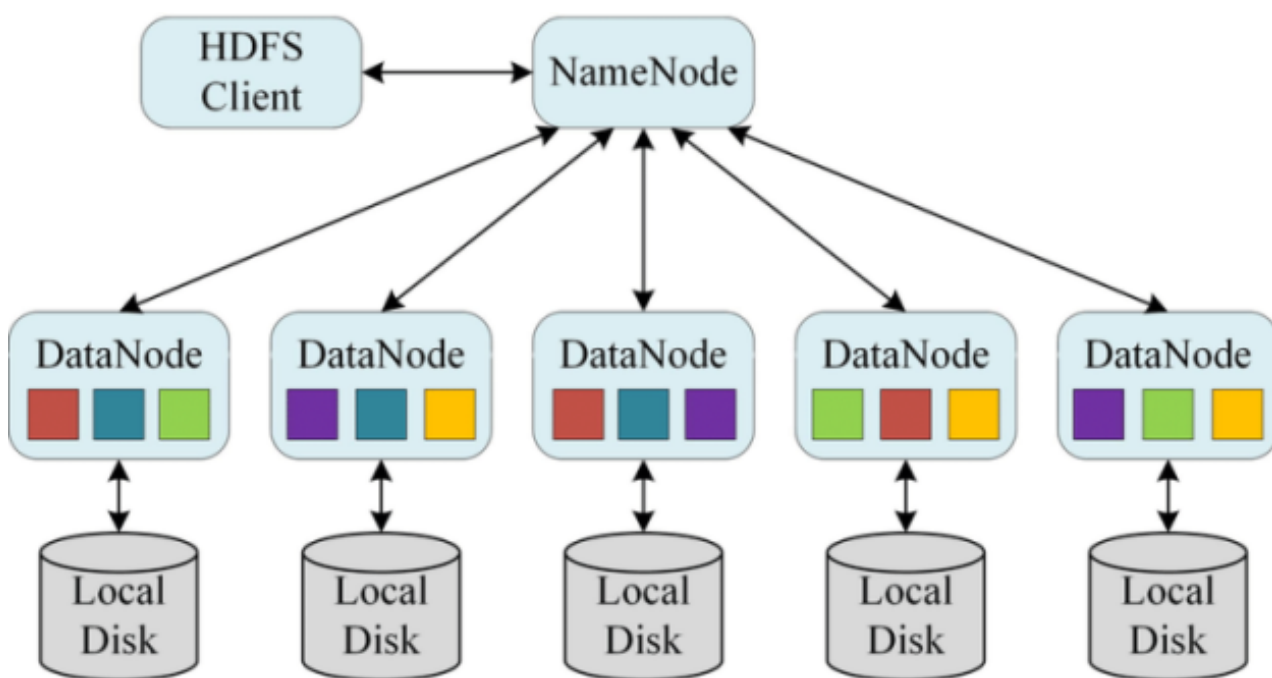


Рисунок 2. Схематическое представление *HDFS*

Кластер *HDFS* включает следующие компоненты [5]:

– управляющий узел, узел имен или сервер имен (*NameNode*) – отдельный, единственный в кластере, сервер с программным кодом для управления пространством имен файловой системы, хранящий дерево файлов, а также метаданные файлов и каталогов. *NameNode* – обязательный компонент кластера *HDFS*, который отвечает за открытие и закрытие файлов, создание и удаление каталогов, управление доступом со стороны внешних клиентов и соответствие между файлами и блоками, дублированными (реплицированными) на узлах данных. Сервер имён раскрывает для всех желающих расположение блоков данных на машинах кластера;

– *secondary NameNode* – вторичный узел имен, отдельный сервер, единственный в кластере, который копирует образ *HDFS* и лог транзакций операций с файловыми блоками во временную папку, применяет изменения, накопленные в логе транзакций к образу *HDFS*, а также записывает его на узел *NameNode* и очищает лог транзакций. *Secondary NameNode* необходим для быстрого ручного восстановления *NameNode* в случае его выхода из строя;

– узел или сервер данных (*DataNode, Node*) – один из множества серверов кластера с программным кодом, отвечающим за файловые операции и работу с блоками данных. *DataNode* – обязательный компонент кластера *HDFS*, который отвечает за запись и чтение данных, выполнение команд от узла *NameNode* по созданию, удалению и репликации блоков, а также периодическую отправку сообщения о состоянии (*heartbeats*) и обработку запросов на чтение и запись, поступающих от клиентов файловой системы *HDFS*. Стоит отметить, что данные проходят с остальных узлов кластера к клиенту мимо узла *NameNode*;

– клиент (*client*) – пользователь или приложение, взаимодействующий через специальный интерфейс (*API – Application Programming Interface*) с распределенной файловой системой. При наличии достаточных прав, клиенту разрешены следующие операции с файлами и каталогами: создание, удаление, чтение, запись, переименование и перемещение. Создавая файл, клиент может явно указать размер блока файла (по умолчанию 64 Мб) и количество создаваемых реплик (по умолчанию значение равно 3-ем).

Взаимодействие узлов имен, узлов данных и клиентов осуществляется по протоколам на основе *TCP/IP*.

3. **Hadoop YARN** – система планирования заданий и управления кластером (*Yet Another Resource Negotiator*), которую также называют *MapReduce 2.0 (MRv2)* – набор системных программ (демонов), обеспечивающих совместное использование, масштабирование и надежность работы распределенных приложений.

Фактически, *YARN* является интерфейсом между аппаратными ресурсами кластера и приложениями, использующих его мощности для вычислений и обработки данных [3].

Архитектура и принцип работы *Hadoop YARN* представлены на рисунке 3 [7].

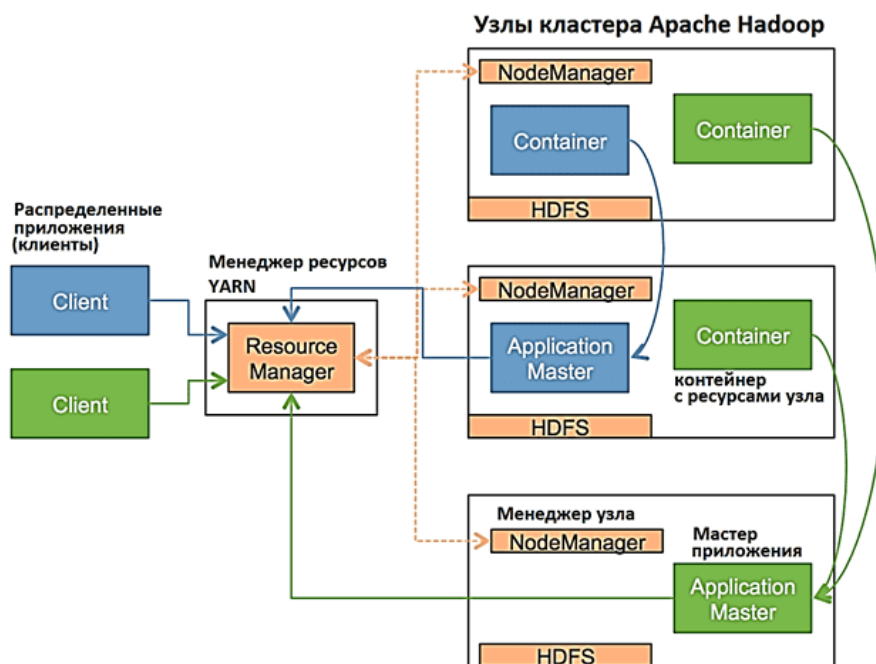


Рисунок 3. Архитектура и принцип работы *Hadoop YARN*

YARN состоит из следующих компонентов [7]:

– *ResourceManager (RM)* – менеджер ресурсов, которых отвечает за распределение ресурсов, необходимых для работы распределенных приложений, и наблюдение за узлами кластера, где эти приложения выполняются. *ResourceManager* включает планировщик ресурсов (*Scheduler*) и диспетчер приложений (*ApplicationsManager, AsM*);

– *ApplicationMaster (AM)* – мастер приложения, ответственный за планирование его жизненного цикла, координацию и отслеживание статуса выполнения, включая динамическое масштабирование потребления ресурсов, управление потоком выполнения, обработку *ошибок* и искажений вычислений, выполнение локальных оптимизаций. Каждое приложение имеет свой экземпляр *ApplicationMaster*. *ApplicationMaster* выполняет произвольный пользовательский код и может быть написан на любом языке программирования благодаря расширяемым протоколам связи с менеджером ресурсов и менеджером узлов;

– *NodeManager (NM)* – менеджер узла – агент, запущенный на узле кластера, который отвечает за отслеживание используемых вычислительных ресурсов (*CPU, RAM* и пр.), управление логами и отправку отчетов об использовании ресурсов планировщику. *NodeManager* управляет абстрактными контейнерами – ресурсами узла, доступными для конкретного приложения;

– Контейнер (*Container*) – набор физических ресурсов (ЦП, память, диск, сеть) в одном вычислительном узле кластера.

4. **Hadoop MapReduce** – платформа программирования и выполнения распределённых *MapReduce*-вычислений, с использованием большого количества компьютеров (узлов, *nodes*), образующих кластер [3].

Суть *MapReduce* состоит в разделении информационного массива на части, параллельной обработке каждой части на отдельном узле и финального объединения всех результатов [8].

Принцип работы *MapReduce* представлен на рисунке 4 [8].

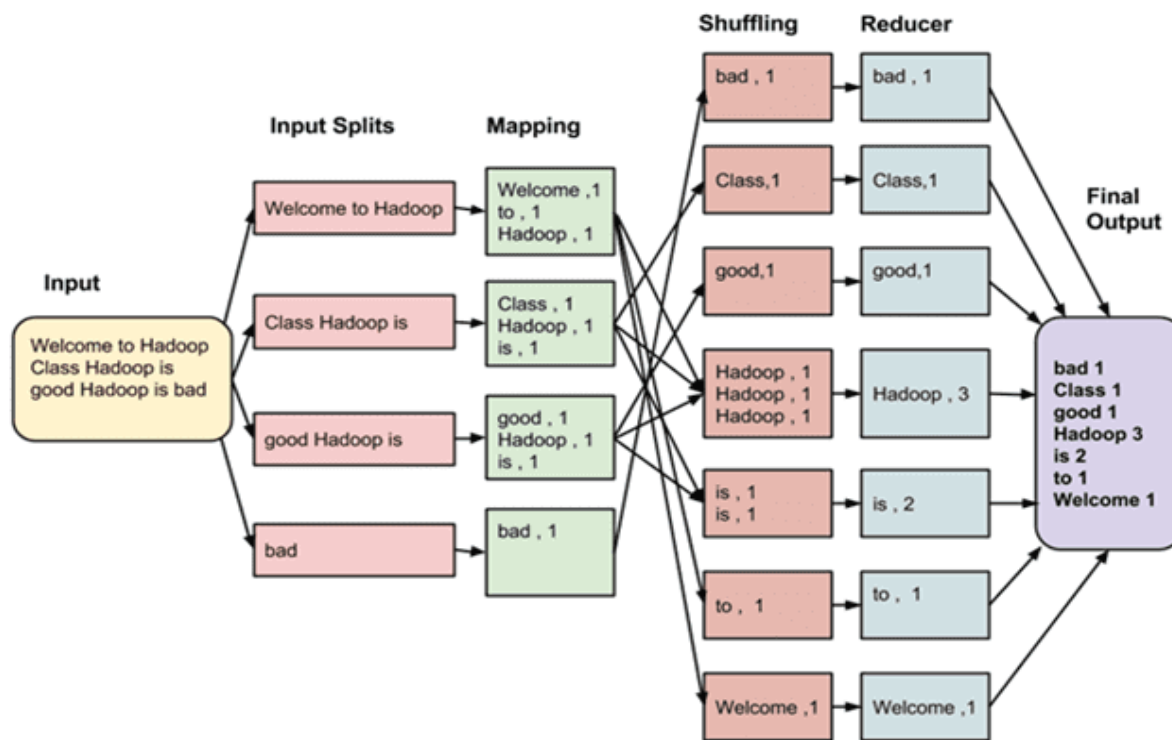


Рисунок 4. Принцип работы *MapReduce*

Вычислительная модель состоит из 3-хшаговой комбинации следующих функций [8]:

– *Map* – предварительная обработка входных данных в виде большого списка значений. При этом главный узел кластера (*master node*) получает этот список, делит его на части и передает рабочим узлам (*worker node*). Далее каждый рабочий узел применяет функцию *Map* к локальным данным и записывает результат в формате «ключ-значение» во временное хранилище.

– *Shuffle* – перераспределение рабочими узлами данных на основе ключей, ранее созданных функцией *Map*, таким образом, чтобы все данные одного ключа лежали на одном рабочем узле.

– *Reduce* – параллельная обработка каждым рабочим узлом каждой группы данных по порядку следования ключей и «склейка» результатов на *master node*. Главный узел получает промежуточные ответы от рабочих узлов и передает их на свободные узлы для выполнения следующего шага. Получившийся после прохождения всех необходимых шагов результат – это и есть решение исходной задачи.

Область применения *Hadoop*.

Платформа активно используется для регулирования рисков и безопасности инфраструктуры, кроме этого, она необходима для оптимизации бизнес-процессов, проведения финансового анализа, выполнения маркетингового анализа и изучения неструктурированной информации, которая собиралась во время продаж [9].

Несистематизированную информацию, собранную из различных источников, нередко называют «озером данных». Как правило, компания не нуждается в таких сведениях, но обязана хранить их по закону. Некоторые организации после завершения анализа данных находят им применение для будущих проектов или задач [9].

При хранении информации в разных источниках и форматах ее анализ, моделирование и прогнозирование затруднено. По сути, «озеро данных» является ненужным для компании, так как не может принести никакой практической пользы. С помощью данной технологии становится возможным распределить и классифицировать сведения, а затем – провести их аналитику и получить различные результаты [9].

Hadoop также часто применяется в следующих целях [9]:

1. *Изучение данных из социальных сетей.* Как правило, сведения из соцсетей позволяют организации лучше понять потребности аудитории. С помощью *Hadoop* анализируются интересы, уровень доходов, уровень образования. Такой подход позволяет настроить таргетированную рекламу, управлять репутацией бренда и повышать отклик из групп и аккаунтов.

2. *Анализ отношения к компании.* Если говорить о том, для чего используется *Hadoop*, то нельзя не упомянуть этот параметр. Технология позволяет анализировать мнения клиентов, которые они высказывают в соцсетях, блогах, отзывах. Удастся проанализировать отношение покупателей к отдельным продуктам / услугам и всему бренду в целом. Это помогает спрогнозировать покупки, скорректировать маркетинговое продвижение товара или оценить имеющуюся репутацию на рынке.

3. *Поддержание безопасности инфраструктуры.* Решение позволяет отслеживать серверные журналы и выявлять любые нарушения безопасности. С помощью технологии можно выявить возможные риски, обнаружить сетевые атаки и спрогнозировать возможные проблемы.

4. *Исследование геоданных.* Компании в области ритейла и производства нередко (с согласия клиентов) собирают сведения об их местоположении. Полученные сведения позволяют в будущем спрогнозировать визиты пользователей или подобрать товары с учетом геолокации. С помощью платформы осуществляется оптимизация и обработка таких геоданных, что упрощает всю процедуру и сокращает время на ее выполнение.

5. *Сбор сведений о поведении клиентов.* Нередко технология оказывается полезной при сборе и обработке данных о поведении и вовлеченности пользователей на сайте. Например, платформа помогает обработать информацию о том, откуда пользователи перешли на сайт, на какую страницу попали, сколько времени провели за просмотром контента, с каких страниц чаще

всего уходили. Анализ подобных данных позволяет организации повысить конверсию ресурса и сделать его более удобным с точки зрения пользователей.

Заключение.

Сегодня *Hadoop* можно встретить в самых разных отраслях – от производства до госсектора. Финансовые и транспортные компании, частные клиники и многие другие организации активно используют данную платформу для управления, анализа и повышения эффективности рабочего процесса. В мире, где огромная часть информации хранится в цифровом виде, а объемы данных становятся все больше, такие платформы, как *Hadoop*, являются просто незаменимыми.

Список литературы

- [1] Big Data: технология, принципы и архитектура [Электронный ресурс]. URL: <http://ipcmagazine.ru/legal-issues/big-data-technology-principles-and-architecture> (Дата обращения: 28.03.2023).
- [2] Top Big Data frameworks: what will tech companies choose in 2023? [Электронный ресурс]. URL: <https://jelvix.com/blog/top-5-big-data-frameworks> (Дата обращения: 28.03.2023).
- [3] Что такое Hadoop, где и зачем используется - хадуп для чайников [Электронный ресурс]. URL: <https://www.bigdataschool.ru/wiki/hadoop> (Дата обращения: 28.03.2023).
- [4] Hadoop [Электронный ресурс]. URL: <https://ru.wikipedia.org/wiki/Hadoop> (Дата обращения: 28.03.2023).
- [5] Что такое HDFS: все о Hadoop Distributed File System [Электронный ресурс]. URL: <https://www.bigdataschool.ru/wiki/hdfs> (Дата обращения: 28.03.2023).
- [6] The overview of the Hadoop Distributed File System HDFS [Электронный ресурс]. URL: https://www.researchgate.net/figure/The-overview-of-the-Hadoop-Distributed-File-System-HDFS_fig4_348387085 (Дата обращения: 28.03.2023).
- [7] Что такое YARN: архитектура, принцип работы, роль в Apache Hadoop [Электронный ресурс]. URL: <https://www.bigdataschool.ru/wiki/yarn> (Дата обращения: 28.03.2023).
- [8] MapReduce Architecture [Электронный ресурс]. URL: <https://www.geeksforgeeks.org/mapreduce-architecture/> (Дата обращения: 28.03.2023).
- [9] Hadoop, где и зачем используется [Электронный ресурс]. URL: <https://www.xelent.ru/blog/chto-takoe-hadoop-gde-i-zachem-ispolzuetsya/> (Дата обращения: 28.03.2023).

REVIEW OF THE SOFTWARE PLATFORM FOR PROCESSING AND STORING BIG DATA ON THE EXAMPLE OF APACHE HADOOP

G.A. Piskun

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

V.F. Alekseev

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

T.M. Voronko

Software Engineer of the Center of Informatization and Innovative Developments of BSUIR, master student of BSUIR

Faculty of Computer Engineering

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

E-mail: voronko232001@gmail.com

Abstract. A review of the software platform for storing and processing big data is made using the example of Apache Hadoop, as one of the most common and effective today. This is a freely distributed set of utilities, libraries and framework for developing and executing distributed programs running on clusters of hundreds and thousands of nodes. The conceptual architecture of the platform is considered through the description of the main modules included in it: Hadoop Common, Hadoop Distributed File System, Hadoop YARN, Hadoop MapReduce, and the scope of this technology is analyzed.

As a result of the analysis, it was found that such platforms for processing and storing big data, such as Apache Hadoop, are one of the most important tools for working with data in the modern world, ensuring infrastructure security and optimizing business processes.

Keywords: big data, architecture, infrastructure.

УДК 001.891

ОРГАНИЗАЦИЯ РАЗДЕЛОВ ЭЛЕКТРОННОЙ ФОРМЫ РЕГИСТРАЦИИ УЧАСТНИКОВ НАУЧНОЙ КОНФЕРЕНЦИИ С ФОРМИРОВАНИЕМ ОНЛАЙН БАЗЫ ДАННЫХ



В.Ф. Алексеев

Доцент кафедры
проектирования

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



Г.А. Пискун

Доцент кафедры
проектирования

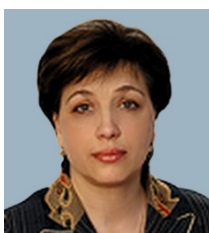
информационно-компьютерных
систем БГУИР, кандидат
технических наук, доцент
piskunbsuir@gmail.com



Д.В. Лихачевский

Декан факультета

компьютерного проектирования
БГУИР, кандидат технических
наук, доцент
likhachevskiyd@bsuir.by



И.Н. Тонкович

Доцент кафедры
проектирования

информационно-компьютерных
систем БГУИР, кандидат
химических наук, доцент
intonkovich@gmail.com



Э.В. Асадчая

Магистрант гр.115401,
ассистент кафедры

проектирования
информационно-компьютерных
систем БГУИР
elina.asadchaya@gmail.com



А.Д. Ларькин

Магистрант гр.215441,
ассистент кафедры

проектирования
информационно-компьютерных
систем БГУИР
anton11061998@gmail.com

В.Ф. Алексеев

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

Г.А. Пискун

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

Д.В. Лихачевский

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

И.Н. Тонкович

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

Э.В. Асадчая

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

А.Д. Ларькин

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

Аннотация. Представлен принцип построения электронной формы регистрации, включающий в себя 6 основных блоков данных. Разработаны алгоритмы организации данных блоков для студенческой научной конференции, в которых определены обязательные и второстепенные маркеры участников. Реализована возможность валидации указанных данных. Показана возможность формирования удаленной базы данных с определением формы и ступени обучения, а также данных о научном руководителе, учреждении высшего образования, кафедре и секции участия.

Ключевые слова: научно-исследовательская работа студентов, научно-исследовательских потенциал, образовательный процесс, принцип построения, алгоритм регистрации, научная конференция.

Введение.

Интеграция науки и образования происходит в разных формах, с различной интенсивностью и во многом определяется уровнем организации образовательного процесса и научной работы на кафедрах вузов. В структурных подразделениях университета предоставляются различные условия, способствующие накоплению научно-исследовательского потенциала, повышению уровня подготовки специалистов, научных и научно-педагогических кадров, оптимизации исследовательской работы [1–12].

Для повышения качества подготовки специалистов во время получения высшего образования или академической степени магистра наук происходит изменение образовательного процесса в направлении развития и реализации творческих способностей обучающихся [1–12]. Наиболее эффективной стратегией для улучшения образовательного процесса является организация научно-исследовательской работы студентов (НИРС).

Одной из наиболее перспективных задач учреждений образования является подготовка всесторонне развитых специалистов, которые способны непрерывно пополнять и углублять свои знания, а также повышать теоретический и профессиональный уровень. В связи с этим постоянно осуществляются меры, направленные на повышение эффективности учебно-образовательного процесса и научно-исследовательской деятельности путем интеграции науки и производства, а также оперативного и гибкого обновления содержания учебных материалов [12–14]. Одним из наиболее популярных видов апробации научных результатов являются гибридные конференции.

В период локдаунов, многие организации перешли на удаленную форму работы и гибридная форма проведения конференций стала, несомненно, удобной как для сотрудников, так и для организаторов их проведения. Гибридные мероприятия, по сути, существовали с момента появления виртуальных, но в последнее время они стали приобретать популярность, поскольку мир уменьшает количество локдаунов и возвращается к нормальной жизни после COVID-19 [15–16].

Принцип построения электронной формы регистрации.

Одним из важнейших этапов организации научной конференции является регистрация участников, сбор и обработка материалов конференции [16]. Для этого наиболее эффективным является следующий принцип, представленный на рисунке 1.

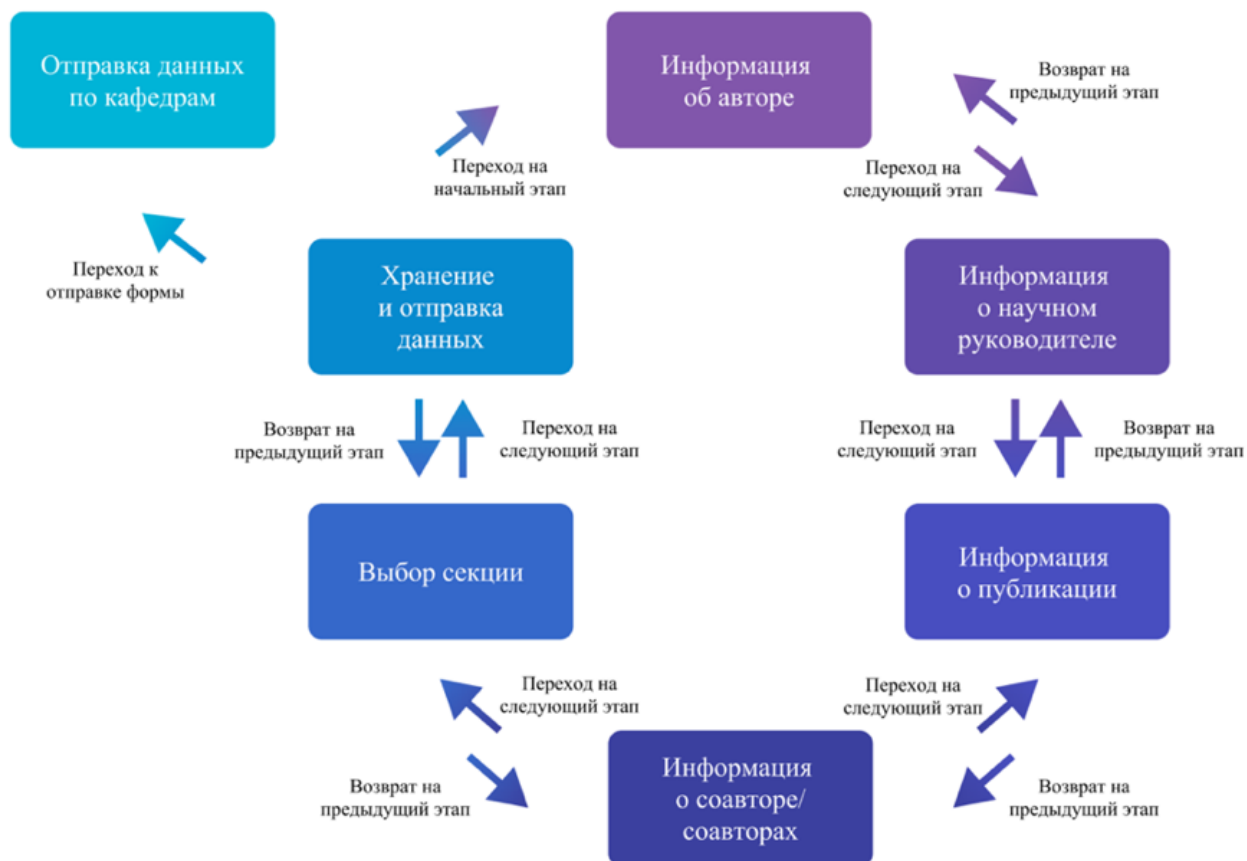


Рисунок 1. Принцип построения электронной формы регистрации

Разработанный авторами принцип построения электронной формы регистрации участников конференции включает в себя следующие блоки:

1. **Информация об авторе.** Данный блок должен включать в себя основную информацию об авторе, например, фамилия, имя, отчество, учреждение образования, степень образования и т.п. Должна быть предусмотрена возможность регистрации нескольких авторов (как правило не более трех).

2. **Информация о научном руководителе/руководителях.** Данный блок необходим для сбора информации о руководителе(ях) научной работы. Сведения включают фамилию, имя, отчество, ученую степень/звание и должность.

3. **Информация о публикации.** В данном разделе может содержаться следующая информация: название научной работы, форма участия, а также информация о докладчике.

4. **Информация о соавторе/соавторах.** В данный раздел можно включить следующую информацию: фамилия, имя, отчество, степень обучения, номер группы.

5. **Выбор секции.** Данный блок включает в себя основные сведения о каждой секции конференции.

6. **Хранение и отправки данных.** Данный раздел включает в себя научную работу, а также согласие на обработку данных.

На каждом этапе регистрации возможен возврат к предыдущему этапу и переход на следующий. На последнем этапе в случае несогласия на обработку персональных данных происходит возврат на первый этап.

Алгоритмы построения разделов электронной формы регистрации.

На рисунках 2–8 представлены алгоритмы всех разделов электронной формы регистрации. Для начала были созданы ссылки на документы (рисунок 2):

- информационное письмо, включающее в себя основные положения по приему материалов и проведению конференции;
- шаблон оформленной статьи;
- сокращение ученых степеней и званий;
- правила оформления использованных источников.



Рисунок 2. Алгоритм построения раздела «Общая информация» электронной формы регистрации

Далее был **создан раздел** для сбора информации об авторе, было принято включить в данный раздел следующую информацию (рисунок 3):

- фамилию, имя, отчество автора;
- страна;
- город;
- учреждения образования;
- кафедра, на которой обучается автор;
- степень обучения;
- номер группы;
- электронная почта.

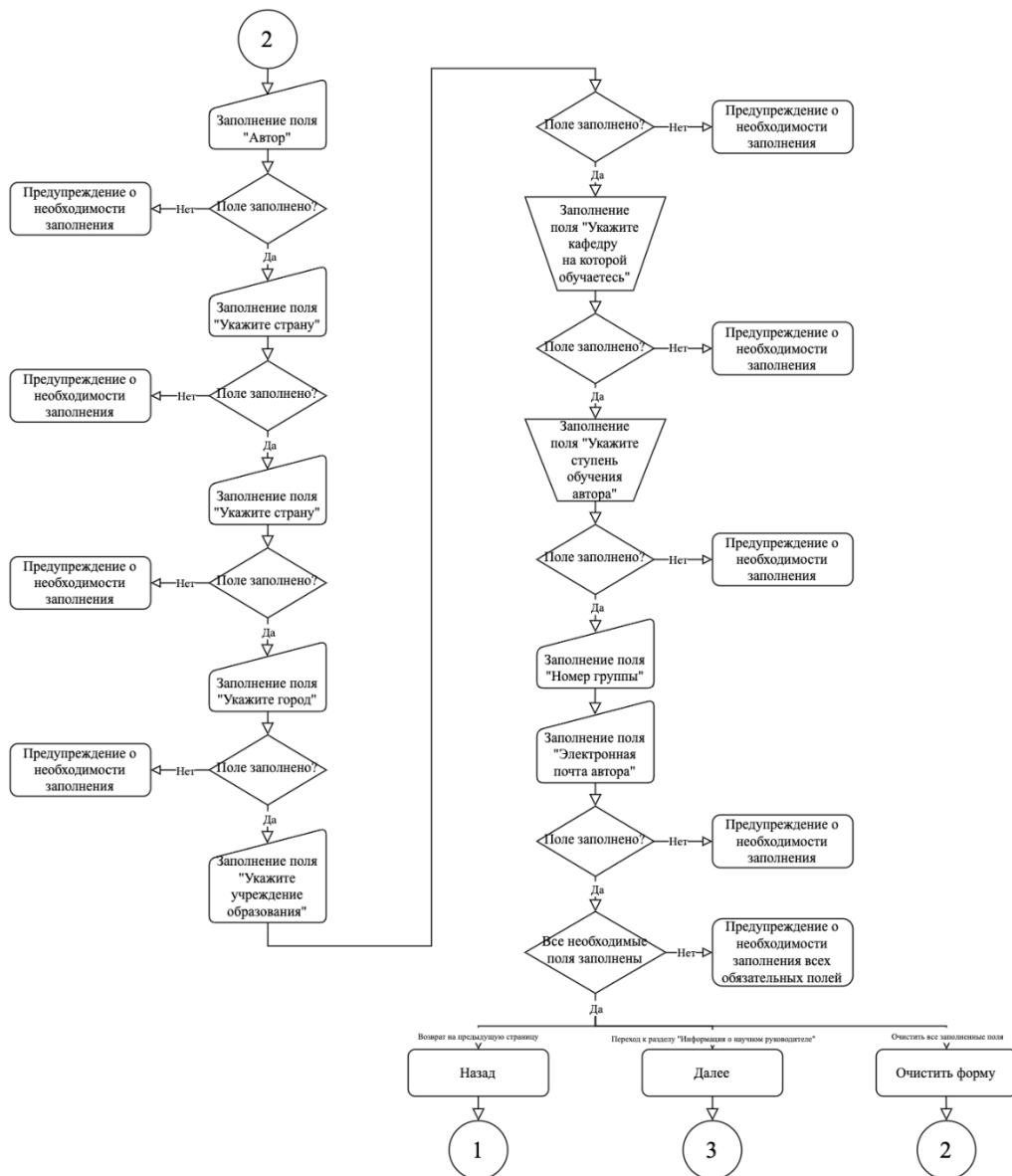


Рисунок 3. Алгоритм построения раздела «Информация об авторе» электронной формы регистрации

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

- научный руководитель (фамилия, имя, отчество);
- ученая степень;
- ученое звание;
- должность;
- вопрос о втором научном руководителе для перехода к следующему разделу:
 - а) если руководитель один, то будет осуществлен переход к разделу Информации о публикации;
 - б) если руководителей несколько, переход осуществляется к разделу следующего научного руководителя.

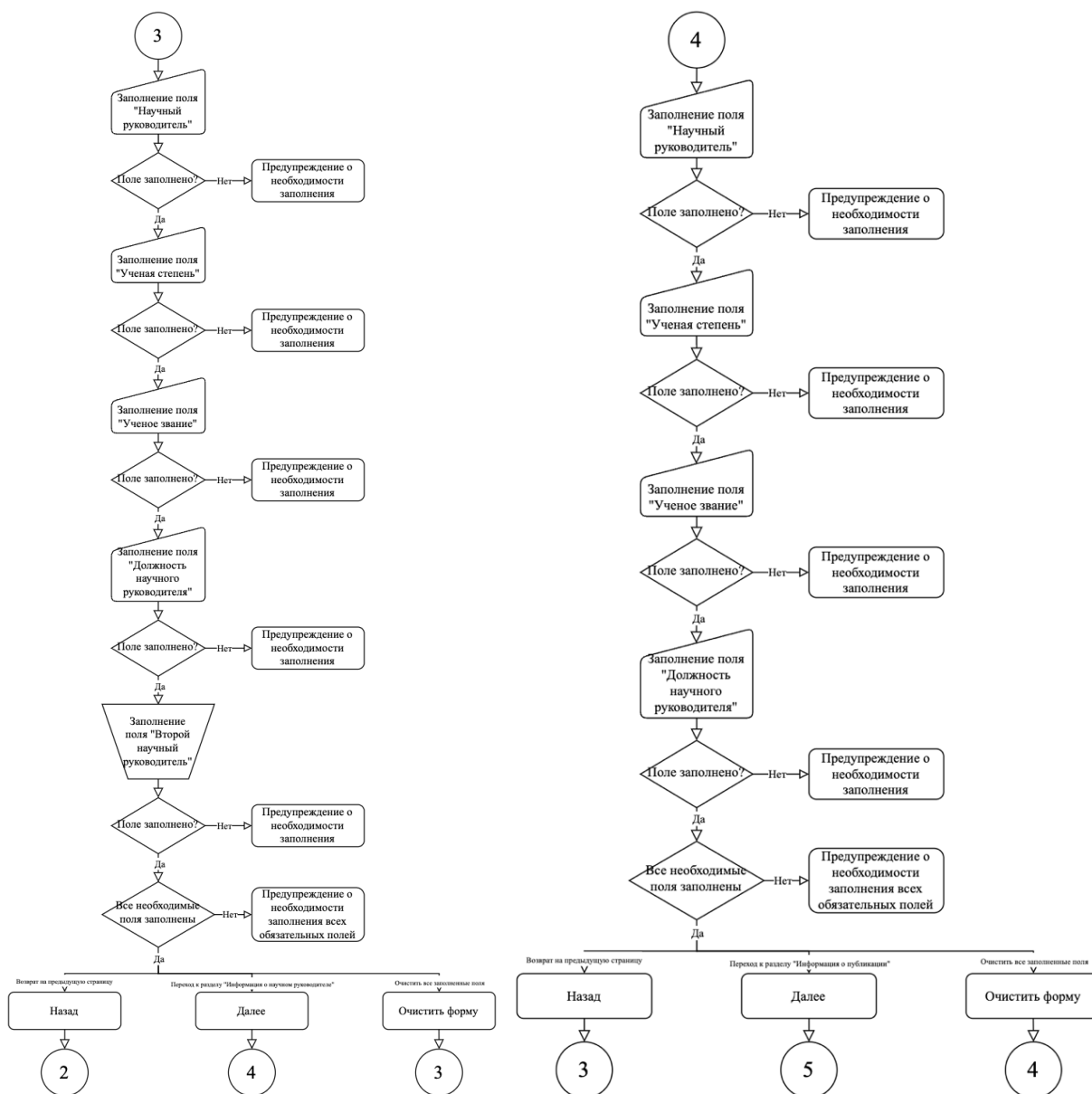


Рисунок 4. Алгоритм построения разделов «Информация о научном руководителе» электронной формы регистрации

Следующий раздел включает в себя информацию о публикации (рисунок 5):

- название статьи (утвержденное научным руководителем);
- форма участия;
- вопрос о соавторстве для перехода к следующему разделу:
 - а) если соавторов нет, то будет осуществлен переход к разделу выбора секции;
 - б) если соавторы есть, переход осуществляется к разделу соавтора.

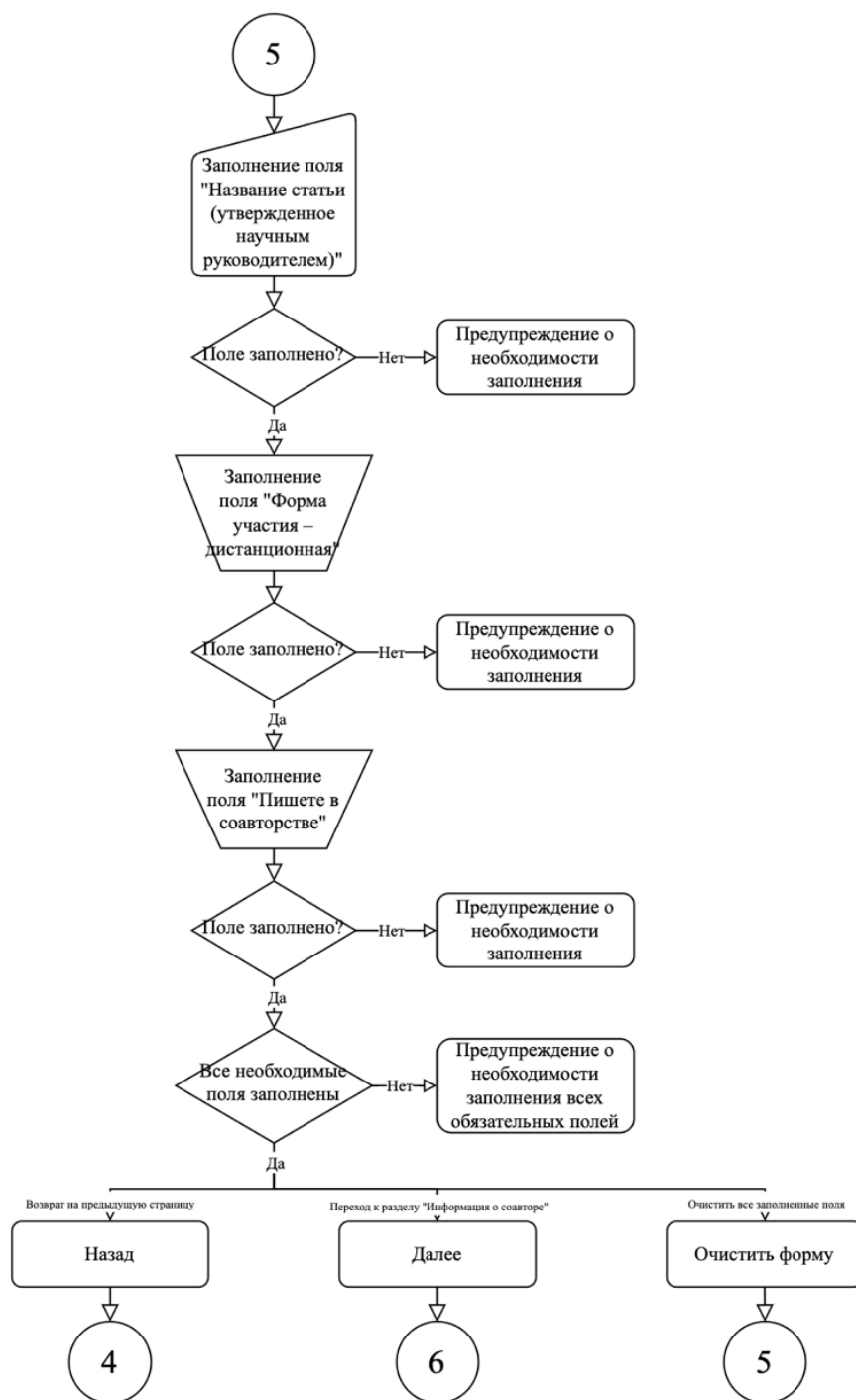


Рисунок 5. Алгоритм построения раздела «Информация о публикации» электронной формы регистрации

Далее идут два раздела для внесения информации о соавторе (рисунок 6) данный раздел включает в себя часть информации из раздела автора и имеет следующие позиции:

- соавтор (фамилия, имя, отчество);
- степень обучения;
- номер группы;
- вопрос о следующем соавторе для перехода к следующему разделу:
 - а) если соавторов нет, то будет осуществлен переход к разделу выбора секции;

б) если соавторы есть, переход осуществляется к разделу следующего соавтора.

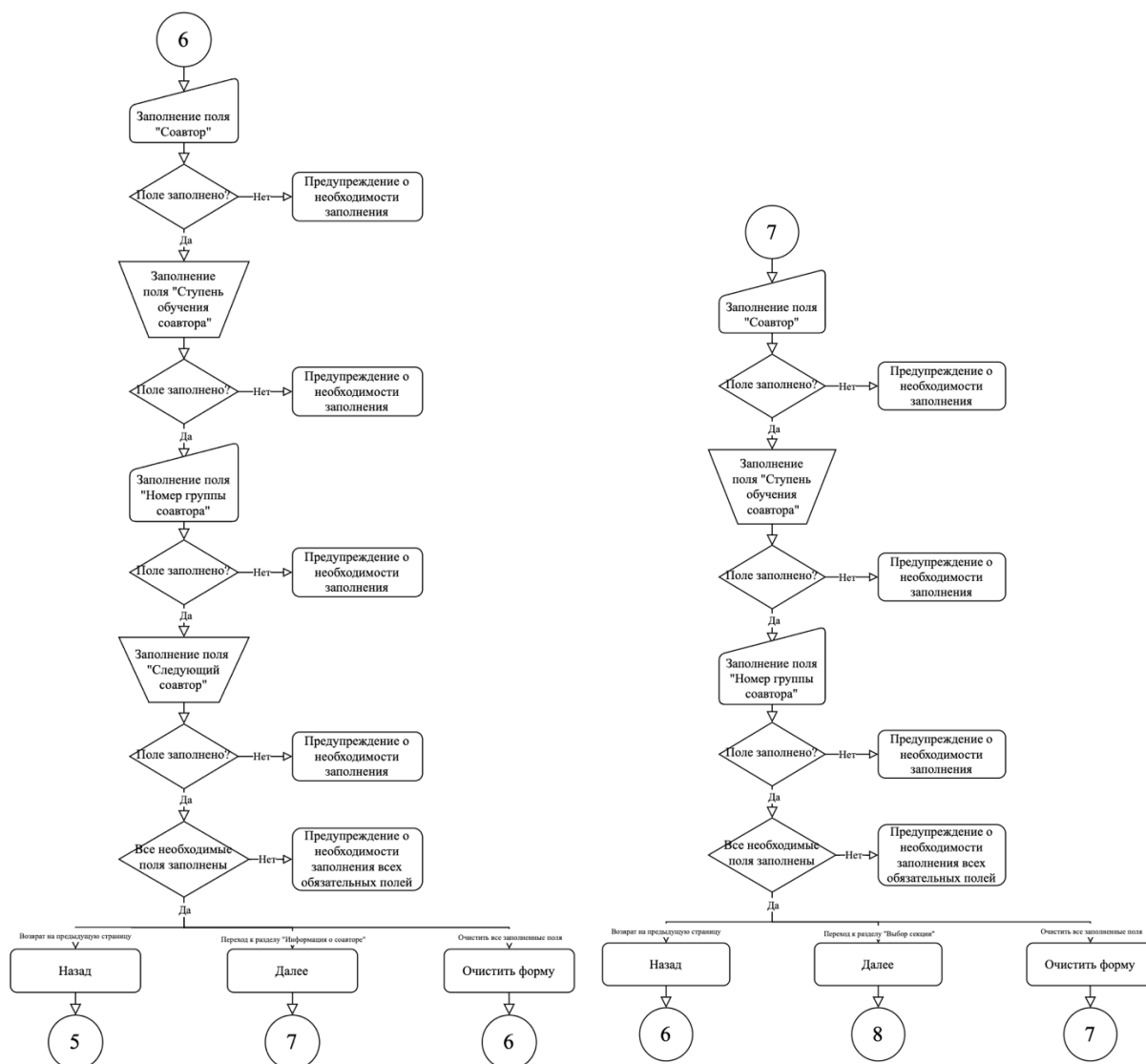


Рисунок 6. Алгоритм построения разделов «Информация о соавторе» электронной формы регистрации

После внесения информации о соавторе идет **раздел для выбора секции**, который включает в себя (рисунок 7):

- основную информацию о секциях и их направлениях;
- выбор секции, в данном поле присутствует пять позиций на выбор (a, b, c, d, e):
 - a) при выборе секции a, происходит перенаправление на электронную почту 1;
 - b) при выборе секции b, происходит перенаправление на электронную почту 2;
 - c) при выборе секции c, происходит перенаправление на электронную почту 3;
 - d) при выборе секции d, происходит перенаправление на электронную почту 4;
 - e) при выборе секции e, происходит перенаправление на электронную почту 5.

Данный раздел имеет непосредственную связь со следующим (последним) разделом.

Последний раздел разработанной электронной формы регистрации предназначен для **хранения и отправки данных**, и содержит следующие поля (рисунок 8):

- файл утвержденной научным руководителем статьи (формата .doc/.docx);

- поле согласия на хранение персональных данных, для перехода к стадии отправки данных:
 - а) Если участник подтверждает согласие, то осуществляется переход на стадию отправки формы;
 - б) Если же участник не подтверждает согласие, то осуществляется переход в самое начало электронной регистрации.

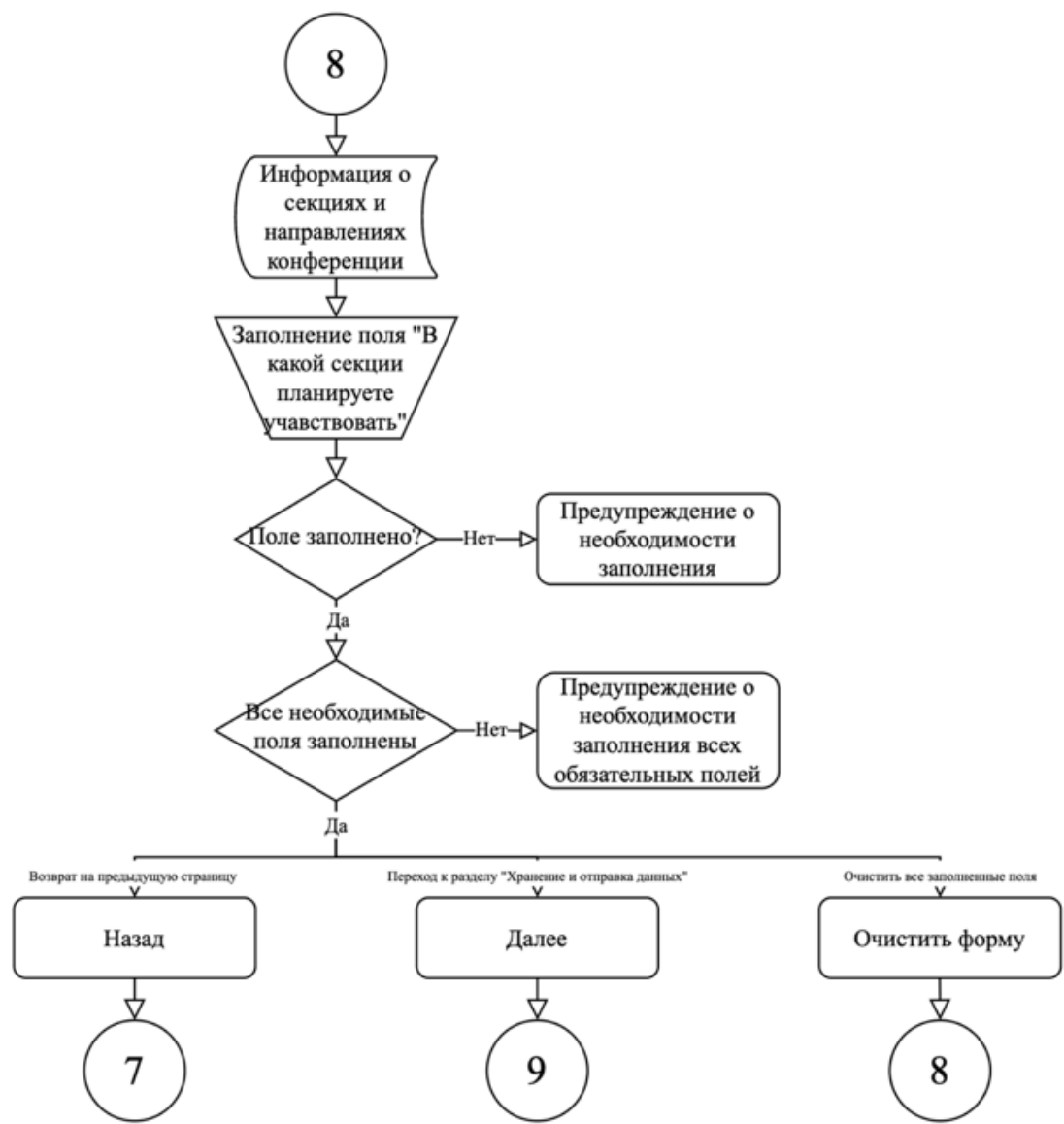


Рисунок 7. Алгоритм построения раздела «Информация о секции» электронной формы регистрации



Рисунок 8. Алгоритм построения раздела «Хранение и отправка данных» электронной формы регистрации

Данный алгоритм электронной формы регистрации был разработан для проведения научной конференции аспирантов, студентов и магистрантов. Поля данной формы могут быть оптимизированы для проведения любой научной конференции в зависимости от необходимой информации.

Заключение.

Предложен принцип построения электронной формы регистрации, реализованный в Google формы. Разработаны алгоритмы организации студенческой научной конференции. Определены обязательные и второстепенные маркеры участников. Реализована возможность валидации указанных данных. Имеется возможность формирования удаленной базы данных с определением формы и ступени обучения, а также данных о научном руководителе, учреждении высшего образования, кафедре и секции, на которой будет заслушиваться выступление.

Список литературы

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

- [2] Батура, М. П. Совершенствование организационной структуры управления научно-исследовательской работой студентов и магистрантов / М. П. Батура, В. Ф. Алексеев, А. П. Кузнецов // Известия Белорусской инженерной академии. – Минск, 2004. – № 1 (17/4). – С.6–9.
- [3] Достанко, А. П. Комплексное планирование НИРС на выпускающей кафедре технического университета / А. П. Достанко, В. Ф. Алексеев, С. В. Бордусов // Навучальна-даследчы прынцып у арганізацыі ўніверсітэцкай адукацыі : зборнік навуковых прац. – Мінск : Белдзяржуніверсітэт імя Максіма Танка, 1998. – С. 11-16.
- [4] Алексеев, В. Ф. Проект программы курса «Основы научных исследований и инженерного творчества» / В. Ф. Алексеев, С. В. Бордусов // Рэалізацыя навучальна-даследчага прынцыпу у сістэме шматўзроўневай ўніверсітэцкай адукацыі : 3б. навук.-метад. матэрыялаў / Пад рэд. Ю.А.Быкадорава – Мінск : БДПУ імя М.Танка, 2000. – С.139-150.
- [5] Батура, М.П. Анализ научно-исследовательской работы студентов учреждения образования «Белорусский государственный университет информатики и радиоэлектроники» / М. П. Батура, В. Ф. Алексеев // Известия Белорусской инженерной академии. – № 1(13)/1' 2002 – с. 6-22.
- [6] Алексеев, В. Ф. Конкретизация образовательных целей как основа учебного процесса / В. Ф. Алексеев, Л. С. Алексеева // Дистанционное обучение – образовательная среда 21 века : материалы II Международной научно-методической конференции, Минск, 26-28 ноября 2001 г. / Бестпринт. – Минск, 2002. – С. 398–400.
- [7] Алексеев, В. Ф. Роль профилирующей кафедры в подготовке высококвалифицированных специалистов / В. Ф. Алексеев // Опыт и проблемы организации научно-исследовательской работы студентов: Сборник научных статей. – Мн.: БГУИР, 2003. – С. 18-21.
- [8] Алексеев, В. Ф. Проблемы и возможные пути их реализации в работе с перспективными выпускниками по привлечению к научным исследованиям / В. Ф. Алексеев, Д. В. Лихачевский, Г. А. Пискун // Высшее техническое образование : проблемы и пути развития = Engineering education: challenges and developments : материалы IX Международной научно-методической конференции, Минск, 1-2 ноября 2018 года / редкол. : В. А. Богуш [и др.]. – Минск : БГУИР, 2018. – С. 9 – 14.
- [9] Алексеев, В. Ф. Научно-методологические основы организации научно-исследовательской работы студентов в высших учебных заведениях / В. Ф. Алексеев // Образовательные технологии в подготовке специалистов: Сб. научных статей. В 5 ч. (По итогам работы МНПК, Минск, 20-21 марта 2003 г.) / Учреждение образования «Минский государственный высший радиотехнический колледж». Под. ред. проф. Н. А. Цырельчука. – Минск: МГВРК, 2003. – Ч.1. – С. 24-39.
- [10] Пискун, Г. А. Особенности применения активных методов обучения при подготовке специалистов в техническом вузе / А. Г. Пискун, В. Ф. Алексеев // Дистанционное обучение – образовательная среда XXI века : материалы IX международной научно-методической конференции (Минск, 3-4 декабря 2015 года). – Минск : БГУИР, 2015. – С. 83 – 84.
- [11] Алексеев, В. Ф. Методология обучения проектированию электронных систем / В. Ф. Алексеев, Д. В. Лихачевский, Г. А. Пискун // Дистанционное обучение – образовательная среда XXI века : материалы XI Международной научно-методической конференции, Минск, 12-13 декабря 2019 г. / редкол. : В. А. Прытков [и др.]. – Минск : БГУИР, 2019. – С. 48-49.
- [12] Алексеев, В. Ф. Инженерное творчество в системе многоуровневого университетского образования / В. Ф. Алексеев, Д. В. Лихачевский, Г. А. Пискун // Дистанционное обучение – образовательная среда XXI века : материалы X международной научно-методической конференции (Минск, 7 - 8 декабря 2017 года). – Минск : БГУИР, 2017. – С. 124 - 125.
- [13] Пискун, Г. А. Оптимизация процесса проведения лабораторных занятий по дисциплине «Физические основы проектирования радиоэлектронных средств» / Г. А. Пискун, В. Ф. Алексеев // Дистанционное обучение – образовательная среда XXI века : материалы IX международной научно-методической конференции (Минск, 3-4 декабря 2015 года). – Минск : БГУИР, 2015. – С. 136 – 138.
- [14] Алексеева, Л. С. Дидактическая специфика деятельности преподавателей и студентов в процессе дистанционного обучения / Л. С. Алексеева, В. Ф. Алексеев, Г. А. Пискун // Дистанционное обучение – образовательная среда XXI века: материалы VIII международной научно-методической конференции. (Минск, 5–6 декабря 2013 года). – Минск: БГУИР, 2013. – С. 59 - 60.
- [15] Асадчая, Э. В. Принцип организации студенческих научных конференций с учетом covid-ограничений / Э. В. Асадчая // Электронные системы и технологии [Электронное издание] : сборник материалов 59-й научной конференции аспирантов, магистрантов и студентов БГУИР, Минск, 17-21 апреля 2023 г. / Белорусский государственный университет информатики и радиоэлектроники ; редкол.: Д. В. Лихачевский [и др.]. – Минск, 2023. – С. 361–362.
- [16] Асадчая, Э. В. Организация научно-исследовательской конференции студентов на кафедре учреждения высшего образования / Э. В. Асадчая, А. О. Царькова // Электронные системы и технологии [Электронное издание] : сборник материалов 59-й научной конференции аспирантов, магистрантов и студентов БГУИР, Минск, 17-21 апреля 2023 г. / Белорусский государственный университет информатики и радиоэлектроники ; редкол.: Д. В. Лихачевский [и др.]. – Минск, 2023. – С. 363–365.

**ORGANIZATION OF SECTIONS OF THE ELECTRONIC FORM
FOR REGISTRATION OF PARTICIPANTS OF A SCIENTIFIC CONFERENCE
WITH THE FORMATION OF ONLINE DATABASE**

V.F. Alexseev

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

G.A. Piskun

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

D.V. Likhachevsky

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

I.N. Tonkovich

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

E.V. Asadchaya

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

A.D. Larkin

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

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

Abstract. The principle of constructing an electronic registration form, which includes 6 main blocks of data, is presented. Algorithms for organizing these blocks for a student scientific conference have been developed, in which mandatory and secondary markers of participants are determined. The possibility of validation of the specified data has been implemented. The possibility of forming a remote database with the definition of the form and level of education, as well as data on the supervisor, institution of higher education, department and section of participation is shown.

Keywords: research work of students, research potential, educational process, construction principle, registration algorithm, scientific conference.

Научное издание

BIG DATA И АНАЛИЗ ВЫСОКОГО УРОВНЯ

BIG DATA AND ADVANCED ANALYTICS

СБОРНИК НАУЧНЫХ СТАТЕЙ
IX МЕЖДУНАРОДНОЙ НАУЧНО-ПРАКТИЧЕСКОЙ КОНФЕРЕНЦИИ

В двух частях

Часть 1

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

В авторской редакции
Ответственный за выпуск *Д. В. Лихачевский*
Компьютерная верстка *И. В. Андриалович*

Подписано в печать 12.05.2023. Формат 60×84 1/8. Бумага офсетная. Гарнитура «Таймс».
Цифровая печать. Усл. печ. л. 56,5. Уч.-изд. л. 53,0. Тираж 25 экз. Заказ 86.

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

