

ISSN 2713-3192  
DOI 10.15622/ia.2025.24.2  
<http://ia.spcras.ru>

ТОМ 24 № 2

**ИНФОРМАТИКА  
И АВТОМАТИЗАЦИЯ**

**INFORMATICS  
AND AUTOMATION**



**СПб ФИЦ РАН**

**Санкт-Петербург  
2025**



# INFORMATICS AND AUTOMATION

Volume 24 № 2, 2025

Scientific and educational journal primarily specialized in computer science, automation, robotics, applied mathematics, interdisciplinary research

Founded in 2002

---

## Founder and Publisher

St. Petersburg Federal Research Center of the Russian Academy of Sciences (SPC RAS)

---

## Editor-in-Chief

**A. L. Ronzhin**, Prof., Dr. Sci., St. Petersburg, Russia

---

## Editorial Council

<b>A. A. Ashimov</b>	Prof., Dr. Sci., Academician of the National Academy of Sciences of the Republic of Kazakhstan, Almaty, Kazakhstan
<b>I. A. Kalyaev</b>	Prof., Dr. Sci., Academician of RAS, Taganrog, Russia
<b>Yu. A. Merkurьев</b>	Prof., Dr. Sci., Academician of the Latvian Academy of Sciences, Riga, Latvia
<b>A. I. Rudskoi</b>	Prof., Dr. Sci., Academician of RAS, St. Petersburg, Russia
<b>V. Sgurev</b>	Prof., Dr. Sci., Academician of the Bulgarian Academy of Sciences, Sofia, Bulgaria
<b>B. Ya. Sovetov</b>	Prof., Dr. Sci., Academician of RAE, St. Petersburg, Russia
<b>V. A. Soyfer</b>	Prof., Dr. Sci., Academician of RAS, Samara, Russia

## Editorial Board

<b>E. Azarov</b>	Prof., Dr. Sci., Minsk, Belarus
<b>O. Yu. Gusikhin</b>	Ph. D., Dearborn, USA
<b>V. Delic</b>	Prof., Dr. Sci., Novi Sad, Serbia
<b>A. Dolgui</b>	Prof., Dr. Sci., St. Etienne, France
<b>M. N. Favorskaya</b>	Prof., Dr. Sci., Krasnoyarsk, Russia
<b>M. Zelezny</b>	Assoc. Prof., Ph.D., Plzen, Czech Republic
<b>H. Kaya</b>	Assoc. Prof., Ph.D., Utrecht, Netherlands
<b>A. A. Karpov</b>	Assoc. Prof., Dr. Sci., St. Petersburg, Russia
<b>S. V. Kuleshov</b>	Dr. Sci., St. Petersburg, Russia
<b>A. D. Khomonenko</b>	Prof., Dr. Sci., St. Petersburg, Russia
<b>D. A. Ivanov</b>	Prof., Dr. Habil., Berlin, Germany
<b>K. P. Markov</b>	Assoc. Prof., Ph.D., Aizu, Japan
<b>R. V. Meshcheryakov</b>	Prof., Dr. Sci., Moscow, Russia
<b>N. A. Moldovian</b>	Prof., Dr. Sci., St. Petersburg, Russia
<b>V. V. Nikulin</b>	Prof., Ph.D., New York, United States
<b>V. Yu. Osipov</b>	Prof., Dr. Sci., Deputy Editor-in-Chief, St. Petersburg, Russia
<b>V. K. Pshikhopov</b>	Prof., Dr. Sci., Taganrog, Russia
<b>H. Samani</b>	Assoc. Prof., Ph.D., Plymouth, UK
<b>J. Savage</b>	Assoc. Prof., Ph.D., Mexico City, Mexico
<b>M. Secujski</b>	Assoc. Prof., Ph.D., Novi Sad, Serbia
<b>A. V. Smirnov</b>	Prof., Dr. Sci., St. Petersburg, Russia
<b>B. V. Sokolov</b>	Prof., Dr. Sci., St. Petersburg, Russia
<b>L. V. Utkin</b>	Prof., Dr. Sci., St. Petersburg, Russia
<b>L. B. Sheremetov</b>	Assoc. Prof., Dr. Sci., Mexico, Mexico

---

**Editor:** A. S. Viktorova

**Interpreter:** Ya. N. Berezina

**Art editor:** N. A. Dormidontova

---

## Editorial office address

SPC RAS, 39 litera A, 14-th line V.O., St. Petersburg, 199178, Russia

e-mail: [ia@sprcras.ru](mailto:ia@sprcras.ru), web: <http://ia.sprcras.ru>

## The journal is indexed in Scopus

The journal is published under the scientific-methodological supervision of Department for Nanotechnologies and Information Technologies of the Russian Academy of Sciences

© St. Petersburg Federal Research Center of the Russian Academy of Sciences, 2025

# ИНФОРМАТИКА И АВТОМАТИЗАЦИЯ

Том 24 № 2, 2025

Научный, научно-образовательный журнал с базовой специализацией  
в области информатики, автоматизации, робототехники, прикладной математики  
и междисциплинарных исследований.

Журнал основан в 2002 году

## Учредитель и издатель

Федеральное государственное бюджетное учреждение науки  
«Санкт-Петербургский Федеральный исследовательский центр Российской академии наук»  
(СПб ФИЦ РАН)

## Главный редактор

А. Л. Ронжин, д-р техн. наук, проф., Санкт-Петербург, РФ

## Редакционный совет

А. А. Ашимов	академик Национальной академии наук Республики Казахстан, д-р техн. наук, проф., Алматы, Казахстан
И. А. Каляев	академик РАН, д-р техн. наук, проф., Таганрог, РФ
Ю. А. Меркурьев	академик Латвийской академии наук, д-р, проф., Рига, Латвия
А. И. Рудской	академик РАН, д-р техн. наук, проф., Санкт-Петербург, РФ
В. Сгурев	академик Болгарской академии наук, д-р техн. наук, проф., София, Болгария
Б. Я. Советов	академик РАО, д-р техн. наук, проф., Санкт-Петербург, РФ
В. А. Соифер	академик РАН, д-р техн. наук, проф., Самара, РФ

## Редакционная коллегия

И. С. Азаров	д-р техн. наук, проф., Минск, Беларусь
О. Ю. Гусихин	д-р наук, Диаборн, США
В. Делич	д-р техн. наук, проф., Нови-Сад, Сербия
А. Б. Долгий	д-р наук, проф. Сент-Этьен, Франция
М. Железны	д-р наук, доцент, Пльзень, Чешская республика
Д. А. Иванов	д-р экон. наук, проф., Берлин, Германия
Х. Кайя	д-р наук, доцент, Утрехт, Нидерланды
А. А. Карпов	д-р техн. наук, доцент, Санкт-Петербург, РФ
С. В. Кулешов	д-р техн. наук, Санкт-Петербург, РФ
К. П. Марков	д-р наук, доцент, Аизу, Япония
Р. В. Мещеряков	д-р техн. наук, проф., Москва, РФ
Н. А. Молдовян	д-р техн. наук, проф., Санкт-Петербург, РФ
В.В. Никулин	д-р наук, проф., Нью-Йорк, США
В.Ю. Осипов	д-р техн. наук, проф., зам. главного редактора, Санкт-Петербург, РФ
В. Х. Пшихопов	д-р техн. наук, проф., Таганрог, РФ
Х. К. Саваж	д-р техн. наук, доцент, Мехико, Мексика
Х. Самани	д-р наук, доцент, Плимут, Соединённое Королевство
М. Сечуиски	д-р техн. наук, доцент, Нови-Сад, Сербия
А. В. Смирнов	д-р техн. наук, проф., Санкт-Петербург, РФ
Б. В. Соколов	д-р техн. наук, проф., Санкт-Петербург, РФ
Л. В. Уткин	д-р техн. наук, проф., Санкт-Петербург, РФ
М. Н. Фаворская	д-р техн. наук, проф., Красноярск, РФ
А. Д. Хомоненко	д-р техн. наук, проф., Санкт-Петербург, РФ
Л. Б. Шереметов	д-р техн. наук, Мехико, Мексика

Выпускающий редактор: А. С. Викторова

Переводчик: Я. Н. Березина

Художественный редактор: Н. А. Дормидонтова

## Адрес редакции

14-я линия В.О., д. 39, лит. А, г. Санкт-Петербург, 199178, Россия

e-mail: [ia@spcras.ru](mailto:ia@spcras.ru), сайт: <http://ia.spcras.ru>

## Журнал индексируется в международной базе данных Scopus

Журнал входит в «Перечень ведущих рецензируемых научных журналов и изданий, в которых должны быть опубликованы основные научные результаты диссертации на соискание ученой степени доктора и кандидата наук»

Журнал выпускается при научно-методическом руководстве Отделения нанотехнологий и информационных технологий Российской академии наук

© Федеральное государственное бюджетное учреждение науки

«Санкт-Петербургский Федеральный исследовательский центр Российской академии наук», 2025  
Разрешается воспроизведение в прессе, а также сообщение в эфир или по кабелю опубликованных в составе печатного периодического издания - журнала «ИНФОРМАТИКА И АВТОМАТИЗАЦИЯ» статей по текущим экономическим, политическим, социальным и религиозным вопросам с обязательным указанием имени автора статьи и печатного периодического издания журнала «ИНФОРМАТИКА И АВТОМАТИЗАЦИЯ»

## CONTENTS

### **Mathematical Modeling and Applied Mathematics**

I. Lebedev

ADAPTIVE REGRESSION MODEL CONSTRUCTION BASED ON THE FUNCTIONAL  
QUALITY ANALYSIS OF THE SEQUENCE SEGMENT PROCESSING 363

K. Krotov

TWO-LEVEL OPTIMIZATION OF TASK DISTRIBUTION INTO BATCHES AND  
SCHEDULING THEIR EXECUTION IN PIPELINE SYSTEMS WITH LIMITED  
BUFFERS 395

I. Chizhov

AN AUDIOCODEC BASED ON THE PERCEPTUAL EQUALITY BETWEEN THE  
ORIGINAL AND RESTORED AUDIO SIGNALS 428

V. Zinov, V. Kartak, Y. Valiakhmetova

SOLVING MULTI-OBJECTIVE RATIONAL PLACEMENT OF LOAD-BEARING  
WALLS PROBLEM VIA GENETIC ALGORITHM 464

V. Kureychik, V. Danilchenko, E. Danilchenko

ROUTING OF AUTONOMOUS DEVICES IN THREE-DIMENSIONAL SPACE 492

D. Nagibin, A. Petrenko, V. Davydenko, I. Kotenko, E. Fedorchenko

INVASIVE APPROACH TO VERIFICATION OF FUNCTIONAL AND STRUCTURAL  
SPECIFICATIONS IMPLEMENTED IN CUSTOM INTEGRATED CIRCUITS 526

### **Artificial Intelligence, Knowledge and Data Engineering**

Nguyen Viet Hung, Nguyen Anh Quan, Nguyen Tan, Tran Trieu Hai, Dang Khanh Trung, Le  
Mai Nam, Bui Thanh Loan, Nguyen Thi Thuy Nga

BUILDING PREDICTIVE SMELL MODELS FOR VIRTUAL REALITY  
ENVIRONMENTS 556

M. Idrissi Alami, A. Ez-zahout, F. Omary

ENHANCED PEOPLE RE-IDENTIFICATION IN CCTV SURVEILLANCE USING DEEP  
LEARNING: A FRAMEWORK FOR REAL-WORLD APPLICATIONS 583

I. Kipyatkova, I. Kagirov, M. Dolgushin

USE OF PRE-TRAINED MULTILINGUAL MODELS FOR KARELIAN SPEECH  
RECOGNITION 604

A. Naveen, I.J. Jacob, A.K. Mandava

DETECTION OF STUDENT ENGAGEMENT VIA TRANSFORMER-ENHANCED  
FEATURE PYRAMID NETWORKS ON CHANNEL-SPATIAL ATTENTION 631

A. Chuvakov, R. Boryaev

FACTORIZING DECISION SUPPORT SYSTEM BASED ON OPTIMIZED QUANTUM  
ALGORITHMS QMC 657

A. Ivaschenko, O. Golovnin, A. Golovnina, E. Dodonova

AN ASSEMBLED MODEL OF MULTILAYER GEOINFORMATION SPACE-TIME 684



## СОДЕРЖАНИЕ

### Математическое моделирование и прикладная математика

И.С. Лебедев  
АДАПТИВНОЕ ПОСТРОЕНИЕ РЕГРЕССИОННЫХ МОДЕЛЕЙ НА ОСНОВЕ  
АНАЛИЗА ФУНКЦИОНАЛА КАЧЕСТВА ОБРАБОТКИ СЕГМЕНТОВ  
ПОСЛЕДОВАТЕЛЬНОСТИ 363

К.В. Кротов  
ДВУХУРОВНЕВАЯ ОПТИМИЗАЦИЯ РАСПРЕДЕЛЕНИЯ ЗАДАНИЙ ПО ПАКЕТАМ  
И РАСПИСАНИЙ ИХ ВЫПОЛНЕНИЯ В КОНВЕЙЕРНЫХ СИСТЕМАХ  
С БУФЕРАМИ ОГРАНИЧЕННЫХ РАЗМЕРОВ 395

И.И. Чижов  
АУДИОКОДЕК НА ОСНОВЕ ПЕРЦЕПТУАЛЬНОГО РАВЕНСТВА ИСХОДНОГО  
И ВОССТАНОВЛЕННОГО ЗВУКОВОГО СИГНАЛА 428

В.И. Зинов, В.М. Картак, Ю.И. Валиахметова  
РЕШЕНИЕ МНОГОКРИТЕРИАЛЬНОЙ ЗАДАЧИ РАЦИОНАЛЬНОГО РАЗМЕЩЕНИЯ  
НЕСУЩИХ СТЕН С ПОМОЩЬЮ ГЕНЕТИЧЕСКОГО АЛГОРИТМА 464

В.В. Курейчик, В.И. Данильченко, Е.В. Данильченко  
МАРШРУТИЗАЦИЯ АВТОНОМНЫХ УСТРОЙСТВ В ТРЁХМЕРНОМ  
ПРОСТРАНСТВЕ 492

Д.В. Нагибин, А.С. Петренко, В.С. Давыденко, И.В. Котенко, Е.В. Федорченко  
ИНВАЗИВНЫЙ ПОДХОД К ВЕРИФИКАЦИИ ФУНКЦИОНАЛЬНО-СТРУКТУРНЫХ  
СПЕЦИФИКАЦИЙ, РЕАЛИЗОВАННЫХ В ЗАКАЗНЫХ ИНТЕГРАЛЬНЫХ СХЕМАХ 526

### Искусственный интеллект, инженерия данных и знаний

Н.В. Хунг, Н.А. Куан, Н. Тан, Т.Т. Хай, Д.Т. Чунг, Л.М. Нам, Б.Т. Лоан, Н.Т.Т. Нга  
СОЗДАНИЕ ПРОГНОЗИРУЮЩИХ МОДЕЛЕЙ ЗАПАХОВ ДЛЯ СРЕД  
ВИРТУАЛЬНОЙ РЕАЛЬНОСТИ 556

М. Идрисси Алами, А. Эз-захут, Ф. Омари  
УЛУЧШЕННАЯ ПОВТОРНАЯ ИДЕНТИФИКАЦИЯ ЛЮДЕЙ В СИСТЕМАХ  
ВИДЕОНАБЛЮДЕНИЯ С ИСПОЛЬЗОВАНИЕМ ГЛУБОКОГО ОБУЧЕНИЯ:  
СТРУКТУРА ДЛЯ РЕАЛЬНЫХ ПРИЛОЖЕНИЙ 583

И.С. Кипяткова, И.А. Кагиров, М.Д. Долгушин  
ПРИМЕНЕНИЕ ПРЕДВАРИТЕЛЬНО ОБУЧЕННЫХ МНОГОЯЗЫЧНЫХ МОДЕЛЕЙ  
ДЛЯ РАСПОЗНАВАНИЯ КАРЕЛЬСКОЙ РЕЧИ 604

А. Навин, И.Д. Джейкоб, А.К. Мандава  
ОПРЕДЕЛЕНИЕ ВОВЛЕЧЕННОСТИ УЧАЩИХСЯ С ПОМОЩЬЮ СЕТЕЙ  
ПИРАМИДАЛЬНЫХ ПРИЗНАКОВ, УЛУЧШЕННЫХ ТРАНСФОРМЕРОМ, С  
КАНАЛЬНО-ПРОСТРАНСТВЕННЫМ ВНИМАНИЕМ 631

А.В. Чуваков, Р.О. Боряев  
СИСТЕМА ПОДДЕРЖКИ ПРИНЯТИЯ ФАКТОРИНГОВЫХ РЕШЕНИЙ НА ОСНОВЕ  
ОПТИМИЗИРОВАННЫХ КВАНТОВЫХ АЛГОРИТМОВ QMC 657

А.В. Иващенко, О.К. Головнин, А.А. Головнина, Е.А. Додонова  
КОМБИНИРОВАННАЯ ГЕОИНФОРМАЦИОННАЯ МНОГОСЛОЙНАЯ  
ПРОСТРАНСТВЕННО-ВРЕМЕННАЯ МОДЕЛЬ 684

И.С. ЛЕБЕДЕВ

**АДАПТИВНОЕ ПОСТРОЕНИЕ РЕГРЕССИОННЫХ МОДЕЛЕЙ  
НА ОСНОВЕ АНАЛИЗА ФУНКЦИОНАЛА КАЧЕСТВА  
ОБРАБОТКИ СЕГМЕНТОВ ПОСЛЕДОВАТЕЛЬНОСТИ**

*Лебедев И.С. Адаптивное построение регрессионных моделей на основе анализа функционала качества обработки сегментов последовательности.*

**Аннотация.** Рассмотрена задача адаптивного построения модели, направленной на повышение показателей качества обработки информационных последовательностей. В методах обработки данных, которые нашли применение во многих прикладных областях, применяемый анализ объектов наблюдения является вычислительно ресурсоемким и в случае изменения свойств данных, требует большого количества итераций. В статье предложена методика выбора сегментов информационной последовательности, полученных разными способами, отличающаяся использованием функционала качества регрессионных моделей обработки подпоследовательностей. Поступающие на вход последовательности объектов наблюдения подвергаются разделению различными предварительно заданными алгоритмами сегментации. На каждом полученном сегменте обучаются заранее выбранные регрессионные модели и, в зависимости от полученных значений вычисленного функционала качества, происходит назначение лучших по качественным показателям моделей на сегменты. Это позволяет формировать агрегационную модель обработки данных. На основе эксперимента на модельных данных и выборках проведена оценка предлагаемой методики. Получены значения показателя качества MSE и MAE для разных алгоритмов обработки и при различном количестве сегментов. Предлагаемая методика дает возможность повысить показатели MSE и MAE за счет сегментации и назначения регрессионных моделей, которые имеют наилучшие показатели на отдельных сегментах. Предложенное решение направлено на дальнейшее усовершенствование ансамблевых методов. Его применение позволяет повысить оперативность настройки базовых алгоритмов в случае трансформации свойств данных и улучшить интерпретируемость результатов. Методика может применяться при разработке моделей и методов обработки информационных последовательностей.

**Ключевые слова:** машинное обучение, адаптивные модели, повышение качества обработки, регрессионные модели.

**1. Введение.** В условиях лавинообразного роста информационных потоков, вызванного внедрением средств сбора и накопления данных во всевозможные процессы, создание эффективных моделей при решении задач регрессии, предсказания поведения системы является проблемным вопросом. Анализ информационных последовательностей, полученных в ходе функционирования разнородных объектов наблюдения, является сложной задачей, поэтому в целях автоматизации действий, направленных на повышение качественных показателей обработки данных применяются саморегулирующиеся вычисления. Их относят к методам, позволяющим автоматически реагировать на динамически изменяющиеся свойства входных данных. Саморегулирующиеся

вычисления могут значительно повысить качественные показатели обработки входных регрессионных последовательностей [1].

Одним из часто используемых методов машинного обучения, применяемых для предсказания поведения системы, является линейная регрессия. В случае наличия в данных случайного статистического шума и близких к линейным зависимостей она позволяет получить достаточно качественное приближение, но при резких изменениях наблюдаемых значений такая модель обработки становится не адекватной. Решение обозначенного вопроса может быть выполнено на основе кусочно-линейной аппроксимации, предполагающей применение сегментированной регрессии. Происходит разделение последовательности данных на отдельные сегменты, а затем на каждом из них строится линейная регрессия [2]. В результате последовательность аппроксимируется набором прямых, ограниченных сегментами.

В случае «линейности» данных сегментированная линейная регрессия может обеспечить хорошую аппроксимацию. Однако для эффективной реализации необходимо решать ряд вопросов, связанных с выбором сегментов. Заданная длина последовательности, количество сегментов, выбросы и шумы, свойства данных могут существенно оказывать влияние на результат. Возникает большое количество возможных вариантов разбиения [3].

При обработке регрессионных зависимостей очень часто возникают ситуации, когда взаимосвязь между ответом и объясняющими переменными резко меняется в определенных точках [4]. Такие точки проявляются во многих областях. Например, в социологии различные срезы общественного мнения, в медицине риски заболеваний могут существенно отличаться от возраста людей, в энергетике генерация электроэнергии – зависеть от климатических, температурных условий, в анализе состояния информационных систем и сетей при определенных внешних воздействиях может проявляться резкий рост информационных, служебных сообщений. Для обнаружения точек, где взаимосвязь между переменными резко меняется, применяются методы на основе сплайнов и методы кусочной регрессии. В первом случае задача состоит в том, чтобы найти наиболее подходящую модель для прогнозирования или для поиска экстремальных значений [5 – 7], а во втором – кусочная регрессия направлена на выявление закономерностей между ответными и объясняющими переменными на интервалах ограниченных точками изменений.

Обнаружение точек изменения является одной из важных проблем для регрессионного анализа. Ее решение зависит от обработки шумов, выбросов, аномальных значений выборки данных. Возникает задача

анализа свойств данных в целях формирования оптимальных выборок для обучения моделей и их последующего использования. Кроме того, существует большое разнообразие методов и способов разбиения последовательностей, которые обуславливают необходимость применения анализа результатов вычислений для оценки и выбора способов, позволяющих достичь оптимальных качественных показателей.

**2. Обзор существующих методов.** Задачи прогнозирования, регрессии решаются с помощью различных методов машинного обучения. Одними из наиболее популярных моделей обработки являются нейронные сети. На сегодняшний день для решения задач используются архитектуры CNN, RNN, GAN, DNN, которые совершенствуются [8]. Постоянно происходит их развитие, разрабатываются новые нейросетевые модели, например KAN. Они обладают свойствами адаптивности, гибкости настройки на решение предметно ориентированных задач, имеют большие возможности по расширению и масштабируемости, позволяют обеспечить высокую скорость обработки данных. Однако при их использовании возникает ряд проблемных вопросов вычислительной сложности, ограничения памяти при обучении на больших наборах данных, связанных с "исчезающими" и "взрывными" градиентами, стабильностью обучения, необходимостью построения архитектуры обработки и хранения данных при решении масштабных задач информационного анализа. Несмотря на все преимущества, позволяющие вычислять скрытые закономерности, по-прежнему для обучения нейросети требуется большое количество данных, что не всегда возможно обеспечить.

В общем виде нейросеть можно рассматривать как суперпозицию функций, где роль функций выполняют нейроны и их совокупности [9 – 11]. В рамках такого представления появляется возможность в качестве альтернативы рассмотреть замену суперпозиции функций, основанной на нейронах, функциями – моделями обработки данных, например реализующими алгоритмы наивного Байеса, деревьев решений, опорных векторов или других нейросетевых структур, последовательность использования которых будет приводить к такому же решению задачи, как и с помощью нейросети, имеющей "сложное" построение.

Определяя правила последовательного применения менее сложных и менее ресурсоемких моделей по отдельности, выбора наиболее подходящих алгоритмов в зависимости от свойств данных и текущих процессов обработки возможно получить общую модель, обладающую лучшей интерпретируемостью результатов, меньшей ресурсоемкостью, и достигающую соизмеримых с нейросетью результатов. Однако выбор

последовательности обработки таких моделей является нетривиальной задачей, подразумевающей оптимизацию протекающих в них процессов.

Каждая из моделей оптимизируется под определенные свойства последовательностей данных. Причем процессы оптимизации методов осуществляются по двум основным направлениям. С одной стороны выполняется «повышение качества» обрабатываемых данных, а с другой – построение эффективной модели обработки [12].

В целях повышения качества обработки данных используются методы формирования пространства признаков. Среди них применительно к рассматриваемой проблематике можно выделить подходы, выполняющие разделение данных, на основе кластеризации, поиска точек разладки временных рядов, обнаружения «дрейфа концепта» при трансформации свойств данных. Проблемные вопросы нахождения точек, где изменяются свойства предикторов и целевой переменной, являются важными для решения многих прикладных задач, использующих сегментированную регрессию. В этих целях применяется большое количество методов, таких как байесовский анализ, метод максимального правдоподобия, квантильная регрессия, непараметрические методы, аналитические способы. Однако сложность аппроксимирующей функции, проходящей через такие точки, часто приводит к большим вычислительным и ресурсным затратам при ее поиске.

Процессы разделения выборки в методах машинного обучения многими исследователями рассматриваются как вспомогательные. В большинстве случаев им уделяют внимание в рамках решения специализированных задач. Они зависят от вида и свойств обрабатываемой информации. Применительно к обработке информационных последовательностей можно выделить ряд основных направлений. Например, в работе [13] используются классические методы кластеризации. Несмотря на их относительную простоту, они позволяют повысить качественные показатели обработки только данных с определенными свойствами. В условиях высокой размерности, когда может существовать более одной точки изменения возникает проблема их обнаружения и локализации. Эта задача может решаться методами сегментации данных, которые используют динамическое программирование [14–16], бинарную сегментацию, байесовские методы [17]. В работах [17–19] предложены байесовские подходы для регрессий по точкам изменения. Однако описанные в них решения имеют высокую вычислительную сложность, требуют большое количество итераций при моделировании цепей Маркова методом Монте-Карло. Работа [20] использует статистические пороги для сегментации последовательности. При таком решении для повышения точности

обработки необходимо учитывать выбросы и шумы объектов наблюдения. В [21] представлен метод сегментированной регрессии (SEG) с использованием линейного перехода для оценки точек изменения регрессионной модели. В [22] предложены методы глубокого обучения. Качественные показатели обработки данных в этом случае зависят от ресурсоемкости и анализа данных. В работе [23] для поиска сегментов применяется расширенный тест Дики-Фуллера (ADF) на стационарность данных. В [24] рассматривается метод, использующий шаблоны для определения сегментов временного ряда. В [25] предложен подход, базирующийся на поиске максимального правдоподобия, применяющий метод динамического программирования. В [26, 27] был реализован метод одновременной многомасштабной оценки точки изменения (SMUCE), определяющий не только сами точки, но и доверительные интервалы. В [28] для разбиения выборки предложен метод бинарной сегментации.

Характеристики рассмотренных методов, предпосылки применения и возникающие при этом сложности представлены в таблице 1.

Разделение выборки на сегменты позволяет определить внутреннюю структуру данных для дальнейшего анализа и обработки, исследовать вероятные связи между объектами наблюдений. В случае рассмотрения информационных последовательностей возникает необходимость применения алгоритмов разделения данных, обнаруживающих новые закономерности.

Второе направление связано с поиском наиболее эффективной модели обработки данных. В простейших задачах применяются базовые алгоритмы, например линейная, логистическая регрессия, метод машинных векторов, нейросетевые подходы. Достижимые ими значения показателей качества обработки зависят от свойств обрабатываемых выборок. Наличие выбросов, линейность данных, независимость переменных оказывают существенное влияние на показатели качества результата обработки.

В целях преодоления обозначенных проблемных вопросов для повышения качества обработки регрессионных последовательностей в [29] используется многомодельный подход, направленный на формирование ансамбля моделей и алгоритмов, сочетающий несколько методов машинного обучения. Могут использоваться системы принятия решений, основанные на правилах, априорных знаниях о данных, системы взвешенного, выборочного голосования, которые оценивают модели только на основе их прошлых результатов прогнозирования [30], каскады простых алгоритмов и глубоких нейронных сетей [31].

Таблица 1. Характеристики рассмотренных методов разделения последовательности

Методы разделения	Ссылка на источник	Характеристики	
		Предпосылки применения	Проблемы применения
Классические методы кластеризации	[13]	Имеют относительную простоту реализации.	Обладают сильной зависимостью от свойств данных.
Методы динамического программирования	[14 – 16], [25]	Применяются при не дифференцируемости целевой функции, дискретном изменении переменных.	Имеют большую ресурсоемкость.
Методы бинарной сегментации	[17], [28]	Простота реализации, высокая скорость обработки.	Большая вероятность ошибочной сегментации.
Байесовские методы	[18, 19]	Могут быть реализованы на выборках небольшого объема, используются для данных высокой размерности.	Характеризуются субъективностью предварительного описания, что может приводить к слабым качественным показателям.
Метод на основе статистических порогов для сегментации последовательности	[20]	Обладают высокой чувствительностью, имеют апробированный математический аппарат.	Качество результатов зависит от анализа и интерпретации данных, (необходимо определять и интерпретировать выбросы и шумы объектов наблюдения).
Метод сегментированной регрессии	[21]	Характеризуется простотой вычислительных алгоритмов.	При увеличении разрядности данных падает точность определения точек изменения, чувствителен к изменчивости данных (сдвигу дисперсии).
Методы глубокого обучения	[22]	Имеют возможность достижения высоких качественных показателей при адекватной модели.	Обладают зависимостью от исходных свойств последовательности, сложностью построения модели, ресурсоемкостью, требуют большое количество итераций.
Метод сегментирования с применением расширенного теста Дики-Фуллера (ADF)	[23]	Адаптирован для рядов, обладающих стационарностью.	В случае нестационарного ряда и необходимости проверки нескольких точек разладки существенно возрастает сложность.
Метод шаблонов для определения сегментов временного ряда	[24]	Использует множество метрик сходства и различия.	Зависит от качества шаблонов и выбранных критериев сходства и различия.
Метод одновременной многомасштабной оценки точки изменения (SMUCE)	[26, 27]	Позволяет оценить число скачков, их расположение, а также доверительные интервалы для точек разладки.	Требует решения ресурсоемкой оптимизационной задачи, в которой на интервалах может существовать достаточно большое число локальных решений.

Все эти методы в той или иной степени улучшают отдельные качественные показатели, однако основными их недостатками является сложность обучения, ресурсоемкость и увеличение времени работы алгоритма. Кроме того, очень часто возникают ситуации, когда неправильно подобранные модели и способы агрегации их результатов ухудшают общий прогноз. А в случае трансформации свойств данных без организации постоянных процессов обучения модель с течением времени может потерять адекватность [32, 33].

В связи с этим возникает задача оценки входной выборки, разделения ее на сегменты с использованием информации о свойствах данных и формирование моделей обработки, показывающих лучшие качественные показатели для вычисленных свойств обрабатываемой последовательности.

**3. Предлагаемый подход.** Адаптивное формирование модели происходит в несколько этапов. В начале осуществляется выбор базовых алгоритмов обработки. На этом шаге, в большинстве случаев, уделяется особое внимание реализации процессов обучения, которые должны обеспечить разнообразие моделей. Это решается различными «манипуляциями» над данными, где происходит формирование признаков пространств, выборок, и использованием принципиально разных моделей обработки. На следующем шаге выполняется анализ базовых моделей, в результате которого, обычно, возникает необходимость исключения тех из них, которые не позволяют достичь высокую точность прогнозирования. После этого выбирается способ обработки полученных результатов.

Большое количество рутинных повторяющихся операций обуславливает необходимость применения анализа результатов вычислений на отдельных этапах формирования модели обработки. При наличии достаточных вычислительных ресурсов становится возможным проводить автоматическое разделение последовательностей данных различными способами применяя разные алгоритмы, например разбиения последовательности или кластеризации. А затем, на отдельных последовательностях сегментов данных вычислять качественные показатели разных заранее выбранных моделей и назначать на каждый отдельный сегмент "лучшую" модель. На рисунке 1 приведены процессы формирования модели и их взаимосвязи.



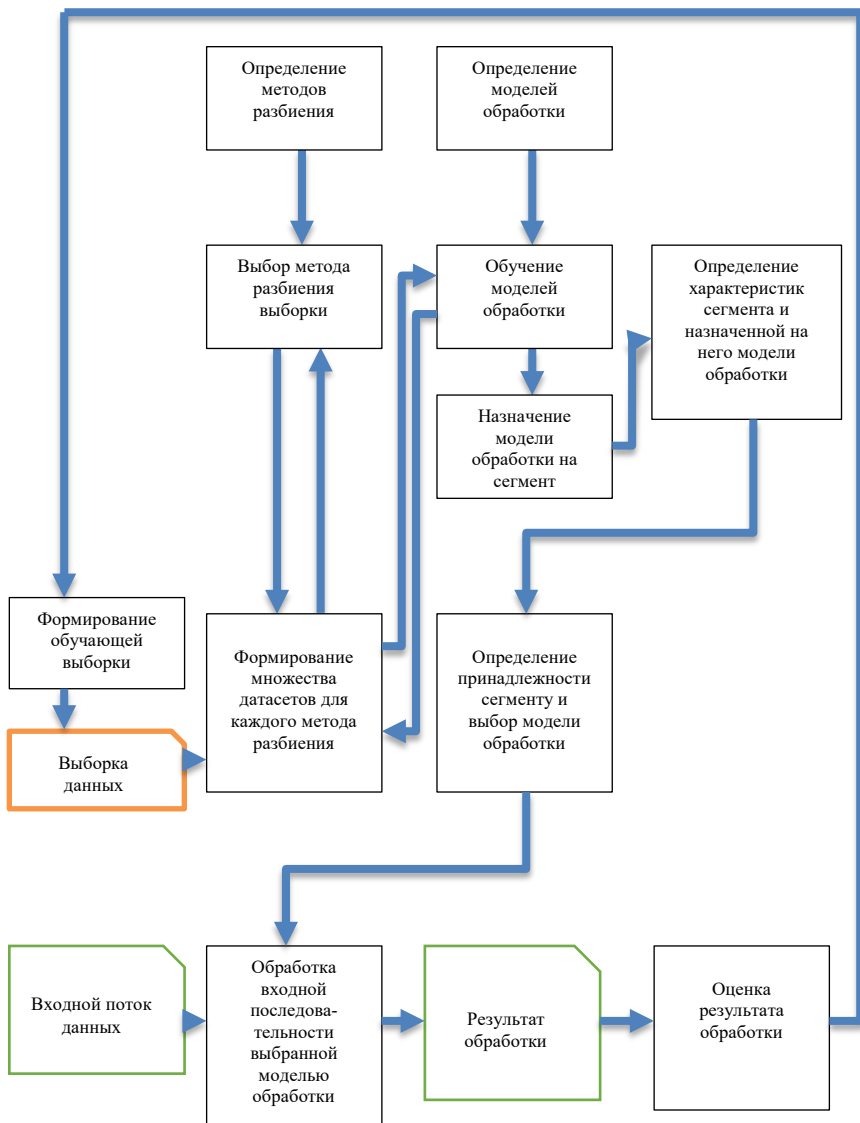


Рис. 1. Процессы формирования модели

Таким образом, в целях улучшения качества обработки в статье предлагается методика, направленная на оценку показателей качества

моделей обработки на отдельных сегментах, полученных разными способами разделения информационной последовательности.

Предполагается, что последовательность данных может быть разделена разными способами. В результате меняется количество объектов наблюдения, состав и свойства сегментов, связанные, например, с распределением данных. А это приводит к тому, что при различных способах разбиения последовательности на каждом сегменте лучшие качественные показатели будут достигать разные модели. В связи с возможной трансформацией свойств данных анализ разделяемых последовательностей необходимо осуществлять постоянно, чтобы оперативно реагировать на возникающие изменения и настраивать модель обработки.

**4. Формальная постановка задачи.** Имеется выборка последовательности объектов наблюдения  $X$ , определены модели обработки  $\{a_1, \dots, a_N\} \in A$  и методы сегментации данных  $\{\mu_1, \dots, \mu_L\} \in \Omega$ . Информационная последовательность  $X$  разбивается на отдельные сегменты. В результате получается множество способов разбиения  $X^\mu$ , сегменты которых обрабатываются моделями  $A$ .

Необходимо найти метод  $\mu^*$  разделения последовательности на сегменты  $X^{\mu^*} = \{X_{\mu^*}^1, \dots, X_{\mu^*}^n\}$  и назначить на каждый сегмент модель обработки  $a_i \in A$ , имеющую лучшее значение функционала качества  $Q(a_i(x), X_{j^*}^{\mu^*}) \rightarrow \max_{a_i \in A, \mu^* \in \Omega}$ .

Таким образом, основной акцент делается на назначение лучших по качественным показателям моделей обработки сегментов, что в определенных случаях позволяет обращать меньше внимания на реализацию вычислительно сложных процедур сегментации данных регрессии, учитывающих свойства объектов наблюдения в последовательностях данных. Возникает задача разработки методики выбора сегментов регрессионной последовательности, полученных разными способами и алгоритмами, отличающейся использованием функционала качества моделей обработки на подпоследовательности, что дает возможность формировать агрегационную модель, осуществляющую назначение лучших по качественным показателям моделей на сегменты.

В отличие от классических подходов регрессионная последовательность данных сначала разделяется выбранными методами сегментации, а затем на сегментах обучаются модели и определяются их характеристики. Базовыми алгоритмами обработки для анализа последовательностей данных могут быть линейная, логистическая регрессии, метод опорных векторов, нейросетевые и другие модели.

Выбор методов сегментации, например алгоритмов кластеризации или поиска точек изменений, определяется ресурсоемкостью и вычислительной сложностью. Предпосылкой предлагаемого метода с одной стороны является предположение о неоднородности данных, а с другой возникающая неоднозначность, например, в условиях высокой размерности, когда может существовать более одной точки в области локализации. Объекты наблюдения регрессионной последовательности могут образовывать разные области, где меняются тренды и свойства данных. В результате получаются сегменты, где может изменяться направление тренда, или в одних случаях будет линейная зависимость, в других нелинейная. Это приводит к тому, разные подпоследовательности внутри  $X$  лучше аппроксимируются различными функциями.

Рассматривая способы разбиения методами  $\mu_i \in \Omega$  последовательностей данных  $\{X_{1\mu_i}^{\mu_i}, \dots, X_{m\mu_i}^{\mu_i}\} \in X^{\mu_i}$ , получая различные комбинации сегментов данных, необходимо выбрать метод разбиения  $\mu^*$  и определить агрегируемую функцию  $a^*$  обработки данных, состоящую из моделей, при которых на сегментах достигаются лучшие показатели качества

$$a^*, \mu^* = \arg \max_{\mu_i \in \Omega, a_i \in A} Q(a_i(x, X^{\mu_i})). \quad (1)$$

Определяя метод разбиения выборки и назначая на каждый сегмент алгоритм, имеющий по результатам обучения самые высокие показатели на сегменте по сравнению с другими выбранными, получается агрегированная модель обработки последовательностей

$$a^*(x) = \left\{ \begin{array}{l} a_{i_{\mu_i}}(x), x \in X_{i_{\mu_i}}^{\mu_i} \\ a_{M_{\mu_i}}(x), x \in X_{M_{\mu_i}}^{\mu_i} \end{array} \right\}. \quad (2)$$

Таким образом, решение направлено на выбор лучшего из заранее определенных методов сегментации и количества сегментов. Это дает возможность сформировать агрегированную модель обработки, где на каждый сегмент назначается свой алгоритм, показывающий в процессе обучения лучший результат на выделенном сегменте.

**5. Реализация метода.** Применение предлагаемой методики предполагает ряд операций, необходимых для настройки моделей и выполнения ими предопределенных задач. В качестве подготовительного

этапа необходимо определить методы разбиения  $\{\mu_l, \dots, \mu_L\} \in \Omega$  последовательности объектов наблюдения и выбрать модели обработки данных  $\{a_1, \dots, a_N\} \in A$ . Для верификации методов разбиения и настройки моделей обработки заранее создается начальная тестовая выборка данных, повторяющая свойства генеральной совокупности. В листинге 1 представлен псевдокод модифицированного алгоритма [33, 34], который реализует формирование модели.

---

**Входные данные:**

$X$  – начальная тестовая выборка,

$\mu_l$  – методы разбиения выборки  $X$  в количестве  $L$ ,

$a_i$  – модели обработки данных, в количестве  $N$ ,

$Q(a(x), X)$  – функционал качества,

$M$  – максимальное количество сегментов.

---

**Выходные данные:**

$\mu^*$  – метод разбиения выборки, где достигается максимальный показатель качества

$m^{\mu^*}$  – количество сегментов, полученных выбранным  $\mu^*$  методом разбиения выборки

$X_{\mu^*}^1, \dots, X_{m^{\mu^*}}^{\mu^*}$  – сегменты выборки  $X$ , обработанной выбранным методом разбиения

$a^{\mu^*}(x, X_{m^{\mu^*}}^{\mu^*})$  – модель обработки выборки  $X$ , сегментированной методом  $\mu^*$

---

**Начало**

1. Цикл перебора методов разбиения выборки: *for*  $l = 1, \dots, L$

2. Цикл, увеличивающий количество сегментов: *for*  $m = 1, \dots, M$

3. Обработка методом разбиения  $\mu_l$  выборки  $X$ , формирование сегментов  $\{X_{1\mu_l}^{\mu_l}, \dots, X_{j\mu_l}^{\mu_l}, \dots, X_{m\mu_l}^{\mu_l}\} \in X^{\mu_l}$  для метода разбиения  $\mu_l$  и количества сегментов  $m$

4. Цикл перебора количества сегментов: *for*  $j = 1, \dots, m$

5. Цикл перебора моделей обработки  $a_i$ : *for*  $i = 1, \dots, N$

6. Обучение модели  $a_i$  на сегменте  $X_{j\mu_l}^{\mu_l}$

7. Конец цикла п. 5

8. Определение лучшей из моделей  $\{a_1, \dots, a_N\} \in A$  на сегменте  $X_{j\mu_l}^{\mu_l}$  по значению

показателя качества модели  $a^{j\mu_l} = \arg \max_{a_i \in A} Q(a_i(x, X_{j\mu_l}^{\mu_l}))$

9. Конец цикла п. 4

10. Определение выборки  $\{X_m^{\mu_l} = \{X_{1\mu_l}^{\mu_l}, \dots, X_{m\mu_l}^{\mu_l}\}$  для выбранного метода и выбранного количества сегментов

---

11. Формирование агрегационной модели из моделей,  $\{a^{1\mu_l}, \dots, a^{m\mu_l}\} \in A$ , показывающих лучшие результаты по значению показателя качества, на  $m$  сегментах на выборке  $X_m^{\mu_l}$  после обработки методом  $\mu_l$

$$a_m^{\mu_l}(x, X_m^{\mu_l}) = \left\{ \begin{array}{l} a^{1\mu_l}(x, X_{1\mu_l}^{\mu_l}), x \in X_{1\mu_l}^{\mu_l} \\ \dots \\ a^{m\mu_l}(x, X_{m\mu_l}^{\mu_l}), x \in X_{m\mu_l}^{\mu_l} \end{array} \right\}$$

12. Конец цикла п. 2

13. Определение количества сегментов разбиения выборки  $X$  методом  $\mu_l$ , при котором был достигнут лучший показатель качества  $m^{\mu_l} = \arg \max_{m \in \{1, \dots, M\}} Q(a_m^{\mu_l}(x, X_m^{\mu_l}))$

Определение сегментов выборки  $X_{m^{\mu_l}}^{\mu_l} = \{X_{1\mu_l}^{\mu_l}, \dots, X_{m^{\mu_l}\mu_l}^{\mu_l}\}$

Определение модели, которая достигает лучшего показателя качества при разбиении методом  $\mu_l$

$$a_{m^{\mu_l}}^{\mu_l}(x, X_{m^{\mu_l}\mu_l}^{\mu_l}) = \arg \max_{a_m^{\mu_l} \in A} Q(a_m^{\mu_l}(x, X_m^{\mu_l}))$$

14. Конец цикла п. 1

15. Окончательный выбор метода разбиения выборки, где достигается максимальный показатель качества

$$\mu^* = \arg \max_{\mu_l \in \mu} Q(a_{m^{\mu_l}}^{\mu_l}(x, X_{m^{\mu_l}\mu_l}^{\mu_l}))$$

Определение количества сегментов, полученных выбранным методом разбиения выборки

$$m^{\mu^*} = \arg \max_{m \in \{1, \dots, M\}} Q(a_m^{\mu^*}(x, X_m^{\mu^*}))$$

Определение сегментов выборки, обработанной выбранным методом разбиения

$$X^{\mu^*} = \{X_{1\mu^*}^{\mu^*}, \dots, X_{m^{\mu^*}\mu^*}^{\mu^*}\}$$

Окончательное формирование модели обработки

$$a^{\mu^*}(x, X^{\mu^*}) = \left\{ \begin{array}{l} a_{1\mu^*}^{\mu^*}(x, X_{1\mu^*}^{\mu^*}), x \in X_{1\mu^*}^{\mu^*} \\ \dots \\ a_{m^{\mu^*}\mu^*}^{\mu^*}(x, X_{m^{\mu^*}\mu^*}^{\mu^*}), x \in X_{m^{\mu^*}\mu^*}^{\mu^*} \end{array} \right\}$$

**Конец**

Листинг 1. Псевдокод алгоритма формирования модели обработки

В представленном алгоритме максимальное число рассматриваемых сегментов из одной последовательности равно  $M^2$ . Условно накладные расходы на усредненное время обучения и обработки для модели  $a_i$  на сегменте можно оценить величиной  $p$ , тогда общая сложность алгоритма составляет  $O(pLN^2M)$ . Количество методов разбиения выборки  $L$  и количество моделей обработки  $N$  обычно относительно небольшие величины. На рост сложности влияет число рассматриваемых сегментов  $M$ , что является существенным ограничением предлагаемого метода.

Однако в целях оптимизации и ускорения алгоритмов обработки временных рядов и информационных последовательностей возможно применение ряда подходов, направленных на анализ и объединения «сходных» по свойствам данных сегментов, отбрасывания способов разбиения выборки при качественных показателях обучения моделей меньше заданного порога, применение распараллеливания процессов обучения для моделей.

Реализация модели предполагает выполнение ряда действий по формированию информационной последовательности и определению методов обработки и сегментации. Каждому методу разбиения задаются параметры, например количество разбиений, функция расстояния для кластеризации, функция определения изменения тренда для методов поиска точки разладки. В целях упрощения можно считать, что методы с разными параметрами являются различными методами разбиения. Затем каждому методу подается на вход выборка данных  $X$  и формируются сегменты  $X_{1\mu_i}^{\mu_i}, \dots, X_{M\mu_i}^{\mu_i}$ . В соответствие методу разбиения  $\mu_i \in \Omega$  ставится  $X_{1\mu_i}^{\mu_i}, \dots, X_{M\mu_i}^{\mu_i} \in X^{\mu_i}$  множество сегментов выборки.

На объектах наблюдения  $x$  каждого из множества сегментов  $X_{1\mu_i}^{\mu_i}, \dots, X_{M\mu_i}^{\mu_i}$ , образованных методом разбиения  $\mu_i \in \Omega$ , осуществляется обучение всех заранее определенных моделей  $\{a_1, \dots, a_N\} \in A$  обработки последовательности.

Далее происходит регуляция системы, где для каждого метода разбиения  $\mu_i \in \Omega$  для всех обученных моделей  $\{a_1, \dots, a_N\} \in A$  на полученных с его применением сегментах  $X_{1\mu_i}^{\mu_i}, \dots, X_{M\mu_i}^{\mu_i}$  сравниваются значения функционала качества  $Q(a_i(x, X_{j\mu_i}^{\mu_i}))$ . Каждому методу разбиения  $\mu_i \in \Omega$  на каждом сегменте  $X_{j\mu_i}^{\mu_i}$  ставятся в соответствие

модели, имеющие лучшие показатели качества. Применяя выражение (1) становится возможным вычислить выбрать метод разбиения  $\mu_i \in \Omega$  при котором лучшие модели показывают лучшие результаты на сегментах. Итоговая модель определяется выражением (2). Формируются кортежи  $\langle a_{j_{\mu_i}}, X_{j_{\mu_i}}^{\mu_i} \rangle$ , где на сегмент назначается модель.

В дальнейшем при обработке данных сегменту ставится в соответствие информация о его свойствах. В зависимости от решаемой задачи в качестве свойств могут использоваться статистические характеристики среднего значения данных сегмента, направление тренда, скорость изменения значений и т.д. Совершенствование предложенной модели возможно с использованием методов постоянного обучения, по аналогии с нейросетевыми решениями [35]. Для этого может быть применена многоуровневая обработка [32, 34, 36], где модели разделяются на уровни. При ее реализации в начале выполняется настройка моделей на обучающей выборке, решающих задачи обработки информации. Производится выбор оптимального метода разбиения  $\mu_i \in \Omega$  и количества сегментов. Выполняется первоначальное обучение. Вычисляются свойства сегментов. Поступающий на вход информационный поток подвергается обработке. Производится анализ и вычисление свойств объектов наблюдения. Они соотносятся со свойствами сегментов, и для их обработки выбирается заранее определенная модель, имеющая лучшие значения функционала качества на схожем по свойствам сегменте обучающей выборки. Предсказанные моделями результаты сравниваются с реальными значениями объектов наблюдения, полученными от регистрирующих систем и устройств. В случае увеличения ошибок выше заранее определенного порога принимается решение о формировании выборки данных. Над выборкой проводятся манипуляции, указанные в алгоритме (листинг 1), происходит настройка и обучение модели. После этого анализируемый поток данных поступает на вход обученной модели.

Таким образом, возможно применение методов самообучения для рассматриваемого решения. Однако эти действия необходимо выполнять в параллельном режиме, чтобы избежать существенного роста вычислительной сложности.

**6. Эксперимент.** Цель проведения эксперимента состояла в оценке повышения качественных показателей обработки информационных последовательностей при применении рассматриваемой методики. Так как предлагаемая методика основана на сегментировании выборки данных, то были рассмотрены два основных подхода к разделению. В первом случае применялся контролируемый подход, когда

осуществлялось нахождение границ сегмента на основе вычисления смены тренда информационной последовательности. Во втором случае был применен подход, разделяющий последовательность на равные по количеству наблюдений сегменты. Оценка предлагаемого метода осуществлялась для задачи прогнозирования одномерной и многомерной последовательности данных.

В предлагаемом методе в качестве базовых моделей в зависимости от задачи обработки можно использовать модели любой сложности. В эксперименте при выборе моделей обработки приоритет отдавался имеющим высокую скорость обучения. Сравнивались результаты линейной регрессии (LR), квадратичной регрессии (QR), машины опорных векторов (SVM), регрессии гауссова процесса (GR), деревьев решений (DT). Сравнение качественных показателей осуществлялось для различных наборов данных. В одном случае сегментирование выборки осуществлялось простым разделением на равное количество объектов наблюдения. В другом – подвергалась обработке алгоритмом оптимального обнаружения точек изменения с линейными вычислительными затратами, реализованного в среде Matlab [37]. Информационная последовательность объектов наблюдения разделялась в первом случае на  $m$  равных частей, а во втором – на  $m$  сегментов с использованием информации о точках изменений. Затем каждый сегмент аппроксимировался моделями регрессии. Поступающий на вход модели объект наблюдения входной последовательности анализировался на принадлежность сегменту. Вычислялись его значения  $\hat{y}_i$  полученные моделью, назначенной на сегмент, и сравнивалось с истинным значением  $y_i$ . Показатели качества обработки данных определялись метриками MSE и MAE.

$$MSE = \frac{1}{n} \sum_{i=1}^n (\hat{y}_i - y_i)^2, \quad (3)$$

$$MAE = \frac{1}{n} \sum_{i=1}^n | \hat{y}_i - y_i |, \quad (4)$$

где  $\hat{y}_i$  – предсказанное значение,  $y_i$  – реальное значение,  $n$  – количество объектов наблюдения.

**6.1. Эксперимент на модельных данных.** Для иллюстрации предлагаемой методики был проведен эксперимент на модельных данных. Рассматриваемая информационная последовательность подвергалась сегментации. Модельные данные эксперимента для одномерной регрессии [38] приведены на рисунке 2.



В верхней части рисунка приведено разбиение на 4 сегмента равных по количеству объектов наблюдения, в нижней части – алгоритмом оптимального обнаружения точек изменения с линейными вычислительными затратами с 4 сегментами.

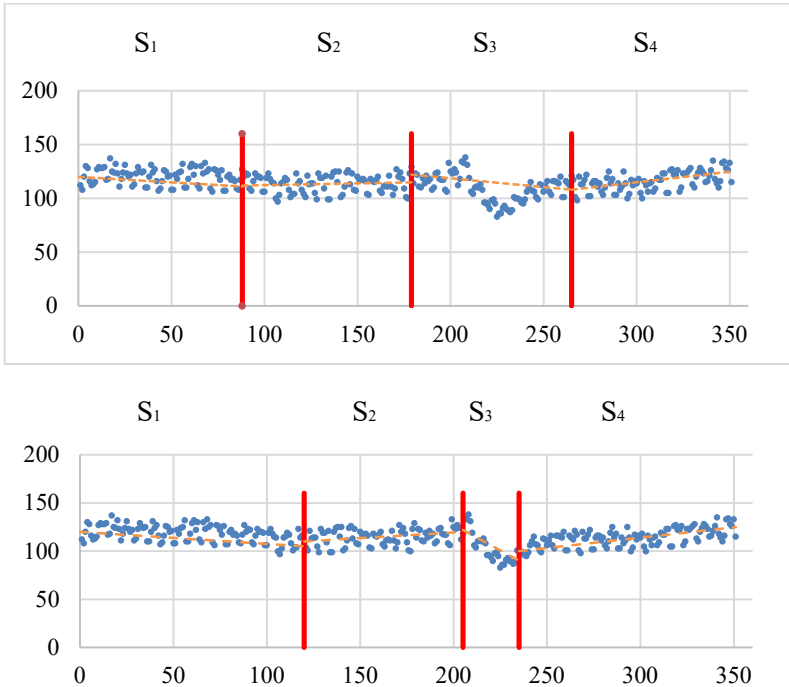


Рис. 2. Деление последовательности на 4 сегмента (на равные части – верхняя часть, алгоритмом оптимального обнаружения точек изменения с линейными вычислительными затратами – нижняя часть)

На рисунке 3 представлены гистограммы функций потерь. На них приведены значения MSE и MAE на каждом из 4 сегментов рисунка 2, полученным разными способами и аппроксимированными линейными и квадратичными функциями.

Для анализируемых данных вычисленные значения выражений (3) и (4) показывают преимущество использования методов разделения последовательности. Значения функций потерь MSE и MAE оказались для данных после сегментирования ниже, чем для всей выборки целиком. Разбиение алгоритмом оптимального обнаружения точек изменения с линейными вычислительными затратами для рассматриваемого случая

оказалось предпочтительнее, чем при разделении на равных по количеству объектов наблюдения сегменты.

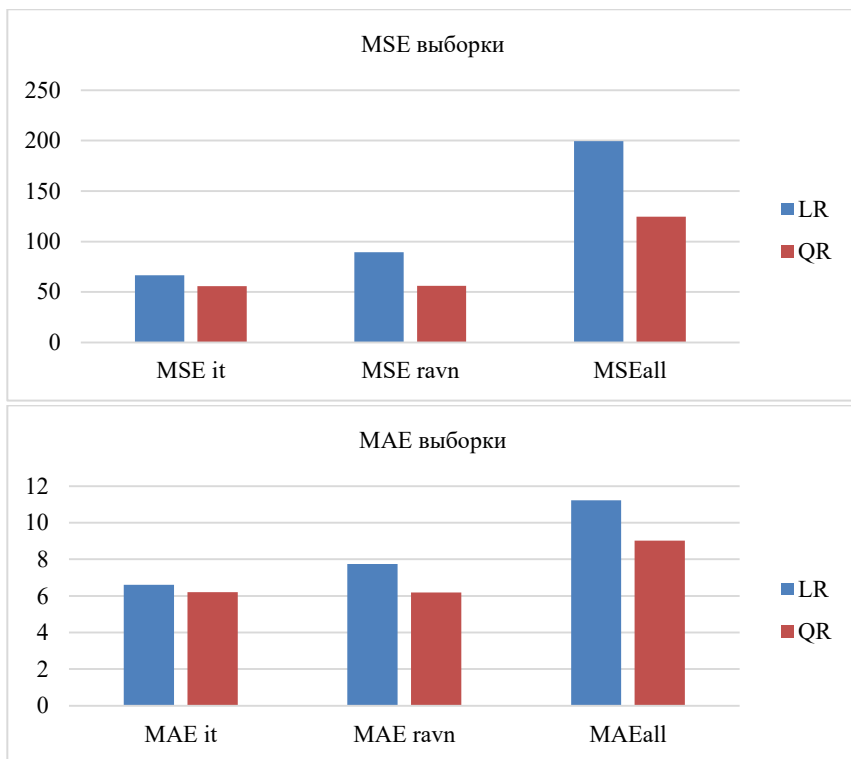


Рис. 3. Функция потерь MSE и MAE при аппроксимации LR и QR на всей выборке и при делении на 4 сегмента равным количеством объектов наблюдения (ravn) и алгоритмом оптимального обнаружения точек изменения с линейными вычислительными затратами (it)

Полученные сегменты могут быть неоднородны. На одних сегментах могут лучше показывать себя одни алгоритмы, на других – другие. Разделив последовательность, например, на 4 сегмента можно увидеть, что на различных сегментах значения функции потерь будут отличаться. На рисунке 4 приведены значения MSE и MAE аппроксимирующих функций на каждом из 4 сегментов, полученных алгоритмом оптимального обнаружения точек изменения с линейными вычислительными затратами и простым разделением на равные отрезки.

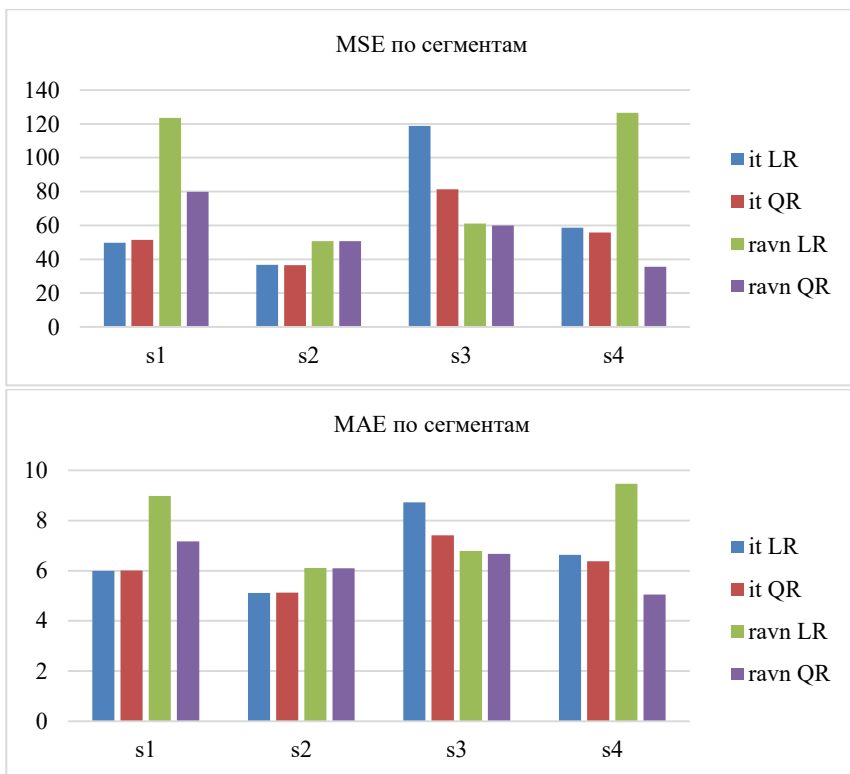


Рис. 4. Значения MSE и MAE алгоритмов LR, QR для сегментов, полученных равным количеством объектов наблюдения (ravn) и алгоритмом оптимального обнаружения точек изменения с линейными вычислительными затратами (it)

Анализ гистограмм рисунка 4 показывает, что при разбиении алгоритмом оптимального обнаружения точек изменения с линейными вычислительными затратами на сегментах  $s_1$ ,  $s_2$  меньшие значения функции потерь показывает линейная регрессия, а на сегментах  $s_3$ ,  $s_4$  – квадратичная регрессия. При простом разделении линейная регрессия показывает лучшие результаты на  $s_2$ , а на остальных сегментах – квадратичная регрессия. При определении значений функционала качества для данных внутри сегментов возможно решение задачи по назначению модели обработки на сегмент, где он показывает лучшие значения.

Далее в эксперименте производилась оценка влияния количества сегментов  $m$  на значения MSE и MAE. Число сегментов увеличивалось

(от 1 (вся выборка) до 35) и оценивались значения функций потерь MSE и MAE для аппроксимирующих функций и способов разбиения.

На рисунке 5 приведена зависимость усредненных значений MSE и MAE линейной и квадратичной регрессии от количества сегментов, полученных при делении последовательности на равные части и при использовании алгоритма оптимального обнаружения точек изменения с линейными вычислительными затратами.

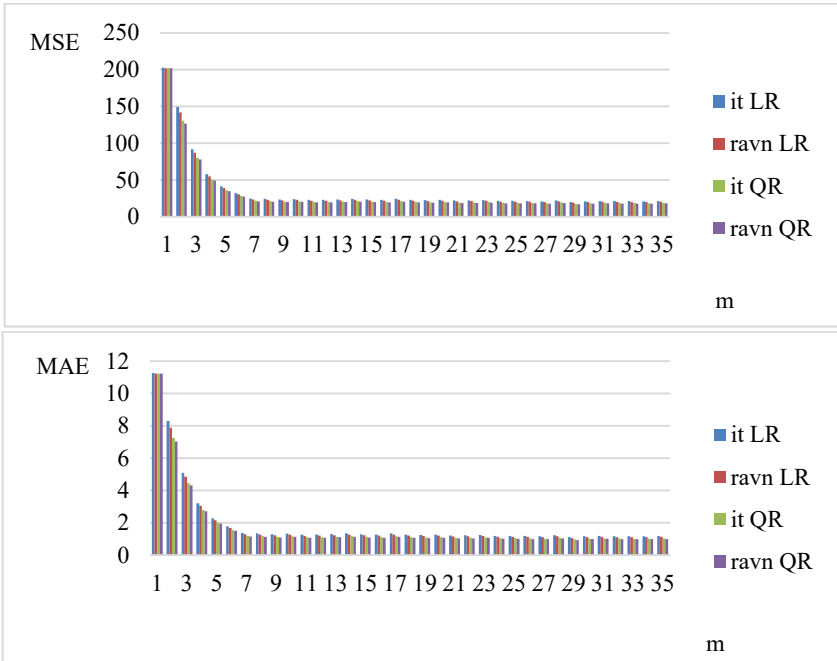


Рис. 5. Зависимость значений MSE и MAE от количества сегментов  $m$  для линейной регрессии при разделении на равные части по количеству объектов (ravn) и алгоритмом оптимального обнаружения точек изменения (it) с линейными вычислительными затратами

Рисунок 5 показывает, что на достигаемые значения MSE и MAE, оказывает влияние количество сегментов. Значения функции потерь для рассматриваемых моделей уменьшаются при увеличении количества сегментов. Кроме того, на графиках видно, что до определенного момента, чем больше сегментов создается, тем меньше становятся MSE и MAE для выбранных аппроксимирующих функций и способов сегментации. Затем показатели функций потерь моделей выходят на

плато и дальнейшее увеличение количества сегментов к существенному улучшению функционала качества не приводят.

В то же время, необходимо отметить, что при неограниченном разрастании количества сегментов может возникнуть ситуация, когда будут появляться «микросегменты» содержащие, недостаточное количество объектов наблюдения для построения адекватной модели. Это требует дополнительного анализа и обработки.

**6.2. Применение методики при появлении точек разрыва.** При анализе последовательности данных могут возникать ситуации, где появляются точки разрыва. При прохождении такой точки может меняться вид регрессионного уравнения.

Несмотря на применение алгоритмов IJD, SEG [28, 37, 39] и ряда других, использующих средние значения, дисперсии данных, изменение наклона линейной регрессии, их результаты зависят от последовательностей данных и могут приводить к существенным отличиям в определении границ и составе объектов наблюдения в сегментах.

На рисунке 6 приведены модельные данные с разрывом.

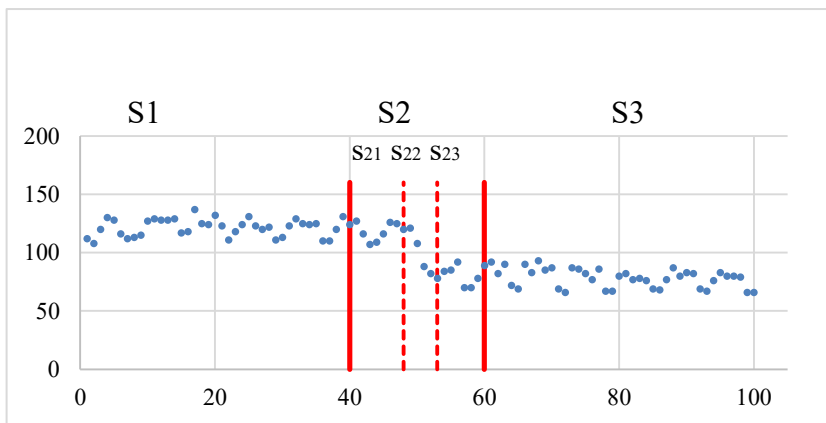


Рис. 6. Локализация окрестностей точек потенциального разрыва дальнейшем делением сегментов

Для обработки подобных последовательностей необходима либо сложная функция аппроксимации, либо определение точки потенциального разрыва как границы сегмента. Повышение сложности функции аппроксимации приводит к увеличению ресурсоемкости, вычислительной сложности, одновременно с этим может возникнуть эффект переобучения модели, а это не гарантирует повышение качества результата обработки.

Определение точки разрыва, как границы сегмента имеет свои сложности, возникающие из-за нечетких окрестностей в неоднозначных случаях. Например, при переходе системы из одного режима работы в другой наблюдаются пограничные ситуации, когда полученные значения не позволяют однозначно трактовать состояние. В этом случае на основе предлагаемой методики возможно зафиксировать модель, получить сегменты, определить показатели качества и в случае, если они оказываются хуже заданного уровня, то произвести дальнейшее деление.

На рисунке 7 показано, что для рассматриваемых данных дальнейшее деление сегмента S2 дает возможность уменьшить значения функции потерь MAE и MSE для линейной и квадратичной регрессии как в случае применения анализа точек изменения, так и в случае простого деления сегмента.

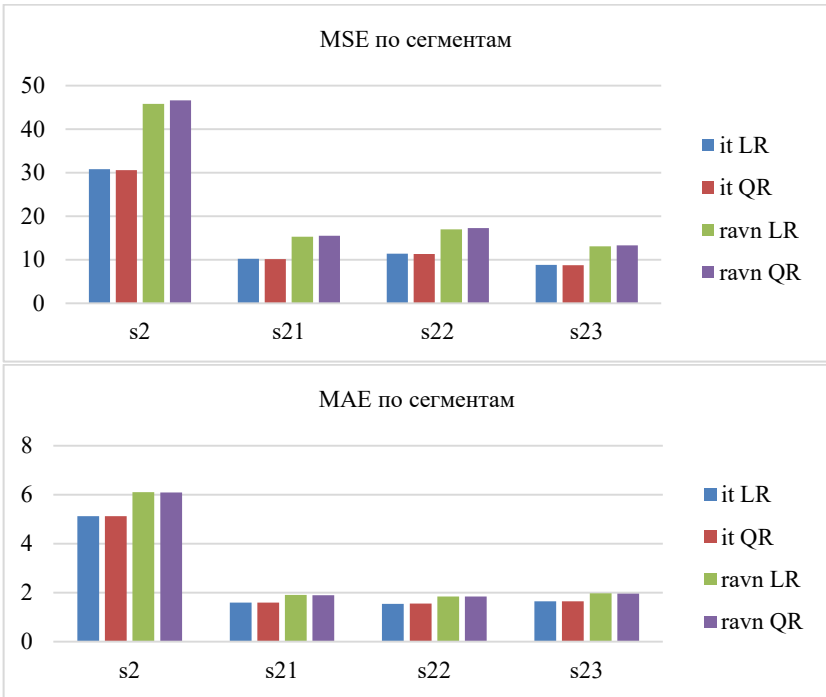


Рис. 7. Значения функции потерь при делении сегмента

Таким образом, для заданной модели можно уменьшать размер сегмента, анализируя выбранный показатель качества. Если значение оказываются хуже заданного уровня, то производить дальнейшее

деление. Затем, определяя значения функционала качества, осуществлять выбор моделей. При этом необходимо также учитывать их ресурсоемкость и сложность процессов обучения.

**6.3. Эксперимент на множественной регрессии.** В условиях высокой размерности, когда может существовать более одной точки изменения, возникает задача их обнаружения и локализации.

В большинстве практических случаев анализируемые данные являются многомерными, что накладывает определенные сложности их предварительной обработки, например при обнаружении выбросов и шумов.

Возникают определенные проблемы, связанные с оптимальным определением точки изменения в регрессионной модели.

Для оценки достигаемых качественных показателей при сегментировании многомерной последовательности данных была применена множественная регрессия. В качестве исследуемого набора данных была использована выборка [38], которая обрабатывалась моделями линейной регрессии LR, регрессия гауссова процесса GR, машина опорных векторов SVM, деревья решений DT.

На рисунке 8 представлены графики изменений MSE(m) и MAE(m) от количества сегментов m для базовых алгоритмов LR, GR, SVM, DT. На графиках прослеживается, что для большинства относительно простых моделей, например LR, SVM, GR к повышению показателей качества обработки может приводить уменьшение сегмента без учета свойств содержащихся в нем объектов наблюдения.

Однако предлагаемая методика, использующая уменьшение размера сегмента, позволяет получить эффект не для всех моделей. Например, для рассматриваемой выборки данных уменьшение размера сегмента оказывает более слабое влияние на алгоритм DT по сравнению с другими. Значение показателей MSE и MAE для него понижается незначительно.

Несмотря на сравнительную простоту реализации предложенного метода имеется ряд ограничений, который необходимо учитывать при формировании моделей обработки. Информационные последовательности могут иметь различные свойства в разных сегментах, что оказывает влияние на выбор наиболее эффективных моделей их обработки.

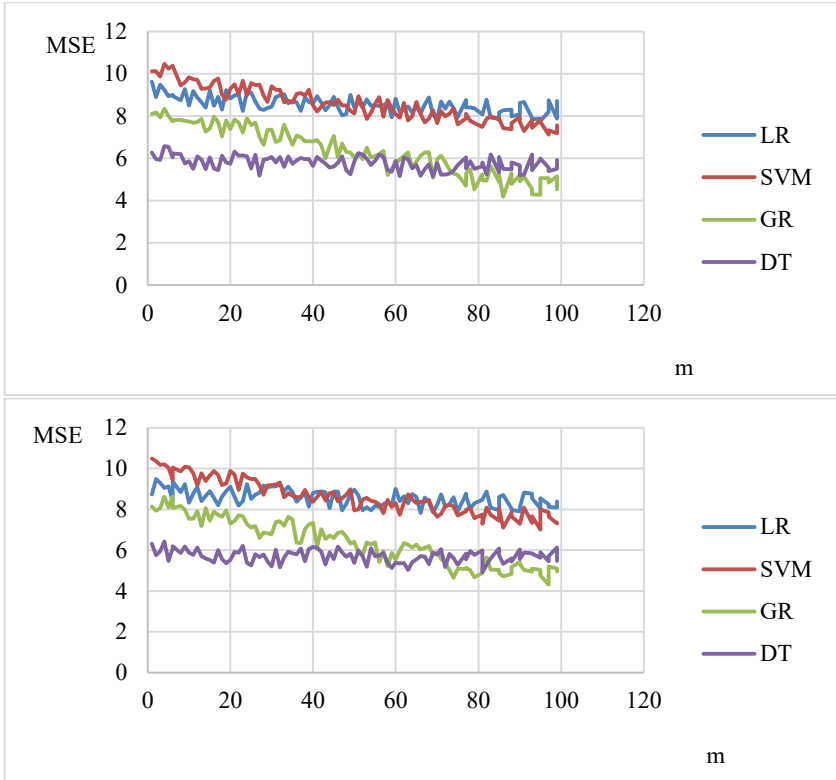


Рис. 8. Зависимость значений MSE различных алгоритмов от количества сегментов  $m$  для множественной регрессии при сегментировании делением на равные отрезки (вверху) и алгоритмом оптимального обнаружения точек изменения с линейными вычислительными затратами (внизу)

**6.4. Эксперимент на данных датасетов.** Далее эксперимент был посвящен анализу ряда различных наборов данных множественной регрессии. В качестве экспериментальных данных рассматривались выборки [38 – 40]. Объем первого набора около 30000 записей, 3 предиктора и 1 целевая переменная, второго – 32000 записей, 123 предиктора и 1 целевая переменная, третьего – 32000 записей, 8 предикторов и 1 целевая переменная.

Данные представлялись временными рядами.

В качестве исследуемых определены модели обработки данных и составленные из них ансамбли LR, SVM, GR, DT, LR+SVM, SVM+GR, не требующие больших вычислительных затрат. Производилась разбиение



выборки на 15 и 25 сегментов алгоритмом оптимального обнаружения точек изменения с линейными вычислительными затратами и разделением на равные по количеству объектов наблюдения, определялся функционал качества каждой модели. В качестве анализируемых показателей были выбраны MSE и MAE.

Результаты качественных показателей, полученных при предсказании для сегментов и всей выборки целиком, представлены в таблицах 2-4.

Таблица 2. Обработка выборки [38]

	Модель	Вся выборка		15 сегментов		25 сегментов	
		MSE	MAE	MSE	MAE	MSE	MAE
Алгоритм обнаружения точек изменения	LR	137,08	9,25	126,33	8,88	<b>116,43</b>	<b>8,52</b>
	SVM	218,13	11,67	201,03	11,20	<b>185,27</b>	<b>10,75</b>
	GR	80,51	7,09	74,20	6,81	<b>68,38</b>	<b>6,53</b>
	DT	46,35	5,38	42,71	5,16	<b>39,37</b>	<b>4,96</b>
	LR+SVM	103,29	8,03	95,19	7,71	<b>87,73</b>	<b>7,40</b>
	SVM+GR	64,02	6,32	59,00	6,07	<b>54,37</b>	<b>5,83</b>
Разделение на равные сегменты	LR	165,87	10,17	152,86	9,77	<b>140,88</b>	<b>9,38</b>
	SVM	263,94	12,83	243,25	12,32	<b>224,18</b>	<b>11,83</b>
	GR	97,42	7,80	89,78	7,49	<b>82,74</b>	<b>7,19</b>
	DT	56,08	5,92	51,68	5,68	<b>47,63</b>	<b>5,45</b>
	LR+SVM	124,98	8,83	115,18	8,48	<b>106,15</b>	<b>8,14</b>
	SVM+GR	77,46	6,95	71,39	6,67	<b>65,79</b>	<b>6,41</b>

Таблица 3. Обработка выборки [41]

	Модель	Вся выборка		15 сегментов		25 сегментов	
		MSE	MAE	MSE	MAE	MSE	MAE
Алгоритм обнаружения точек изменения	LR	0,0025	0,0300	0,0024	0,0291	<b>0,0022</b>	<b>0,0282</b>
	SVM	0,0049	0,0553	0,0046	0,0536	<b>0,0043</b>	<b>0,0520</b>
	GR	0,0016	0,0316	0,0015	0,0307	<b>0,0014</b>	<b>0,0297</b>
	DT	0,0009	0,0237	<b>0,0008</b>	0,0230	<b>0,0008</b>	<b>0,0223</b>
	LR+SVM	0,0036	0,0474	0,0034	0,0460	<b>0,0032</b>	<b>0,0446</b>
	SVM+GR	0,0016	0,0316	0,0015	0,0307	<b>0,0014</b>	<b>0,0297</b>
Разделение на равные сегменты	LR	0,0036	0,0360	0,0034	0,0349	<b>0,0032</b>	<b>0,0339</b>
	SVM	0,0071	0,0664	0,0066	0,0644	<b>0,0062</b>	<b>0,0624</b>
	GR	0,0023	0,0379	0,0022	0,0368	<b>0,0020</b>	<b>0,0357</b>
	DT	0,0013	0,0284	0,0012	0,0276	<b>0,0011</b>	<b>0,0268</b>
	LR+SVM	0,0052	0,0569	0,0049	0,0552	<b>0,0046</b>	<b>0,0535</b>
	SVM+GR	0,0023	0,0379	0,0022	0,0368	<b>0,0020</b>	<b>0,0357</b>

Таблица 4. Обработка выборки [42]

	Модель	Вся выборка		15 сегментов		25 сегментов	
		MSE	MAE	MSE	MAE	MSE	MAE
Алгоритм обнаружения точек изменения	LR	6,9169	0,7100	6,5081	0,6887	<b>6,1235</b>	<b>0,6680</b>
	SVM	9,7344	0,8400	9,1591	0,8148	<b>8,6178</b>	<b>0,7904</b>
	GR	4,7089	0,5700	4,4306	0,5529	<b>4,1688</b>	<b>0,5363</b>
	DT	3,7249	0,4500	3,5048	0,4365	<b>3,2976</b>	<b>0,4234</b>
	LR+SVM	7,9524	0,6300	7,4824	0,6111	<b>7,0402</b>	<b>0,5928</b>
	SVM+GR	5,1076	0,5500	4,8057	0,5335	<b>4,5217</b>	<b>0,5175</b>
Разделение на равные сегменты	LR	11,6896	0,9230	10,9987	0,8953	<b>10,3487</b>	<b>0,8685</b>
	SVM	16,4511	1,0920	15,4789	1,0592	<b>14,5641</b>	<b>1,0275</b>
	GR	7,9580	0,7410	7,4877	0,7188	<b>7,0452</b>	<b>0,6972</b>
	DT	6,2951	0,5850	5,9230	0,5675	<b>5,5730</b>	<b>0,5504</b>
	LR+SVM	13,4396	0,8190	12,6453	0,7944	<b>11,8979</b>	<b>0,7706</b>
	SVM+GR	8,6318	0,7150	8,1217	0,6936	<b>7,6417</b>	<b>0,6727</b>

Результаты в таблицах 2-4 показывают, что уменьшение «размера» сегмента последовательности, в основном, уменьшают функции потерь MSE и MAE от 2 до 5%. Предложенный метод, во-первых, облегчает сложность процессов анализа. Во-вторых, он может приводить к тому, что при обработке появляется возможность построения более простой аппроксимирующей функции. В-третьих, применение сегментации дает возможность локализовывать группы объектов наблюдения и нивелировать влияние различных эффектов, связанных с шумовыми данными и выбросами. Однако для применения предлагаемого решения, необходимо анализировать, чтобы данные внутри сегмента повторяли свойства генеральной совокупности, в противном случае могут возникнуть проблемы репрезентативности обрабатываемой выборки и модель не будет адекватна.

Полученные результаты показывают, что оценка потенциальных качественных показателей при обучении моделей на сегментах дает возможность определить для каждого сегмента алгоритм обработки, обладающий лучшими показателями функционала качества.

Сегментация выборок данных во многих случаях создает предпосылки для повышения качественных показателей обработки. Получаемые различными методами сегменты характеризуются разными свойствами. В связи с чем при их обработке результаты алгоритмов могут существенно отличаться. Формирование сегмента информационной последовательности данных обычно приводит к возможности построению более простой аппроксимирующей функции, уменьшению

вычислительной сложности. Однако может вызывать проблемы, связанные с необходимостью более детального анализа данных на предмет выбросов и шумов, которые на практике могут быть сложны в интерпретации.

Уменьшение размера сегмента посредством увеличения их количества целесообразно до определенного предела. После его достижения существенного прироста качественных показателей моделей обработки данных не происходит. Назначение алгоритмов с лучшими качественными показателями на сегменты позволяет повысить качественные показатели обработки выборки.

Представленная методика направлена на совершенствование моделей обработки данных. При ее реализации возможно параллельное функционирование алгоритмов.

**7. Заключение.** Основное преимущество предлагаемой методики построения регрессионных моделей на основе анализа функционала качества обработки сегментов последовательности – отсутствие необходимости использовать сложные методы обнаружения точек изменения.

Причем, когда размерность данных относительно небольшая точки изменения свойств интуитивно понятны, но с ростом размерности данных интерпретировать такие точки становится затруднительно. Кроме того, каждый применяемый метод будет давать свой результат разделения данных, не совпадающий с другими методами, что может сказываться на результате обработки данных.

В информационных последовательностях могут присутствовать выбросы, однако определение объектов наблюдения, подлежащих исключению, является сложной задачей. Определить ситуацию, связанную с выбросом или закономерным появлением объекта наблюдения не всегда представляется возможным, что влияет на результат обнаружения границ сегментов.

В предложенной методике предлагается осуществлять деление выборки данных на сегменты, а далее "подбирать" на них аппроксимирующие модели. Критерием выбора модели является заранее определенный показатель качества, который сравнивается с достигаемыми значениями других моделей, что позволяет выбрать на отрезке лучшую модель.

Предлагаемая методика является относительно простой, позволяет быстро разделять последовательности для различных видов данных регрессии. Эксперименты показывают, что, сегментируя выборку данных и применяя различные модели аппроксимации, можно достигать определенного уровня значений показателя качества.

В то же время предложенная методика построения регрессионных моделей обработки данных обладает определенными ограничениями. Во-первых, рассматриваемая выборка должна обладать свойствами генеральной совокупности. В случае несоблюдения этого свойства будет возникать проблема адекватности предлагаемой модели обработки данных. Во-вторых, в процессе разделения данных на сегменты могут возникать ситуации, когда в обучающей выборке оказывается слишком мало объектов наблюдения, что не позволяет правильно оценить свойства и обучить модели на этих сегментах. В-третьих, при использовании предлагаемого решения желательно проводить анализ данных, чтобы избежать проблем переобучения.

Тем не менее предложенная методика позволяет формировать сегменты информационной последовательности на основе использования функционала качества моделей обработки информационных последовательностей. А это, в свою очередь, даст возможность реализовать агрегационную модель, где выполняется назначение лучших по качественным показателям моделей на сегменты.

### Литература

1. Chen H.Y., Chen C. Evaluation of Calibration Equations by Using Regression Analysis: An Example of Chemical Analysis // *Sensors*. 2022. vol. 22. no. 2. DOI: 10.3390/s22020447.
2. Schober P., Vetter T.R. Segmented Regression in an Interrupted Time Series Study Design // *Anesthesia and Analgesia*. 2021. vol. 132. no. 3. pp. 696–697.
3. Bozpolat E. Investigation of the self-regulated learning strategies of students from the faculty of education using ordinal logistic regression analysis // *Educational Sciences: Theory & Practice*. 2016. no. 16(1). pp. 301–318.
4. Jarantow S.W., Pisors E.D., Chiu M.L. Introduction to the use of Linear and Nonlinear Regression Analysis in Quantitative Biological Assays // *Current Protocols*. 2023. no. 3. DOI: 10.1002/cpz1.801.
5. Britzger D. The Linear Template Fit // *The European Physical Journal C*. 2022. vol. 82(8). DOI: 10.1140/epjc/s10052-022-10581-w.
6. Perperoglou A., Sauerbrei W., Abrahamowicz M., Schmid M. A review of spline function procedures in R // *BMC Medical Research Methodology*. 2019. vol. 19. pp. 1–16.
7. Ren J., Tapert S., Fan C.C., Thompson W.K. A semi-parametric Bayesian model for semi-continuous longitudinal data // *Statistics in Medicine*. 2022. vol. 41. no. 13. pp. 2354–2374.
8. Taye M.M. Theoretical Understanding of Convolutional Neural Network: Concepts, Architectures, Applications, Future Directions // *Computation*. 2023. vol. 11. no. 3. DOI: 10.3390/computation11030052.
9. Колмогоров А.Н. О представлении непрерывных функций нескольких переменных в виде суперпозиции непрерывных функций одного переменного // *Доклады АН СССР*. 1957. Т. 114. № 5. С. 953–956.
10. Giasi F., Poggio T. Representation Properties of Networks: Kolmogorov's Theorem is Irrelevant. *Neural Computation*. 1989. vol. 4. no. 1. pp. 465–469.
11. Parhi R., Nowak R.D. Banach Space Representer Theorems for Neural Networks and Ridge Splines // *Journal of Machine Learning Research*. 2021. vol. 22(1). pp. 1960–1999.

12. Marques H.O., Swersky L., Sander J., Campello R.J., Zimek A. On the evaluation of outlier detection and one-class classification: a comparative study of algorithms, model selection, and ensembles // *Data Mining and Knowledge Discovery*. 2023. vol. 37. no. 4. pp. 1473–1517.
13. Li Y., Guo X., Lin W., Zhong M., Li Q., Liu Z., Zhong W., Zhu Z. Learning dynamic user interest sequence in knowledge graphs for click-through rate prediction // *IEEE Transactions on Knowledge and Data Engineering*. 2023. vol. 35. no. 1. pp. 647–657.
14. Rinaldo A., Wang D., Wen Q., Willett R., Yu Y. Localizing changes in highdimensional regression models // *The 24th International Conference on Artificial Intelligence and Statistics*. 2021. pp. 2089–2097.
15. Aue A., Rice G., Sönmez O. Detecting and dating structural breaks in functional data without dimension reduction // *Journal of the Royal Statistical Society. Series B, Statistical Methodology*. 2018. vol. 80. no. 3. pp. 509–529.
16. Data A., Zou H., Banerjee S. Bayesian high-dimensional regression for change point analysis // *Statistics and its Interface*. 2019. vol. 12. no. 2. pp. 253–264. DOI: 10.4310/SII.2019.v12.n2.a6.
17. Melnyk I., Banerjee A. A spectral algorithm for inference in hidden semi-Markov models // *Journal of Machine Learning Research*. 2017. vol. 18. no. 35. pp. 1–39.
18. Haynes K., Fearnhead P., Eckley I.A. A computationally efficient nonparametric approach for changepoint detection // *Statistics and Computing*. 2017. vol. 27. pp. 1293–1305. DOI: 10.1007/s11222-016-9687-5.
19. Muggeo V. Estimating regression models with unknown break-points // *Statistics in Medicine*. 2003. vol. 22(19). pp. 3055–3071.
20. Lu K.P., Chang S.T. A fuzzy classification approach to piecewise regression models // *Applied Soft Computing Journal*. 2018. vol. 69. pp. 671–688.
21. Bardwell L., Fearnhead P. Bayesian detection of abnormal segments in multiple time series // *Bayesian Analysis*. 2017. vol. 12. no. 1. pp. 193–218.
22. Huang J., Chen P., Lu L., Deng Y., Zou Q. WCDForest: a weighted cascade deep forest model toward the classification tasks // *Applied Intelligence*, 2023. vol. 53. no. 23. pp. 29169–29182. DOI: 10.1007/s10489-023-04794-z.
23. Tong W., Wang Y., Liu D. An Adaptive Clustering Algorithm Based on Local-Density Peaks for Imbalanced Data Without Parameters // *IEEE Transactions on Knowledge and Data Engineering*. 2023. vol. 35. no. 4. pp. 3419–3432.
24. Lu K.P., Chang S.T. Fuzzy maximum likelihood change-point algorithms for identifying the time of shifts in process data // *Neural Computing and Applications*. 2019. vol. 31. pp. 2431–2446.
25. Nevendra M., Singh P. Software defect prediction using deep learning // *Acta Polytechnica Hungarica*. 2021. vol. 18. no. 10. pp. 173–189.
26. Tallman E., West M. Bayesian predictive decision synthesis // *Journal of the Royal Statistical Society. Series B: Statistical Methodology*. 2024. vol. 86. no. 2. pp. 340–363.
27. Korkas K., Fryzlewicz P. Multiple change-point detection for non-stationary time series using wild binary segmentation. *Statistica Sinica*. 2017. vol. 27. pp. 287–311. DOI: 10.5705/ss.202015.0262.
28. Silva R.P., Zarpelão B.B., Cano A., Junior S.B. Time Series Segmentation Based on Stationarity Analysis to Improve New Samples Prediction // *Sensors*. 2021. vol. 21(21). DOI: 10.3390/s21217333.
29. Barzegar V., Laflamme S., Hu C., Dodson J. Multi-Time Resolution Ensemble LSTMs for Enhanced Feature Extraction in High-Rate Time Series // *Sensors*. 2021. vol. 21(6). DOI: 10.3390/s21061954.
30. Si S., Zhao J., Cai Z., Dui H. Recent advances in system reliability optimization driven by importance measures // *Frontiers of Engineering Management*. 2020. vol. 7. no. 3. pp. 335–358.

31. Xu S., Song Y., Hao X. A Comparative Study of Shallow Machine Learning Models and Deep Learning Models for Landslide Susceptibility Assessment Based on Imbalanced Data // *Forests*. 2022. vol. 13. no. 11. DOI: 10.3390/f13111908.
32. Лебедев И.С. Адаптивное применение моделей машинного обучения на отдельных сегментах выборки в задачах регрессии и классификации // *Информационно-управляющие системы*. 2022. № 3(118). С. 20–30.
33. Тихонов Д.Д., Лебедев И.С. Метод формирования сегментов информационной последовательности с использованием функционала качества моделей обработки // *Научно-технический вестник информационных технологий, механики и оптики*. 2024. Т. 24. № 3. С. 474–482.
34. Lebedev I.S., Sukhoparov M.E. Adaptive Learning and Integrated Use of Information Flow Forecasting Methods // *Emerging Science Journal*. 2023. vol. 7. no. 3. pp. 704–723.
35. Osipov V., Nikiforov V., Zhukova N., Miloserdov D. Urban traffic flows forecasting by recurrent neural networks with spiral structures of layers // *Neural Computing and Applications*. 2020. vol. 32. no. 18. pp. 14885–14897.
36. Lebedev I.S., Sukhoparov M.E. Improving the Quality Indicators of Multilevel Data Sampling Processing Models Based on Unsupervised Clustering // *Emerging Science Journal*. 2024. vol. 8. no. 1. pp. 355–371.
37. Jin H., Yin G., Yuan B., Jiang F. Bayesian hierarchical model for change point detection in multivariate sequences // *Technometrics*. 2022. vol. 64. no. 2. pp. 177–186.
38. Power Supply dataset. URL: <http://www.cse.fau.edu/~xqzhu/stream.html> (дата обращения: 16.05.2024).
39. Lu K.-P., Chang S.-T. An Advanced Segmentation Approach to Piecewise Regression Models // *Mathematics*. 2023. vol. 11(24). DOI: 10.3390/math11244959.
40. Energy generation dataset. URL: [https://www.kaggle.com/nicholasjhana/energy-consumption-generation-prices-and-weather/data?select=energy\\_dataset.csv](https://www.kaggle.com/nicholasjhana/energy-consumption-generation-prices-and-weather/data?select=energy_dataset.csv) (дата обращения: 16.05.2024).
41. Pima Indians Diabetes Database URL: <https://www.kaggle.com/datasets/uciml/pima-indians-diabetes-database> (дата обращения: 16.05.2024).
42. E-Commerce Data URL: <https://www.kaggle.com/datasets/carrie1/ecommerce-data> (дата обращения: 16.05.2024).

**Лебедев Илья Сергеевич** — д-р техн. наук, профессор, главный научный сотрудник, лаборатория интеллектуальных систем, Санкт-Петербургский Федеральный исследовательский центр Российской академии наук (СПб ФИЦ РАН). Область научных интересов: методы машинного обучения, представление и обработка слабоструктурированных данных, применение методов искусственного интеллекта в системах информационной безопасности. Число научных публикаций — 200. [isl\\_box@mail.ru](mailto:isl_box@mail.ru); 14-я линия В.О., 39, 199178, Санкт-Петербург, Россия; р.т.: +7(812)508-3311.

**Поддержка исследований.** Исследование выполнено за счет гранта Российского научного фонда № 25-21-00269, <https://rscf.ru/project/25-21-00269/>.

I. LEBEDEV

**ADAPTIVE REGRESSION MODEL CONSTRUCTION BASED  
ON THE FUNCTIONAL QUALITY ANALYSIS OF THE SEQUENCE  
SEGMENT PROCESSING**

*Lebedev I. Adaptive Regression Model Construction Based on the Functional Quality Analysis of the Sequence Segment Processing.*

**Abstract.** The article considers the problem of constructing an adaptive model aimed at improving the quality indicators of processing information sequences. In data processing techniques that have found application in many application areas, the applied analysis of observation objects is computationally resource-intensive and requires many iterations in case of changes in data properties. The article proposes a technique for selecting segments of an information sequence obtained in different ways, which differs in the use of the quality functional of regression models for processing subsequences. The sequences of observation objects received at the input of the model are divided by various specified segmentation algorithms. Pre-selected regression models are trained on each obtained segment and, depending on the obtained values of the calculated quality functional, the best models in terms of quality indicators are assigned to the segments. This allows us to form an aggregation model for data processing. Based on the experiment on model data and samples, the proposed technique is assessed. The values of the quality indicator MSE and MAE are obtained for different processing algorithms and with a different number of segments. The proposed method makes it possible to increase the MSE and MAE indicators by segmentation and assignment of regression models that have the best indicators on individual segments. The proposed solution is aimed at further improvement of ensemble methods. Its application allows to increase the efficiency of setting up basic algorithms in case of data property transformation and to improve the interpretability of results. The method can be used in developing models and methods for processing information sequences.

**Keywords:** machine learning, adaptive models, improving the quality of processing, regression models.

**References**

1. Chen H.Y., Chen C. Evaluation of Calibration Equations by Using Regression Analysis: An Example of Chemical Analysis. *Sensors*. 2022. vol. 22. no. 2. DOI: 10.3390/s22020447.
2. Schober P., Vetter T.R. Segmented Regression in an Interrupted Time Series Study Design. *Anesthesia and Analgesia*. 2021. vol. 132. no. 3. pp. 696–697.
3. Bozpolat E. Investigation of the self-regulated learning strategies of students from the faculty of education using ordinal logistic regression analysis. *Educational Sciences: Theory & Practice*. 2016. no. 16(1). pp. 301–318.
4. Jarantow S.W., Pisors E.D., Chiu M.L. Introduction to the use of Linear and Nonlinear Regression Analysis in Quantitative Biological Assays. *Current Protocols*. 2023. no. 3. DOI: 10.1002/cpz1.801.
5. Britzger D. The Linear Template Fit. *The European Physical Journal C*. 2022. vol. 82(8). DOI: 10.1140/epjc/s10052-022-10581-w.
6. Perperoglou A., Sauerbrei W., Abrahamowicz M., Schmid M. A review of spline function procedures in R. *BMC Medical Research Methodology*. 2019. vol. 19. pp. 1–16.
7. Ren J., Tapert S., Fan C.C., Thompson W.K. A semi-parametric Bayesian model for semi-continuous longitudinal data. *Statistics in Medicine*. 2022. vol. 41. no. 13. pp. 2354–2374.

8. Taye M.M. Theoretical Understanding of Convolutional Neural Network: Concepts, Architectures, Applications, Future Directions. *Computation*. 2023. vol. 11. no. 3. DOI: 10.3390/computation11030052.
9. Kolmogorov A.N. [On the representation of continuous functions of many variables by superposition of continuous functions of one variable and addition]. *Reports of the USSR Academy of Sciences – Doklady AN SSSR*. 1957. vol. 114. no. 5. pp. 953–956.
10. Giroisi F., Poggio T. Representation Properties of Networks: Kolmogorov’s Theorem is Irrelevant. *Neural Computation*. 1989. vol. 4. no. 1. pp. 465–469.
11. Parhi R., Nowak R.D. Banach Space Representer Theorems for Neural Networks and Ridge Splines. *Journal of Machine Learning Research*. 2021. vol. 22(1). pp. 1960–1999.
12. Marques H.O., Swersky L., Sander J., Campello R.J., Zimek A. On the evaluation of outlier detection and one-class classification: a comparative study of algorithms, model selection, and ensembles. *Data Mining and Knowledge Discovery*. 2023. vol. 37. no. 4. pp. 1473–1517.
13. Li Y., Guo X., Lin W., Zhong M., Li Q., Liu Z., Zhong W., Zhu Z. Learning dynamic user interest sequence in knowledge graphs for click-through rate prediction. *IEEE Transactions on Knowledge and Data Engineering*. 2023. vol. 35. no. 1. pp. 647–657.
14. Rinaldo A., Wang D., Wen Q., Willett R., Yu Y. Localizing changes in highdimensional regression models. *The 24th International Conference on Artificial Intelligence and Statistics*. 2021. pp. 2089–2097.
15. Aue A., Rice G., Sönmez O. Detecting and dating structural breaks in functional data without dimension reduction. *Journal of the Royal Statistical Society. Series B, Statistical Methodology*. 2018. vol. 80. no. 3. pp. 509–529.
16. Datta A., Zou H., Banerjee S. Bayesian high-dimensional regression for change point analysis. *Statistics and its Interface*. 2019. vol. 12. no. 2. pp. 253–264. DOI: 10.4310/SII.2019.v12.n2.a6.
17. Melnyk I., Banerjee A. A spectral algorithm for inference in hidden semi-Markov models. *Journal of Machine Learning Research*. 2017. vol. 18. no. 35. pp. 1–39.
18. Haynes K., Fearnhead P., Eckley I.A. A computationally efficient nonparametric approach for changepoint detection. *Statistics and Computing*. 2017. vol. 27. pp. 1293–1305. DOI: 10.1007/s11222-016-9687-5.
19. Muggeo V. Estimating regression models with unknown break-points. *Statistics in Medicine*. 2003. vol. 22(19). pp. 3055–3071.
20. Lu K.P., Chang S.T. A fuzzy classification approach to piecewise regression models. *Applied Soft Computing Journal*. 2018. vol. 69. pp. 671–688.
21. Bardwell L., Fearnhead P. Bayesian detection of abnormal segments in multiple time series. *Bayesian Analysis*. 2017. vol. 12. no. 1. pp. 193–218.
22. Huang J., Chen P., Lu L., Deng Y., Zou Q. WCDForest: a weighted cascade deep forest model toward the classification tasks. *Applied Intelligence*, 2023. vol. 53. no. 23. pp. 29169–29182. DOI: 10.1007/s10489-023-04794-z.
23. Tong W., Wang Y., Liu D. An Adaptive Clustering Algorithm Based on Local-Density Peaks for Imbalanced Data Without Parameters. *IEEE Transactions on Knowledge and Data Engineering*. 2023. vol. 35. no. 4. pp. 3419–3432.
24. Lu K.P., Chang S.T. Fuzzy maximum likelihood change-point algorithms for identifying the time of shifts in process data. *Neural Computing and Applications*. 2019. vol. 31. pp. 2431–2446.
25. Nevendra M., Singh P. Software defect prediction using deep learning. *Acta Polytechnica Hungarica*. 2021. vol. 18. no. 10. pp. 173–189.
26. Tallman E., West M. Bayesian predictive decision synthesis. *Journal of the Royal Statistical Society. Series B: Statistical Methodology*. 2024. vol. 86. no. 2. pp. 340–363.



27. Korkas K., Fryzlewicz P. Multiple change-point detection for non-stationary time series using wild binary segmentation. *Statistica Sinica*. 2017. vol. 27. pp. 287–311. DOI: 10.5705/ss.202015.0262.
28. Silva R.P., Zarpelão B.B., Cano A., Junior S.B. Time Series Segmentation Based on Stationarity Analysis to Improve New Samples Prediction. *Sensors*. 2021. vol. 21(21). DOI: 10.3390/s21217333.
29. Barzegar V., Laflamme S., Hu C., Dodson J. Multi-Time Resolution Ensemble LSTMs for Enhanced Feature Extraction in High-Rate Time Series. *Sensors*. 2021. vol. 21(6). DOI: 10.3390/s21061954.
30. Si S., Zhao J., Cai Z., Dui H. Recent advances in system reliability optimization driven by importance measures. *Frontiers of Engineering Management*. 2020. vol. 7. no. 3. pp. 335–358.
31. Xu S., Song Y., Hao X. A Comparative Study of Shallow Machine Learning Models and Deep Learning Models for Landslide Susceptibility Assessment Based on Imbalanced Data. *Forests*. 2022. vol. 13. no. 11. DOI: 10.3390/f13111908.
32. Lebedev I.S. [Adaptive application of machine learning models on separate segments of a data sample in regression and classification problems]. *Informatsionno-upravliaushchie sistemy – Information and Control Systems*. 2022. no. 3(118). pp. 20–30.
33. Tikhonov D.D., Lebedev I.S. [Method for generating information sequence segments using the quality functional of processing models]. *Nauchno-tehnicheskij vestnik informacionnyh tehnologij, mehaniki i optiki – Scientific and Technical Journal of Information Technologies, Mechanics and Optics*. 2024. vol. 24. no. 3. pp. 474–482.
34. Lebedev I.S., Sukhoparov M.E. Adaptive Learning and Integrated Use of Information Flow Forecasting Methods. *Emerging Science Journal*. 2023. vol. 7. no. 3. pp. 704–723.
35. Osipov V., Nikiforov V., Zhukova N., Miloserdov D. Urban traffic flows forecasting by recurrent neural networks with spiral structures of layers. *Neural Computing and Applications*. 2020. vol. 32. no. 18. pp. 14885–14897.
36. Lebedev I.S., Sukhoparov M.E. Improving the Quality Indicators of Multilevel Data Sampling Processing Models Based on Unsupervised Clustering. *Emerging Science Journal*. 2024. vol. 8. no. 1. pp. 355–371.
37. Jin H., Yin G., Yuan B., Jiang F. Bayesian hierarchical model for change point detection in multivariate sequences. *Technometrics*. 2022. vol. 64. no. 2. pp. 177–186.
38. Power Supply dataset. Available at: <http://www.cse.fau.edu/~xqzhu/stream.html> (accessed 16.05.2024).
39. Lu K.-P., Chang S.-T. An Advanced Segmentation Approach to Piecewise Regression Models. *Mathematics*. 2023. vol. 11(24). DOI: 10.3390/math11244959.
40. Energy generation dataset. Available at: [https://www.kaggle.com/nicholasjhana/energy-consumption-generation-prices-and-weather/data?select=energy\\_dataset.csv](https://www.kaggle.com/nicholasjhana/energy-consumption-generation-prices-and-weather/data?select=energy_dataset.csv) (accessed 16.05.2024).
41. Pima Indians Diabetes Database. Available at: <https://www.kaggle.com/datasets/uciml/pima-indians-diabetes-database> (accessed 16.05.2024).
42. E-Commerce Data. Available at: <https://www.kaggle.com/datasets/carrie1/ecommerce-data> (accessed 16.05.2024).

**Lebedev Ilya** — Ph.D., Dr.Sci., Professor, Chief scientific officer, Laboratory of intelligent systems, St. Petersburg Federal Research Center of the Russian Academy of Sciences (SPC RAS). Research interests: machine learning methods, representation and processing of weakly structured data, application of artificial intelligence methods in information security systems. The number of publications — 200. [isl\\_box@mail.ru](mailto:isl_box@mail.ru); 39, 14-th Line V.O., 199178, St. Petersburg, Russia; office phone: +7(812)508-3311.

**Acknowledgements.** This research is supported by RSF (grant № 25-21-00269, <https://rscf.ru/project/25-21-00269/>).

К.В. КРОТОВ  
**ДВУХУРОВНЕВАЯ ОПТИМИЗАЦИЯ РАСПРЕДЕЛЕНИЯ  
ЗАДАНИЙ ПО ПАКЕТАМ И РАСПИСАНИЙ ИХ ВЫПОЛНЕНИЯ  
В КОНВЕЙЕРНЫХ СИСТЕМАХ С БУФЕРАМИ  
ОГРАНИЧЕННЫХ РАЗМЕРОВ**

*Кротов К.В. Двухуровневая оптимизация распределения заданий по пакетам и расписаний их выполнения в конвейерных системах с буферами ограниченных размеров.*

**Аннотация.** Существующие на данный момент математические модели и алгоритмы обеспечивают оптимизацию расписаний выполнения единичных заданий либо фиксированных пакетов заданий на приборах конвейерных систем, содержащих буферы ограниченных размеров. Эти модели и алгоритмы не позволяют осуществлять поиск оптимальных решений по группированию однотипных заданий в пакеты и по последовательностям пакетов для реализации операций с ними на приборах конвейерных систем. Повышение эффективности использования ресурсов конвейерных систем достигается путем оптимизации решений по группированию однотипных заданий в пакеты и по последовательностям пакетов для проведения операций с ними. Решение этой задачи выполнено в работе посредством привлечения подхода, реализующего двухуровневую оптимизацию, который позволяет сформировать иерархию подзадач поиска эффективных решений. Привлечение упомянутого подхода предполагает разработку математических моделей иерархических игр, позволяющих идентифицировать эффективные решения рассматриваемого вида. Осуществлено построение двух математических моделей иерархических игр, использование которых позволяет реализовать оптимизацию составов пакетов на верхнем уровне ведущим игроком и оптимизацию расписаний выполнения пакетов в конвейерных системах на нижнем уровне ведомым игроком. Способ определения оптимальных решений каждым из игроков предусматривает заданный в игре порядок ходов и обмен решений между ними в процессе игры. Первая математическая модель иерархической игры реализует определение эффективных решений при учете простоев обрабатывающих приборов в процессе реализации операций с пакетами. Вторая математическая модель игры реализует определение эффективных решений при учете общего времени ожидания буферами размещения в них заданий, с которыми завершились операции на предшествующих приборах. Для этого сформированы выражения, позволяющие определять простои буферов в ожидании готовности заданий из пакетов к размещению на основе временных характеристик процессов выполнения операций с ними на приборах рассматриваемых систем. В основу алгоритма определения оптимальных решений по порядкам осуществления операций с пакетами на нижнем уровне в каждой из иерархических игр положена разработанная математическая модель процессов реализации действий с пакетом в указанных системах и соответствующий алгоритм моделирования. Реализация рассматриваемого подхода к оптимизации позволила получить результаты, которые показали, что использование буферов позволяет значительно повысить эффективность процессов осуществления операций с пакетами на приборах рассматриваемых систем; увеличение размеров промежуточных буферов позволяет в большей степени повысить эффективность указанных процессов при значительных неоднородностях значений временных параметров, их характеризующих; использование первой модели иерархической игры позволяет добиться большего повышения эффективности процессов в сравнении со второй моделью.

**Ключевые слова:** конвейерные системы, расписания, буферы ограниченного размера, пакеты заданий.

---

**1. Введение.** В различных сферах практической деятельности требуется реализация операций с заданиями, входящими в разные наборы (принадлежащие разным типам), в конвейерных системах (КС). Задания одного типа имеют одинаковое время их выполнения на приборах КС, а также одинаковое время переналадок приборов на реализацию операций с ними на этих приборах. Производительность процессов выполнения заданий разных типов в КС увеличивается за счет их группирования в пакеты. Тогда требуется осуществить распределение заданий по пакетам и определить последовательность пакетов для их выполнения на приборах КС при максимально эффективном использовании их (приборов) ресурсов. Ранее автором был предложен метод многоуровневой оптимизации расписаний [1], использование которого позволяет решить рассматриваемую задачу. Этот метод предполагает формирование иерархии подзадач поиска оптимальных решений по составам пакетов и расписаниям их выполнения на приборах КС, которые образуют обобщенное решение рассматриваемой задачи. С использованием этого метода реализовано решение ряда задач оптимизации составов пакетов заданий и расписаний их выполнения на приборах КС [1 – 3]. Решение этих задач выполнено в предположении о наличии между приборами буферов неограниченных размеров, в которых находятся задания в случае, если приборы не готовы к реализации операций с ними. Применение указанного подхода возможно также в случае, когда в состав конвейерных систем включены промежуточные буферы ограниченных размеров. Включение в состав этих систем буферов ограниченных размеров позволяет уменьшить зависимость производительности процессов реализации операций с заданиями в таких системах от неоднородностей времени выполнения заданий на разных приборах КС, а также от неоднородностей времени переналадок приборов. Тогда актуальным является применение метода иерархической оптимизации, предусматривающего построение математических моделей иерархических игр, участники которых (ведущий игрок на верхнем уровне и ведомый игрок на нижнем уровне) реализуют оптимизацию решений по группированию заданий в пакеты (ведущий игрок) и расписаниям выполнения пакетов заданий на приборах КС (ведомый игрок). Оптимизация расписаний выполнения пакетов заданий на приборах КС осуществляется с использованием модели многостадийных процессов реализации

операций с ними в этих системах, в состав которых входят буферы ограниченных размеров [4].

Работа [5] является одной из первых работ, в которой синтезирована математическая модель процесса реализации операций с единичными заданиями (ЕЗ) в системе с двумя приборами и одним промежуточным буфером ограниченного размера. Модель процесса осуществления операций с единичными заданиями в системах с произвольным количеством приборов и промежуточными буферами ограниченных размеров предложена в [6]. Для определения последовательностей заданий с целью реализации операций с ними на приборах рассматриваемых систем в [6] использованы эвристические правила. Математическая модель *mixed-integer linear programming* (MILP) определения последовательностей заданий для осуществления операций в потоковой системе, включающей обрабатывающие устройства с параллельно функционирующими приборами и предваряющие их буферы (*Batch Processing Machine* (BPM)), предложена в [7]. Использование этой модели позволяет сформировать оптимальное решение по назначению прибора, входящего в BPM, для каждого задания, а также сформировать последовательности выполнения заданий на этих приборах. В [8] решена аналогичная задача идентификации эффективных решений для ЕЗ с использованием разработанного эвристического алгоритма. В [9] оценка эффективности решений осуществляется на основе значений, вычисляемых с использованием предложенной в работе математической модели, аналогичной модели в [6]. Поиск решений по последовательностям ЕЗ для проведения операций с ними осуществляется в [9] с использованием эволюционного алгоритма распределения, реализующего генерацию новых решений на основе вероятностных моделей, полученных с учетом родительских решений. В [10] предложена модель MILP, позволяющая оптимизировать решения по последовательностям ЕЗ для их выполнения на приборах КС, в которые включены буферы заданных размеров. Модель MILP в [11] позволяет определять эффективные последовательности заданий, выполняемых в системе, содержащей: BPM, машины последовательной обработки пакетов заданий, промежуточные буферы заданных размеров. В [12] рассматривается модель MILP, используемая для определения последовательностей заданий, выполняемых в системах с параллельными маршрутами, общими обрабатывающими приборами, включенными в оба маршрута, и единичными промежуточными буферами между приборами. В [13] рассматривается КС, в которой накопитель имеется только за

последним прибором. Модель MILP в [13] позволяет определять оптимальные решения по последовательностям выполнения заданий на приборах указанных систем с учетом ограниченного размера накопителя и директивных сроков извлечения заданий из него. Модель MILP в [13] используется только для определения допустимости сформированных решений, а поиск оптимальных решений осуществляется с использованием вероятностного метода локального поиска с запретами. В [14] маршруты выполнения ЕЗ в конвейерной системе предусматривают возможность их (заданий) повторной обработки на приборах, а сама система включает буферы ограниченных размеров между приборами. В [15] предложены модели MILP оптимизации решений по последовательностям ЕЗ для реализации операций с ними в системе с блокированием приборов при отсутствии возможности у этих приборов приступить к выполнению заданий. Система в [16] имеет два параллельных маршрута осуществления операций с ЕЗ. Оптимизация последовательностей заданий осуществляется с использованием алгоритма дискретной китовой стаи (DWSA).

В [17] рассматривается задача оптимизации последовательностей выполнения ЕЗ на приборах BPM, которые предусматривают буферы ограниченных размеров, с учетом проведения предварительного технического обслуживания (ПТО) приборов, включенных в BPM. С учетом стохастических параметров функционирования приборов, характеризующих отказы, осуществляется определение вида последовательностей выполнения ЕЗ на приборах и периодичность ПТО.

В [18] рассматривается двух приборная система Flexible Flow Shop (FFS), содержащая буфер, который выделяется заданиям на все время их обработки. Оптимизационная задача состоит в определении пары устройств на первой и второй стадиях, назначаемых для выполнения конкретных ЕЗ с учетом буфера ограниченной емкости (длительности выполнения различных заданий на приборах первой и второй стадий являются единичными, буфер может быть разделен между несколькими заданиями). С учетом введенных условий в [18] предложен полиномиальный алгоритм оптимизации назначения заданий на пары устройств и порядков их выполнения на них.

В [19] рассматривается система сбора данных с распределенных узлов на центральную обрабатывающую станцию, на которой размер буфера хранения данных ограничен. Эта система проинтерпретирована в [19] как двух приборная система Flow Shop с общим буфером,

реализующая обработку ЕЗ (соответствующая часть буфера выделяется заданию в момент времени начала передачи данных от узла на станцию и освобождается в момент времени окончания их обработки этой станцией). В [19] построены эвристики для определения эффективных решений с учетом следующих предположений: 1) время выполнения операции с ЕЗ на первой стадии пропорционально размеру размещаемого в буфере задания; 2) время выполнения операции с ЕЗ на второй стадии не меньше времени выполнения операций на первой стадии.

Продолжением работ [18, 19] является работа [20], в которой также рассмотрена двух приборная система Flow Shop с общим буфером ограниченного размера, используемым заданиями в течение всего времени их обработки приборами этой системы. В [20] предложен полиномиальный алгоритм, позволяющий получить оптимальные решения за ограниченное время при выполнении следующих условий: 1) потребности в буфере со стороны заданий не обязательно пропорциональны длительностям их выполнения на первой стадии; 2) минимальное время обработки заданий на первой стадии не меньше наибольшего времени обработки заданий на второй стадии (либо минимальное время обработки заданий на второй стадии не меньше наибольшего времени обработки заданий на первой).

В [21] решена задача управления комплектацией заказов на предприятиях Интернет-торговли (маркет-плейсах). Система, рассматриваемая в [21], интерпретируется как Flow Shop, в которой обрабатываемыми приборами выступают комплектовщики, оперирующими с товарами, размещенными в соответствующих зонах, соединенных конвейером для перемещения контейнеров с заказами (контейнеры интерпретируются как задания). Контейнеры с заказами клиентов накапливаются в буфере соответствующего комплектовщика. В том случае, если буфер заполнен, конвейер останавливается. В статье разработана модель MILP оптимизации порядка контейнеров, соответствующих формируемому в рассматриваемой системе заказам (порядок контейнеров может быть различным для разных комплектовщиков (различных приборов системы Flow Shop)).

Решение задачи, аналогичной рассмотренной в [17], осуществлено в [22] при учете ПТО транспортеров, реализующих размещение ЕЗ в накопителях и перемещение ЕЗ из накопителей на последующие приборы. Для оптимизации расписаний выполнения ЕЗ в КС и периодичности проведения ПТО транспортеров в [22] предложена модель MILP.

В работах [23, 24] решается задача оптимизации расписания обработки данных приложениями и распределения памяти (пространства буферов) между приложениями, используемой ими для обмена данными. В качестве модели взаимодействия приложений через буферы с постоянными количествами формируемых и потребляемых данных рассматривается граф синхронного потока данных. Оптимизация решений осуществляется путем имитационного моделирования рассматриваемых процессов, при реализации которого учитываются различные параметры функционирования приложений, характеристики данных и обрабатывающих систем.

В работе [25] решается задача оптимизации расписаний в транспортно-логистической сфере деятельности, связанная с организацией рейсов авиакомпаний и функционированием аэропортов. Рейсы, выполняемые разными воздушными судами, интерпретируются как единичные задания, а аэропорты посадки/вылета рассматриваются как буферы для накопления пассажиров. Расписания рейсов воздушных судов оптимизируются с точки зрения возможности реализации стыковочных рейсов с учетом размеров буферов (количества пассажиров, которые могут быть размещены в аэропорту) с использованием имитационной модели рассматриваемых процессов.

Математические модели и методы оптимизации, рассмотренные в [5 – 25], позволяют обеспечить решение задач поиска оптимальных расписаний выполнения единичных заданий в обрабатывающих системах разного вида, содержащих буферы между приборами. Использование этих моделей для оптимизации решений по составам пакетов заданий (ПЗ) и расписаниям их выполнения на приборах систем рассматриваемого вида является не возможным.

В [26] разработан способ определения последовательностей пакетов, эффективных с точки зрения времени окончания действий с ними в двух приборной системе, имеющей в составе промежуточный буфер ограниченного размера. При этом составы пакетов являются неизменными, то есть неизменно количество заданий в них (рассматриваются фиксированные пакеты, составы которых не оптимизируются). Задача представлена как задача коммивояжера, решение которой обеспечивается точными методами. В [27] решается аналогичная задача: фиксированные пакеты, двух приборная система, имеющая в составе буфер. Модель, введенная в [27], представляет собой модифицированную модель из [26]. В ней учтено время заполнения конвейера заданиями из пакетов для их выполнения на приборах рассматриваемой системы. В [28] исследуется проблема

определения последовательностей пакетов (с неизменным количеством заданий в них) для их выполнения в системе с двумя приборами и буфером между ними.

В работе [29] оптимизируются составы партий выпускаемой продукции различного вида и расписания ее (продукции) производства в двух приборной (производство и упаковка) системе Flow Shop в химической промышленности с общим буфером (складом) для ее хранения. Под партиями подразумевается количество продукции, размещаемой в резервуарах, объемы резервуаров являются заданными, в каждом резервуаре размещается продукция только одного вида. Оптимизируется количество продукции в резервуарах заданных объемов и порядок выпуска этой продукции (с учетом ограниченного объема хранилища для резервуаров и требования минимизации количества их очисток). Для оптимизации решений разработан полиномиальный алгоритм и эвристические схемы.

В работе [30] реализуется оптимизация порядка обработки пакетов заданий разных типов в системах Flexible Flow Shop (FFS), в которых на каждой стадии находятся машины пакетной обработки (Batch Processing Machine (BPM)), включающие в себя заданное количество приборов, выполняющих задания параллельно. Между BPM в систему включены буферы для хранения пакетов заданий. В сформированной в [30] модели количество пакетов заданий каждого типа является заданным, составы пакетов не оптимизируются (предполагается, что количество заданий в пакетах не превышает количество приборов в BPM). Наличие промежуточных буферов в системе интерпретируется в модели путем задания условия блокировки предшествующего прибора в случае, если буфер заполнен. Так как составы пакетов не оптимизируются, а определяется порядок выполнения заданного количества пакетов заданий каждого типа, то простои приборов, входящих в BPM, не фиксируются. Также не определяется эффективность использования ограниченных ресурсов хранения заданий из пакетов. Решение задачи оптимизации порядка выполнения ПЗ в системе FFS с промежуточными буферами решается в [30] с использованием генетических алгоритмов.

В [31] решается задача оптимизации составов ПЗ разных типов и порядка их выполнения в FFS, состоящих из машин с параллельно функционирующими приборами (BPM) и промежуточными буферами между ними. Размеры партий каждого продукта задаются равными, оптимизируется количество заданий в них. Также оптимизируется их (партий) размещение в буферах, состоящих из секций. Для



оптимизации решений рассматриваемого вида применены генетические алгоритмы.

Таким образом, в [26 – 28] рассматриваются способы определения оптимальных решений по последовательностям выполнения фиксированных ПЗ в двух приборных системах. При этом ограниченным является как количество пакетов, так и количество обрабатывающих приборов в системах. В работе [29] оптимизируется количество продукции химической промышленности (количество продукции в партиях), разливаемой по резервуарам при их размещении с учетом размеров в ограниченном буфере. При этом обрабатывающая система является двух приборной, а буфер представляет собой хранилище общего пользования. В работах [30, 31] рассматриваются системы FFS, включающие BPM и промежуточные буферы ограниченных размеров. Модель взаимодействия приборов друг с другом и приборов с буферами с точки зрения передачи между ними пакетов заданий не соответствует рассматриваемой системе конвейерного типа с машинами последовательной обработки (с Job Processing Machine (JPM)). При этом в [30, 31] не решается задача общего вида, предполагающая возможность определения при оптимизации решений произвольного количества пакетов и произвольного количества заданий в них.

Рассмотренные в [26 – 31] модели не учитывают возможность передачи отдельных заданий из пакетов между приборами непосредственно в момент времени окончания реализации операций с ними на этих приборах (возможность перемещения каждого задания на следующий прибор (либо в буфер) сразу после завершения операций с ним на предыдущем приборе). Перемещение пакетов между приборами в этих работах реализуется только в момент времени окончания выполнения всех заданий, включенных в них.

**2. Математическая модель процессов выполнения пакетов заданий в конвейерных системах, содержащих буферы ограниченных размеров.** В работе рассматриваются многостадийные процессы выполнения заданий из  $N$  наборов в конвейерных системах, количество приборов в которых  $L > 2$ . Окончание на приборах операций с заданиями одного типа и переход к осуществлению операций с заданиями другого типа связаны с переналадкой приборов. В состав системы включены промежуточные буферы ограниченных размеров, следующие за соответствующими им устройствами обработки. Задания, с которыми завершилось выполнение операций на приборе, могут быть размещены в буфере, следующим за ним. Прибор,

предшествующий некоторому буферу, блокируется в том случае, если этот буфер заполнен.

Модели процессов выполнения пакетов заданий в конвейерных системах, модели иерархических игр оптимизации решений по группированию заданий в пакеты и по последовательностям выполнения пакетов заданий в КС формируются с использованием обозначений [1, 4]:  $i$  – тип задания ( $i = \overline{1, N}$ );  $n^i$  – количество заданий  $i$ -го типа ( $i = \overline{1, N}$ );  $l$  – индекс прибора КС ( $l = \overline{1, L}$ );  $\pi^l$  – последовательность выполнения пакетов заданий на  $l$ -м приборе КС ( $l = \overline{1, L}$ );  $\mathbf{m} = (m_1, m_2, \dots, m_N)^T$  – вектор количества ПЗ каждого типа;  $\mathbf{A} = \|a_{ih}\|_{N \times m^{\max}}$  – матрица, каждый элемент которой соответствует количеству заданий  $i$ -го типа в  $h$ -х пакетах;  $N_p$  – количество ПЗ;  $\mathbf{T} = \|t_{li}\|_{L \times N}$  – матрица времен выполнения заданий  $i$ -х типов на  $l$ -х приборах;  $\mathbf{T}^l = \|t_{il}^l\|_{N \times N}$  ( $l = \overline{1, L}$ ) – матрицы времен переналадки приборов;  $j$  – индекс позиции пакета в последовательности  $\pi^l$  ( $l = \overline{1, L}$ );  $\mathbf{P} = \|p_{ij}\|_{N \times N_p}$  – матрица позиций ПЗ  $i$ -х типов в последовательностях  $\pi^l$  их выполнения на  $l$ -х приборах КС ( $p_{ij} = 1$ , если ПЗ  $i$ -го типа размещен в последовательностях  $\pi^l$  в  $j$ -й позиции;  $p_{ij} = 0$ , если ПЗ  $i$ -го типа не размещен в последовательности  $\pi^l$  в  $j$ -й позиции);  $\mathbf{R} = \|r_{ij}\|_{N \times N_p}$  – матрица количества заданий в пакетах, занимающих в последовательностях  $\pi^l$   $j$ -е позиции;  $N_j$  – количество заданий в пакетах, которые занимают в  $\pi^l$   $j$ -е позиции ( $j = \overline{1, N_p}$ );  $\mathbf{T}^{ol} = \|t_{jq}^{ol}\|_{N_p \times Q}$  ( $l = \overline{1, L}$ ) – матрицы начала операций с  $q$ -ми заданиями ( $q = \overline{1, N_j}$ ) из пакетов, занимающих в  $\pi^l$  в  $j$ -е позиции ( $Q = \max_j(N_j)$ ) на  $l$ -х приборах. Через  $b$  обозначен размер буферов.

Также при синтезе моделей иерархических игр использованы обозначения, позволяющие охарактеризовать использование ограниченного ресурса хранения заданий из пакетов (использование

буферов):  $t_{jq}^{01l}$  – моменты времени, в которые задания из пакетов добавляются в буферы, следующие за  $l$ -ми приборами;  $t_{jq}^{012}$  – моменты времени, в которые задания из пакетов удаляются из буферов, следующих за  $l$ -ми приборами.

В работе [4] предложены математические модели многостадийных процессов выполнения заданий, распределенных по пакетам, на приборах конвейерных системах, содержащих два и три обрабатывающих прибора, а также промежуточные буферы (модели обеспечивающие вычисление значений  $t_{jq}^{01}$  ( $j = \overline{1, N_p}$ ;  $q = \overline{1, N_j}$ ;  $l = \overline{1, L}$ ) для заданий в соответствующих пакетах). Полученные математические модели обобщены для произвольного количества  $L$  приборов. Обобщенная математическая модель рассматриваемого вида представлена следующими выражениями:

1. выражения для вычисления значений  $t_{lq}^{01}$  ( $q = \overline{1, N_1}$ ):

$$- q=1: t_{l1}^{01} = 0; \quad (1)$$

$$- 1 < q \leq b+1: t_{lq}^{01} = t_{l,(q-1)}^{01} + \sum_{s=1}^N t_{1s} \cdot P_{s1}; \quad (2)$$

$$- b+1 < q \leq N_1: t_{lq}^{01} = \max(t_{l,(q-1)}^{01} + \sum_{s=1}^N t_{1s} \cdot P_{s1}; t_{l,(q-b)}^{02}); \quad (3)$$

2. выражения для вычисления значений  $t_{jq}^{01}$  ( $q = \overline{1, N_j}$ ;  $j = \overline{1, N_p}$ ):

$$- q=1: t_{j1}^{01} = \max(t_{(j-1), N_{j-1}}^{01} + \sum_{s=1}^N t_{1s} \cdot P_{s,(j-1)} + t_{i_{j-1}, i_j}^1; t_{(j-1), (N_{j-1}-b+1)}^{02}); \quad (4)$$

$$- 1 < q \leq b: t_{jq}^{01} = \max(t_{j,(q-1)}^{01} + \sum_{s=1}^N t_{1s} \cdot P_{s,j}; t_{(j-1), (N_{j-1}-b+q)}^{02}); \quad (5)$$

$$- \quad b+1 \leq q \leq N_j : t_{jq}^{0l} = \max(t_{j,(q-1)}^{0l} + \sum_{s=1}^N t_{ls} \cdot P_{sj} ; t_{j,(q-b)}^{02}) ; \quad (6)$$

3. выражения для вычисления значений  $t_{lq}^{0l}$  ( $q = \overline{1, N_1}$  ;  $l = \overline{2, L-1}$ ):

$$- \quad q=1 : t_{l1}^{0l} = t_{l1}^{0l-1} + \sum_{s=1}^N t_{(l-1),s} \cdot P_{s1} ; \quad (7)$$

$$- \quad 1 < q \leq b+1 : t_{lq}^{0l} = \max(t_{lq}^{0l-1} + \sum_{s=1}^N t_{(l-1),s} \cdot P_{s1} ; t_{l,(q-1)}^{0l} + \sum_{s=1}^N t_{ls} \cdot P_{s1}) ; \quad (8)$$

-  $b+1 < q \leq N_j$  :

$$t_{lq}^{0l} = \max(t_{lq}^{0l-1} + \sum_{s=1}^N t_{(l-1),s} \cdot P_{s1} ; t_{l,(q-1)}^{0l} + \sum_{s=1}^N t_{ls} \cdot P_{s1} ; t_{l,(q-b)}^{0l+1}) ; \quad (9)$$

4. выражения для вычисления значений  $t_{jq}^{0l}$  ( $q = \overline{1, N_j}$  ;  $j = \overline{2, N_p}$  ;  $l = \overline{2, L-1}$ ):

-  $q=1$  :

$$t_{j1}^{0l} = \max(t_{j1}^{0l-1} + \sum_{s=1}^N t_{(l-1),s} \cdot P_{sj} ; t_{(j-1),N_{j-1}}^{0l} + \sum_{s=1}^N t_{s,(j-1)} \cdot P_{s,(j-1)} + t_{i_{j-1},i_j}^l ; t_{(j-1),(N_{j-1}-b+1)}^{0l+1}) ; \quad (10)$$

-  $1 < q \leq b$  :

$$t_{jq}^{0l} = \max(t_{jq}^{0l-1} + \sum_{s=1}^N t_{(l-1),s} \cdot P_{sj} ; t_{j,(q-1)}^{0l} + \sum_{s=1}^N t_{ls} \cdot P_{sj} ; t_{(j-1),(N_{j-1}-b+q)}^{0l+1}) ; \quad (11)$$

-  $b+1 \leq q \leq N_j$  :

$$t_{jq}^{0l} = \max(t_{jq}^{0l-1} + \sum_{s=1}^N t_{(l-1),s} \cdot P_{sj} ; t_{j,(q-1)}^{0l} + \sum_{s=1}^N t_{ls} \cdot P_{sj} ; t_{j,(q-b)}^{0l+1}) ; \quad (12)$$

5. выражения для вычисления значений  $t_{lq}^{0L}$  ( $q = \overline{1, N_1}$ ):

$$- q=1: t_{l1}^{0L} = t_{l1}^{0L-1} + \sum_{s=1}^N t_{(L-1),s} \cdot p_{s1} = \sum_{l=1}^{L-1} \sum_{s=1}^N t_{ls} \cdot p_{s1}; \quad (13)$$

$$- 1 < q \leq N_1: t_{lq}^{0L} = \max(t_{lq}^{0L-1} + \sum_{s=1}^N t_{(L-1),s} \cdot p_{s1}; t_{1,(q-1)}^{0L} + \sum_{s=1}^N t_{Ls} \cdot p_{s1}); \quad (14)$$

6. выражения для вычисления значений  $t_{jq}^{0L}$  ( $q = \overline{1, N_j}, j = \overline{2, N_p}$ ):

$$- q=1: t_{j1}^{0L} = \max(t_{j1}^{0L-1} + \sum_{s=1}^N t_{(j-1),s} \cdot p_{sj}; t_{(j-1),N_{j-1}}^{0L} + \sum_{s=1}^N t_{Ls} \cdot p_{s,(j-1)} + t_{i_{j-1},i_j}^L); \quad (15)$$

$$- 1 < q \leq N_j: t_{jq}^{0L} = \max(t_{jq}^{0L-1} + \sum_{s=1}^N t_{(j-1),s} \cdot p_{sj}; t_{j,(q-1)}^{0L} + \sum_{s=1}^N t_{Ls} \cdot p_{s1}). \quad (16)$$

Алгоритм математического моделирования хода многостадийного процесса выполнения заданий, распределенных по пакетам, включенных в последовательности  $\pi^l$  в  $j$ -х позициях, предусматривает непосредственную интерпретацию выражений (1)-(16) для вычисления значений  $t_{jq}^{0L}$  ( $q = \overline{1, N_j}, j = \overline{1, N_p}; l = \overline{1, L}$ ).

Последовательность шагов алгоритма моделирования рассматриваемых процессов в КС с буферами представлена в [4].

Алгоритм определения оптимальных решений по последовательностям выполнения пакетов заданий на  $l$ -х приборах КС предусматривает изменение на каждой итерации вида матриц  $\mathbf{P}$  и  $\mathbf{R}$ , а также определение значений  $t_{jq}^{0L}$  временных характеристик рассматриваемого многостадийного процесса [1]. В соответствии с видом системы вычисление временных характеристик процессов выполнения ПЗ на ее приборах – значений  $t_{jq}^{0L}$  ( $j = \overline{1, N_p}; q = \overline{1, N_j}; l = \overline{1, L}$ ), – соответствующих матрицам  $\mathbf{P}$  и  $\mathbf{R}$ , осуществляется путем использования модели (1)-(16),

учитывающей наличие промежуточных буферов, а не модели, предложенной в [1].

**3. Математические модели иерархических игр оптимизации составов пакетов заданий и последовательностей пакетов для их выполнения на приборах конвейерных систем, содержащих буферы ограниченных размеров.** В соответствии с декомпозицией обобщенной задачи оптимизации расписаний выполнения пакетов заданий в КС определены иерархически упорядоченные подзадачи. На первом уровне – оптимизация решений по группированию однотипных заданий в пакеты, на втором уровне – оптимизация последовательностей выполнения пакетов заданий на приборах КС, в состав которых включены буферы ограниченных размеров. Одним из возможных подходов к решению рассматриваемой задачи является теория иерархических игр. Ее использование предусматривает закрепление за каждым из уровней игроков, осуществляющих определение оптимальных решений на них: первый уровень – ведущий игрок, второй уровень – ведомый игрок. Применение рассматриваемого аппарата предусматривает взаимодействие игроков, связанное с обменом решениями между ними (между уровнями игры, за которыми они закреплены), которое будет обеспечивать достижение цели ведущим игроком. Действия ведущего игрока направлены на достижение им цели, связанной с минимизацией времени выполнения пакетов заданий в системе рассматриваемого вида. Ведомый игрок оптимизирует расписания выполнения ПЗ в КС, имеющие вид  $[P, R, \{T^{0l} | l = \overline{1, L}\}]$ , соответствующие решениям  $[m, A]$  по их составам, максимизируя эффективность использования ресурсов (минимизируя время простоя приборов при проведении операций с заданиями из пакетов, минимизируя время простоя буферов в ожидании добавления в них заданий из пакетов).

Порядок взаимодействия ведущего и ведомого игроков в иерархической игре следующий:

1) первый ход делает ведущий игрок, формируя решение  $[m, A] \in N_1$  по составам ПЗ, и передает это решение на второй уровень, где  $N_1$  – множество допустимых решений на первом уровне;

2) второй ход делает ведомый игрок, оптимизируя для полученного с первого уровня решения  $[m, A] \in N_1$  по составам ПЗ расписания их выполнения на приборах КС; таким образом, ведомый игрок на втором уровне реализует решение задачи вида:

$$[\mathbf{P}, \mathbf{R}, \{\mathbf{T}^{0l} | l = \overline{1, L}\}]^* = \arg \min_{[\mathbf{P}, \mathbf{R}, \{\mathbf{T}^{0l} | l = \overline{1, L}\}] \in N_2([\mathbf{m}, \mathbf{A}])} f_2([\mathbf{m}, \mathbf{A}], [\mathbf{P}, \mathbf{R}, \{\mathbf{T}^{0l} | l = \overline{1, L}\}]),$$

где  $N_2([\mathbf{m}, \mathbf{A}])$  – множество допустимых расписаний выполнения ПЗ в КС, соответствующее решению  $[\mathbf{m}, \mathbf{A}] \in N_1$  с верхнего уровня;

3) полученное на втором уровне оптимальное расписание выполнения ПЗ в КС вида  $[\mathbf{P}, \mathbf{R}, \{\mathbf{T}^{0l} | l = \overline{1, L}\}]^*$ , соответствующее решению  $[\mathbf{m}, \mathbf{A}] \in N_1$ , ведомый игрок передает на верхний уровень с целью оценки оптимальности сформированного ведущим игроком решения по составам ПЗ (то есть, оценка оптимальности решения по составам ПЗ осуществляется на первом уровне на основе оптимального расписания выполнения ПЗ в КС  $[\mathbf{P}, \mathbf{R}, \{\mathbf{T}^{0l} | l = \overline{1, L}\}]^*$ , полученного для решения  $[\mathbf{m}, \mathbf{A}] \in N_1$ );

4) ведущий игрок формирует новое решение  $[\mathbf{m}, \mathbf{A}] \in N_1$  по составам ПЗ, передает его игроку второго уровня. Определение лучшего состава ПЗ  $[\mathbf{m}, \mathbf{A}]^* \in N_1$  реализуется путем решения задачи определения  $\max_{[\mathbf{m}, \mathbf{A}] \in N_1} f_1([\mathbf{m}, \mathbf{A}], [\mathbf{P}, \mathbf{R}, \{\mathbf{T}^{0l} | l = \overline{1, L}\}]^*)$ , то есть задачи определения  $[\mathbf{m}, \mathbf{A}]^* = \arg \max_{[\mathbf{m}, \mathbf{A}] \in N_1} f_1([\mathbf{m}, \mathbf{A}], [\mathbf{P}, \mathbf{R}, \{\mathbf{T}^{0l} | l = \overline{1, L}\}]^*)$ .

Тогда решением иерархической игры являются: 1) решение по составам ПЗ  $[\mathbf{m}, \mathbf{A}]^* \in N_1$ , 2) решение  $[\mathbf{P}, \mathbf{R}, \{\mathbf{T}^{0l} | l = \overline{1, L}\}]^*$ , обеспечивающее выполнение следующего условия:  $[\mathbf{m}, \mathbf{A}]^* = \arg \max_{[\mathbf{m}, \mathbf{A}] \in N_1} f_1([\mathbf{m}, \mathbf{A}], [\mathbf{P}, \mathbf{R}, \{\mathbf{T}^{0l} | l = \overline{1, L}\}]^*)$ . Эти решения обеспечивают ситуацию равновесия в рассматриваемой двухуровневой иерархической игре двух лиц (являются решением игры).

Для введенной в рассмотрение задачи требуется разработка математических моделей иерархических игр определения оптимальных решений по включению заданий в пакеты и расписаний выполнения пакетов, в соответствии с которыми реализуются операции с ними на приборах КС. Полученные модели обеспечат исследование возможности использования буферов ограниченных размеров для уменьшения влияния неоднородностей значений временных параметров процессов осуществления операций с заданиями в КС на эффективность этих процессов.

Действия ведущего игрока направлены на достижение им цели, связанной с минимизацией времени реализации процесса выполнения

ПЗ в системах рассматриваемого вида. Через  $t_{N_p, N_{N_p}}^{0L}$  обозначен момент времени начала выполнения последнего задания в последнем  $N_p$ -м пакете в последовательности  $\pi^L$  на последнем  $L$ -м приборе, выражение вида  $\sum_{h=1}^N t_{Lh} \cdot p_{h, N_p}$  определяет время выполнения задания, включенного в этот пакет на  $L$ -м приборе. Тогда выражение  $t_{N_p, N_{N_p}}^{0L} + \sum_{h=1}^N t_{Lh} \cdot p_{h, N_p}$  позволяет определить момент времени окончания выполнения последнего задания в последнем  $N_p$ -м пакете в последовательности  $\pi^L$  на последнем  $L$ -м приборе, то есть момент времени окончания реализации указанного процесса в КС. Тогда с точки зрения реализуемой цели ведущего игрока, направленной на минимизацию длительности процесса выполнения ПЗ в КС, требуется минимизация момента времени окончания выполнения последнего задания в последнем  $N_p$ -м пакете в последовательности  $\pi^L$  на последнем  $L$ -м приборе, который определяется с использованием предложенного выражения. В этом случае представленное выражение требуется использовать в качестве критерия оптимизации ведущим игроком решений по составам ПЗ на верхнем уровне игры.

Критерии оптимизации последовательностей  $\pi^l$  ( $l = \overline{1, L}$ ) выполнения пакетов на приборах КС (на втором уровне в иерархических играх) формируются таким образом, чтобы обеспечивалось эффективное использование выделяемых для этого ресурсов. К таким ресурсам относятся: 1) время, в течение которого приборы КС задействованы в проведении действий с заданиями, включенными в пакеты (требуется минимизировать простой приборов при осуществлении операций с заданиями из пакетов); 2) время, выделенное буферам для хранения заданий, с которыми завершены действия на предшествующих приборах (требуется минимизировать общее время ожидания буферами размещения в них заданий из пакетов, действия с которыми завершены на предшествующих приборах). В соответствии с указанными видами ресурсов сформированы две математические модели иерархических игр определения наилучших решений рассматриваемого вида.

В иерархической игре первого вида критерием оптимизации последовательностей  $\pi^l$  ( $l = \overline{1, L}$ ) выполнения ПЗ на приборах КС



на нижнем ее уровне является суммарное время ожидания приборами готовности заданий, входящих в пакеты, к выполнению (вид критерия получен в [1]). То есть на нижнем уровне иерархической игры первого вида оптимизация последовательностей выполнения ПЗ в КС осуществляется с учетом требования (условия) эффективного использования одного из ресурсов этих систем – времени функционирования приборов при реализации заданий, включенных в пакеты. Полученная на основе выполненных рассуждений модель иерархической игры первого вида имеет следующую форму [1]:

– первый уровень игры (ведущий игрок):

$$\min_{[\mathbf{M}, \mathbf{A}] \in N_1} f_1([\mathbf{M}, \mathbf{A}], [\mathbf{P}, \mathbf{R}, \{\mathbf{T}^{0l} \mid l = \overline{1, L}\}]^*), \quad (17)$$

где  $f_1 = t_{N_p, N_p}^{0L} + \sum_{h=1}^N t_{Lh} \cdot p_{h, N_p}$  ;

– второй уровень игры (ведомый игрок):

$$\min_{[\mathbf{P}, \mathbf{R}, \{\mathbf{T}^{0l} \mid l = \overline{1, L}\}] \in N_1([\mathbf{M}, \mathbf{A}])} f_2([\mathbf{M}, \mathbf{A}], [\mathbf{P}, \mathbf{R}, \{\mathbf{T}^{0l} \mid l = \overline{1, L}\}]), \quad (18)$$

где

$$f_2 = \sum_{l=1}^L t_{11}^{0l} + \sum_{l=1}^L \sum_{j=2}^{N_p} \left[ t_{j1}^{0l} - \left[ t_{j-1, N_{j-1}}^{0l} + \sum_{i=1}^N t_{li} \cdot p_{i, j-1} \right] \right] + \sum_{l=2}^L \sum_{j=l}^{N_p} \sum_{q=2}^{N_j} \left[ t_{jq}^{0l} - \left[ t_{j, q-1}^{0l} + \sum_{i=1}^N t_{li} \cdot p_{ij} \right] \right]. \quad (19)$$

В (19) первое слагаемое соответствует простоям приборов на начальной стадии заполнения конвейерной системы заданиями из пакетов для проведения операций с ними. Второе слагаемое в (19) – простои приборов на стадии их переналадки, связанной с переопределением вида операций для заданий в следующих пакетах в  $\pi^l (l = \overline{1, L})$ . Третье слагаемое в (19) – простои, связанные с ожиданием готовности заданий в пакетах к осуществлению операций с ними на приборах.

Ограничения на множества допустимых решений имеют вид:

$$\sum_{h=1}^{m_i} a_{ih} = n^i; \quad i = \overline{1, N}; \quad (20)$$

$$\sum_{j=1}^{N_p} p_{ij} = m_i; \quad i = \overline{1, N}; \quad (21)$$

$$\sum_{j=1}^{N_p} r_{ij} = \sum_{h=1}^{m_i} a_{ih}; \quad i = \overline{1, N}. \quad (22)$$

Они соответствуют: требованию включения в пакеты всех заданий каждого типа – (20), требованию размещения в последовательностях  $\pi^l$  всех пакетов каждого  $i$ -го типа – (21), требованию размещения в последовательностях  $\pi^l$  всех пакетов, составы которых определены на первом уровне игры – (22). С учетом рассмотренной выше процедуры взаимодействия игроков представленная двухкритериальная модель является моделью иерархической игры оптимизации составов пакетов и расписаний их выполнения на приборах КС (при учете требования минимизации простоев приборов при реализации ими операций с заданиями из пакетов). Оптимизация составов пакетов (на верхнем уровне) при условии минимизации общей длительности процесса их выполнения обеспечивается определением оптимальных решений по их порядкам для реализации операций с ними на приборах КС, формируемых ведомым игроком на нижнем уровне для каждого решения, получаемого им от ведущего игрока с верхнего уровня. Оценка оптимальности решений, формируемых ведущим игроком, осуществляется на основе оптимальных расписаний их выполнения в КС, полученных от ведомого игрока и сформированных им для соответствующих решений по составам ПЗ.

Второй вариант математической модели иерархической игры предусматривает, что на втором уровне определены оптимальных расписаний выполнения ПЗ в КС осуществляется с учетом требования эффективного использования ресурса времени промежуточных буферов. Характеристиками процесса реализации операций с заданиями из пакетов в КС являются моменты времени добавления их (заданий) в буферы и удаления их из буфера ( $t_{jq}^{6l1}$  и  $t_{jq}^{6l2}$ ). Значение

$t_{jq}^{\bar{0}l1}$  определяется как момент времени окончания операций с  $q$ -м заданием из пакета, который занимает в последовательностях  $\pi^l$   $j$ -ю позицию, на  $l$ -м приборе:  $t_{jq}^{\bar{0}l1} = t_{jq}^{\bar{0}l} + \sum_{i=1}^N t_{li} \cdot p_{ij}$  (при соблюдении требования блокирования  $l$ -го прибора в случае отсутствия места в следующем за ним буфере). Значение  $t_{jq}^{\bar{0}l2}$  определяется как момент времени начала операций с  $q$ -м заданием из пакета, который занимает в последовательностях  $\pi^l$   $j$ -ю позицию, на  $(l+1)$ -м приборе:  $t_{jq}^{\bar{0}l2} = t_{jq}^{\bar{0}l+1}$ . Критерий оптимизации расписаний выполнения ПЗ в КС в модели игры второго вида характеризует суммарные длительности интервалов времени, в течение которых буферы, следующие за  $l$ -ми приборами ( $l = \overline{1, L-1}$ ), не используются.

Введены обозначения для характеристик процесса выполнения ПЗ в КС, соответствующих использованию ограниченного ресурса хранения заданий: 1) длительности интервалов времени ожидания буферами добавления в них  $q$ -х заданий из пакетов, которые размещены в последовательностях  $\pi^l$  в  $j$ -х позициях –  $W_{jq}^{\bar{0}l}$ ; 2) длительности интервалов времени ожидания буферами добавления в них первых заданий из пакетов, которые включены в последовательностях  $\pi^l$  в  $j$ -х позициях –  $W_{j1}^{\bar{0}l}$ ; 3) длительности интервалов времени ожидания буферами добавления в них первых заданий из первых пакетов в  $\pi^l$  –  $W_{11}^{\bar{0}l}$ . Виды последовательностей, представленные на рисунке 1, использованы при синтезе выражений для определения значений этих характеристик. Вычисление значений характеристик, позволяющих определить эффективность использования ресурсов хранения, осуществляется выражениями вида:

$$W_{jq}^{\bar{0}l} = \max(0; t_{jq}^{\bar{0}l1} - t_{jq-1}^{\bar{0}l2}), \quad q = \overline{2, N_j}, \quad j = \overline{1, N_p}, \quad l = \overline{1, L-1};$$

$$W_{j1}^{\bar{0}l} = \max(0; t_{j1}^{\bar{0}l1} - t_{j-1, N_{j-1}}^{\bar{0}l2}), \quad j = \overline{2, N_p}, \quad l = \overline{1, L-1};$$

$$W_{11}^{\bar{0}l} = t_{11}^{\bar{0}l1} = t_{11}^{\bar{0}l} + \sum_{i=1}^N t_{li} \cdot p_{i1}, \quad l = \overline{1, L-1}.$$

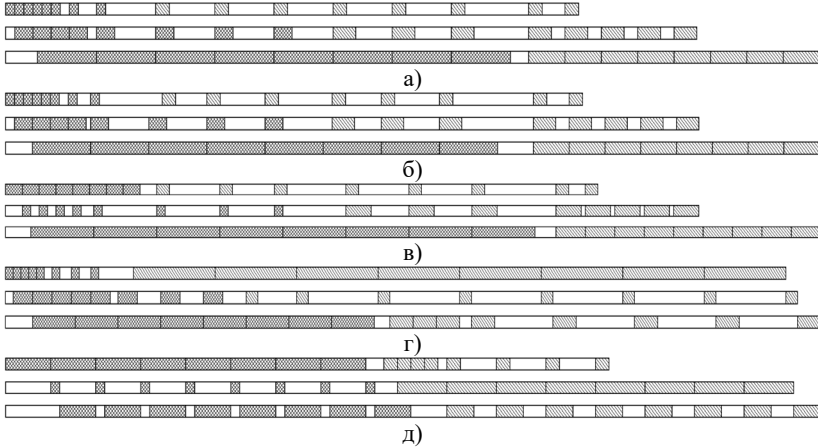


Рис. 1. Последовательности выполнения ПЗ в КС, различающиеся значениями длительностей  $t_{li}$  выполнения заданий разных типов, используемые при

формировании выражений для вычисления значений  $W_{jq}^{\delta l}$ ,  $W_{j1}^{\delta l}$  и  $W_{11}^{\delta l}$

- а)  $t_{1i1} \leq t_{2i1}, t_{2i1} \leq t_{3i1}, t_{1i2} \leq t_{2i2}, t_{2i2} \leq t_{3i2}$ ; б)  $t_{1i1} \leq t_{2i1}, t_{2i1} \leq t_{3i1}, t_{1i2} \leq t_{2i2}, t_{2i2} \leq t_{3i2}$ ; в)  $t_{1i1} > t_{2i1}, t_{2i1} \leq t_{3i1}, t_{1i2} \leq t_{2i2}, t_{2i2} \leq t_{3i2}$ ; г)  $t_{1i1} \leq t_{2i1}, t_{2i1} \leq t_{3i1}, t_{1i2} > t_{2i2}, t_{2i2} \leq t_{3i2}$ ; д)  $t_{1i1} > t_{2i1}, t_{2i1} \leq t_{3i1}, t_{1i2} > t_{2i2}, t_{2i2} \leq t_{3i2}$

Представленные выражения преобразованы с учетом способов вычисления значений  $t_{jq}^{\delta l1}$  и  $t_{jq}^{\delta l2}$ . Итоговый вид выражений для вычисления длительностей интервалов времени ожидания буферами размещения в них одного задания из одного пакета в последовательностях  $\pi^l$  следующий (при условии, что  $q \neq 1$  и  $j \neq 1$ ):

$$W_{jq}^{\delta l} = \max(0; (t_{jq}^{ol} + \sum_{i=1}^N t_{li} \cdot p_{ij}) - t_{jq-1}^{ol+1}), q = \overline{2, N_j}, j = \overline{1, N_p}, l = \overline{1, L-1};$$

$$W_{j1}^{\delta l} = \max(0; (t_{j1}^{ol} + \sum_{i=1}^N t_{li} \cdot p_{ij}) - t_{j-1, N_{j-1}}^{ol+1}), j = \overline{2, N_p}, l = \overline{1, L-1}.$$

Тогда общая длительность интервалов времени ожидания всеми буферами добавления в них заданий из пакетов (при  $q = \overline{2, N_j}$ ,

$j = \overline{1, N_p}, l = \overline{1, L-1}$ ) определяется выражением  $\sum_{l=1}^{L-1} \sum_{j=1q=2}^{N_p} W_{jq}^{\delta l}$ , общая

длительность интервалов времени ожидания всеми буферами добавления в них первых заданий из пакетов ( $j = \overline{2, N_p}, l = \overline{1, L-1}$ )

определяется выражением  $\sum_{l=1}^{L-1} \sum_{j=2}^{N_p} W_{j1}^{\delta l}$ , общая длительность

интервалов времени ожидания всеми буферами добавления в них первых заданий из первых пакетов (при  $l = \overline{1, L-1}$ ) определяется

выражением  $\sum_{l=1}^{L-1} \sum_{j=2}^{N_p} W_{j1}^{\delta l}$ .

Итоговый вид математической модели иерархической игры оптимизации решений по группированию заданий в пакеты и по последовательностям пакетов для их выполнения на приборах КС следующий:

$$\min_{[\mathbf{M}, \mathbf{A}] \in N_1} f_1([\mathbf{M}, \mathbf{A}], [\mathbf{P}, \mathbf{R}, \{\mathbf{T}^{0l} | l = \overline{1, L}\}]^*), \quad (23)$$

где  $f_1 = t_{N_p, N_p}^{0L} + \sum_{i=1}^N t_{Li} \cdot p_{i, N_p}$ ;

$$\min_{[\mathbf{P}, \mathbf{R}, \{\mathbf{T}^{0l} | l = \overline{1, L}\}] \in N_1([\mathbf{M}, \mathbf{A}])} f_2([\mathbf{M}, \mathbf{A}], [\mathbf{P}, \mathbf{R}, \{\mathbf{T}^{0l} | l = \overline{1, L}\}]), \quad (24)$$

где

$$f_2 = \sum_{l=1}^{L-1} W_{11}^{\delta l} + \sum_{l=1}^{L-1} \sum_{j=2}^{N_p} W_{j1}^{\delta l} + \sum_{l=1}^{L-1} \sum_{j=1q=2}^{N_p} W_{jq}^{\delta l},$$

$$W_{jq}^{\delta l} = \max(0; (t_{jq}^{0l} + \sum_{i=1}^N t_{li} \cdot p_{ij}) - t_{jq-1}^{0l+1}), \quad q = \overline{2, N_j}, j = \overline{1, N_p}, l = \overline{1, L-1};$$

$$W_{j1}^{\delta l} = \max(0; (t_{j1}^{0l} + \sum_{i=1}^N t_{li} \cdot p_{ij}) - t_{j-1, N_{j-1}}^{0l+1}), \quad j = \overline{2, N_p}, l = \overline{1, L-1};$$

$$W_{11}^{\bar{0}l} = t_{11}^{0l} + \sum_{i=1}^N t_{li} \cdot p_{i1}, \quad l = \overline{1, L-1}.$$

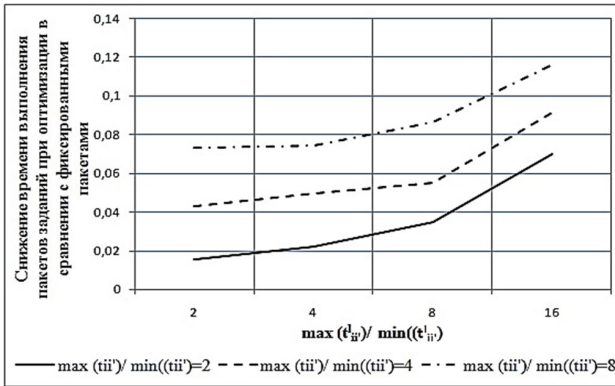
Ограничения для полученной модели иерархической игры аналогичны ограничениям (20)-(22) для модели игры первого вида. Ограничение на размеры буферов в обеих математических моделях иерархических игр отсутствует, так как условие блокирования предшествующего прибора в ожидании освобождения одной позиции для выполненного задания в буфере учтено в модели (1)-(16).

Определение решений по группированию заданий в пакеты и по последовательностям пакетов для осуществления операций с ними на приборах реализуется с использованием метода поиска лучших решений указанного вида, предложенного в [1], определение наилучших решений по последовательностям пакетов осуществления операций на приборах КС выполняется с использованием алгоритма, также предложенного в [1]. При реализации выбора решений указанного видам расчет временных характеристик процессов осуществления операций с пакетами на приборах КС проводится не с использованием модели, предложенной в [1], а с использованием модели (1)-(16), в которой учтено наличие промежуточных буферов ограниченного размера между приборами.

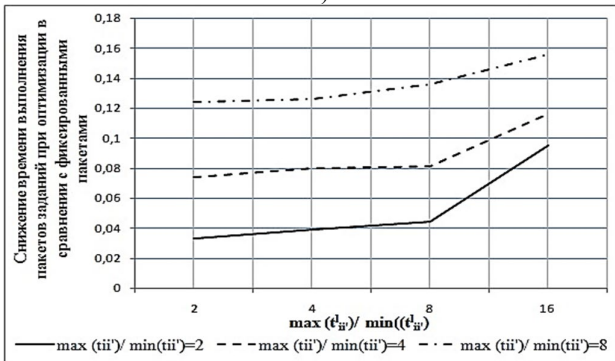
**4. Исследование возможностей использования буферов для повышения эффективности процессов выполнения пакетов заданий в конвейерных системах.** Проводимые исследования направлены на то, чтобы удостовериться в том, что: 1) применение моделей иерархических игр обеспечивает определение оптимальных решений по группированию заданий в пакеты и по последовательностям пакетов для реализации операций с ними на приборах КС при учете эффективного использования различных ресурсов приборов (в том числе с ограничениями на размеры буферов); 2) применение буферов позволяет снизить влияние неоднородностей временных параметров процессов реализации действий с заданиями на приборах КС на общее время процесса их выполнения в КС. Указанные исследования проведены путем решения задач оптимизации решений рассмотренных видов при варьировании значений их (задач) параметров. Параметры задач заданы таким образом, что  $\max(t_{ii})/\min(t_{ii}) \in \{2,4,8\}$  (неоднородность времени, в течение которого осуществляются операции с заданиями разных типов) и  $\max(t_{i'v}^l)/\min(t_{i'v}^l) \in \{2,4,8,16\}$  (неоднородность интервалов времени, в течение которых осуществляются переналадки приборов на выполнение заданий разных типов),  $b \in \{1,2,4,8\}$ . Оставшиеся параметры задач:  $n=5$ ,

$n^i=12, L=5$ . Результатом экспериментов с моделями являются установленные зависимости отношения  $f_{y\partial n} = (f_1^{\text{фикс}} - f_1^{\text{взн}}) / f_1^{\text{фикс}}$  от параметров задач ( $f_{y\partial n}$  – доля уменьшения длительности процесса выполнения заданий из пакетов (УДП) для найденных оптимальных решений по включению в них единичных заданий и по их (пакетов) последовательностям в сравнении с фиксированными пакетами, в которые включены все задания каждого типа,  $f_1^{\text{фикс}}$  – значение критерия на первом уровне для пакетов, в которые включены все задания каждого типа;  $f_1^{\text{взн}}$  – значение критерия на первом уровне для лучших решений по группированию заданий в пакеты).

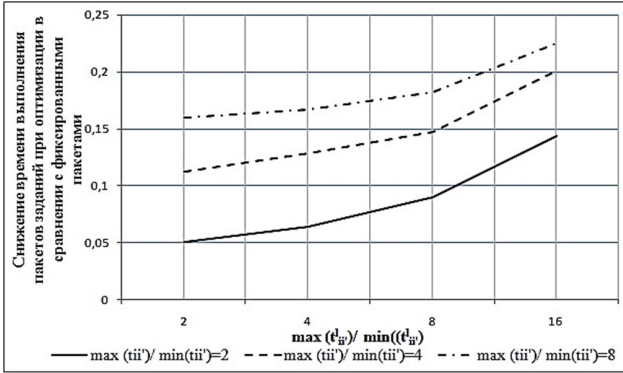
Результаты исследований моделей иерархических игр в виде зависимости  $f_{y\partial n}$  от параметров задач, представлены на рисунках 2 и 3.



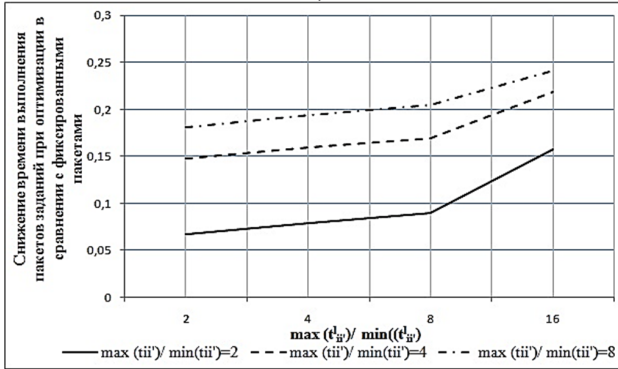
а)



б)

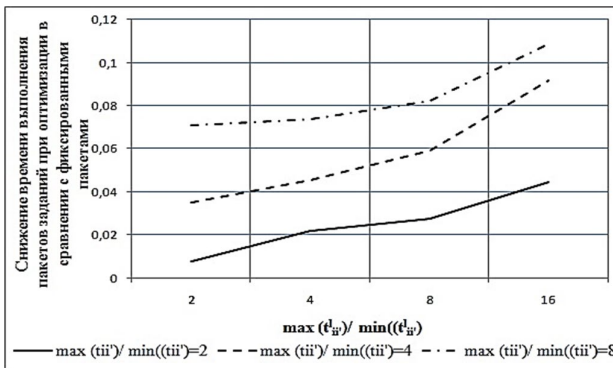


Б)



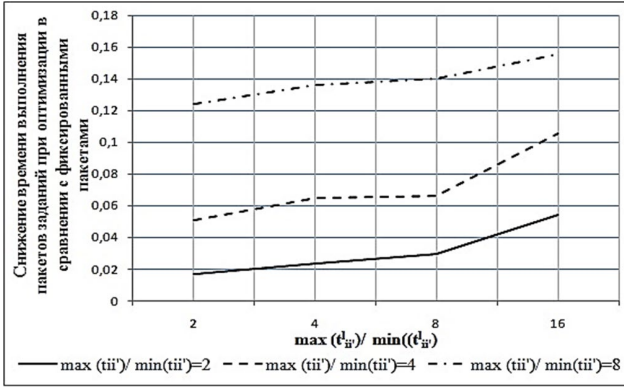
Г)

Рис. 2. Сокращение времени реализации процессов выполнения ПЗ в КС для модели иерархической игры первого вида: а)  $b=1$ ; б)  $b=2$ ; в)  $b=4$ ; г)  $b=8$

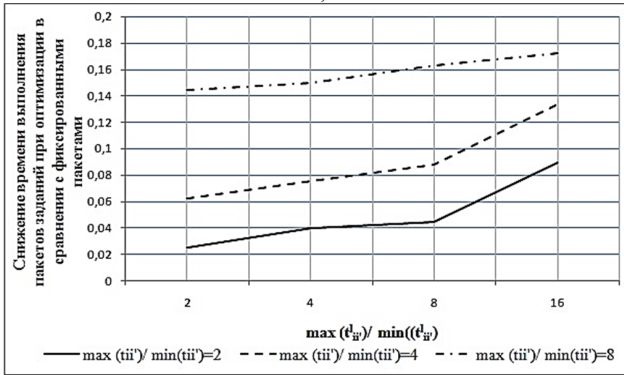


а)

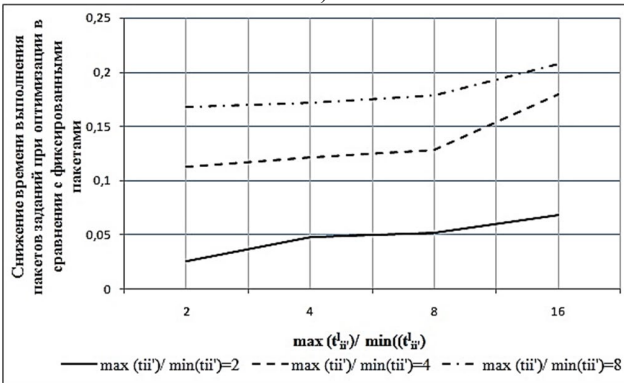




б)



в)



г)

Рис. 3. Сокращение времени реализации процессов выполнения ПЗ в КС для модели иерархической игры второго вида: а)  $b=1$ ; б)  $b=2$ ; в)  $b=4$ ; г)  $b=8$

В соответствии с полученными зависимостями значений  $f_{yop}$  от параметров задач сделаны выводы, касающиеся особенностей операций с пакетами на приборах рассматриваемых систем:

1) увеличение размера промежуточных буферов (с  $b=1$  до  $b=8$ ) позволяет значительно увеличить эффективность рассматриваемых процессов; при  $\max(t_{ii})/\min(t_{ii}) \in \{2,4\}$  увеличение количества позиций в буфере (в указанных пределах) обеспечивает рост эффективности процессов от 0,02 до 0,15; при  $\max(t_{ii})/\min(t_{ii})=8$  увеличение количества позиций в буфере обеспечивает рост эффективности процессов выполнения заданий из пакетов от 0,08 до 0,23; таким образом, при больших значениях неоднородности времени выполнения заданий разных типов на приборах (заданной отношением  $\max(t_{ii})/\min(t_{ii})$ ) использование промежуточных буферов позволяет в большей степени повысить эффективность рассматриваемых процессов;

2) использование буфера более значительно влияет на эффективность процесса выполнения заданий, входящих в пакеты, при значительной неоднородности интервалов времени, в течение которых проводятся на приборах операции с заданиями разных типов (заданной отношением  $\max(t_{ii})/\min(t_{ii})$ );

3) при увеличении размера буфера рост эффективности процессов выполнения пакетов заданий является не значительным при  $\max(t'_{i'v})/\min(t'_{i'v}) \in \{2,4\}$ ; при  $\max(t'_{i'v})/\min(t'_{i'v}) \in \{8,16\}$  увеличение размера буфера обеспечивает более значительный рост эффективности рассматриваемых процессов;

4) при отношениях  $\max(t_{ii})/\min(t_{ii})$  и  $\max(t'_{i'v})/\min(t'_{i'v})$ , характеризующих малыми значениями, использование модели иерархической игры первого вида (группирование заданий по пакетам и определение последовательностей выполнения пакетов с учетом простоев приборов) обеспечивает более значительный рост эффективности рассматриваемых процессов по сравнению со второй моделью; при больших значениях отношений  $\max(t_{ii})/\min(t_{ii})$  и  $\max(t'_{i'v})/\min(t'_{i'v})$  полученные с использованием первой и второй моделей результаты являются сопоставимыми с точки зрения значений эффективности процессов выполнения ПЗ в КС.

Таким образом, включение в КС промежуточных буферов позволяет уменьшить влияние неоднородностей времени, в течение которого проводятся операции с заданиями разных типов на приборах КС, и неоднородностей времени, в течение которого реализуются

переналадки приборов, на общее время протекания рассматриваемых процессов. Увеличение размеров буферов позволяет снизить это влияние в большей степени.

**5. Заключение.** В статье предложена математическая модель многостадийных процессов выполнения пакетов заданий в конвейерных системах, в состав которых включены промежуточные буферы ограниченных размеров. Использование этой модели допустимо при произвольном количестве приборов в КС и произвольном количестве пакетов заданий в последовательностях их выполнения на этих приборах. Особенностью предложенной модели указанных процессов является учет перемещения заданий, включенных в пакеты, между приборами (либо между приборами и следующими за ними буферами) непосредственно в момент времени окончания реализации действий с ними (а не в момент времени окончания выполнения всех заданий, входящих в пакет). Также в полученной модели модифицированы выражения для определения моментов времени начала выполнения заданий из пакетов на  $l$ -х приборах КС. Для этого в указанные выражения включены моменты времени начала выполнения заданий из пакетов на  $(l+1)$ -х приборах, следующих за рассматриваемыми. Это позволяет непосредственно в самой модели многостадийных процессов выполнения пакетов заданий в конвейерных системах учесть условие блокирования предшествующих приборов в случае отсутствия места в буфере для размещения заданий.

С целью оптимизации решений по составам пакетов заданий и расписаний их выполнения на приборах КС проведена разработка двух математических моделей иерархических игр оптимизации решений указанного вида (применение аппарата иерархических игр предусматривает установленный порядок ходов и обмен решениями между игроками на уровнях иерархии с целью оптимизации соответствующих решений). В обеих моделях иерархических игр критерий оптимизации состав пакетов на первом уровне соответствуют общей длительности процессов выполнения пакетов в указанных системах, а критерии идентификации лучших решений на втором уровне соответствуют эффективности использования ресурсов обрабатывающих приборов этих систем. В частности, получены выражения, позволяющие определять общее время ожидания буферами размещения в них заданий, с которыми завершились операции на предшествующих приборах, а также критерий оптимизации расписаний выполнения ПЗ в КС, учитывающий эффективность использования этого вида ресурса.

Результаты решения тестовых задач оптимизации решений рассматриваемого вида показали, что: 1) использование буферов позволяет значительно повысить эффективность процессов осуществления операций с пакетами заданий на приборах указанных систем; 2) увеличение размеров промежуточных буферов позволяет в большей степени повысить эффективность рассматриваемых процессов при больших значениях неоднородностей временных параметров выполнения заданий разных типов на приборах КС, содержащих промежуточные буферы ограниченных размеров.

### Литература

1. Кротов К.В. Комплексный метод определения эффективных решений по составам партий данных и расписаниям их обработки в конвейерных системах // Вычислительные технологии. Изд-во Института вычислительных технологий СО РАН. 2018. Т. 23. № 3. С. 58–76.
2. Кротов К.В., Скاتков А.В. Построение комплексных расписаний выполнения пакетов заданий при формировании комплектов в заданные директивные сроки // Информатика и автоматизация. 2021. Т. 20(3). С. 654–689. DOI: 10.15622/ia.2021.3.6.
3. Кротов К.В., Скатков А.В. Оптимизация планирования выполнения пакетов заданий в многостадийных системах при ограничениях и формировании комплектов // Компьютерные исследования и моделирование. 2021. Т. 13. № 5. С. 917–946.
4. Кротов К.В. Математическое моделирование процессов выполнения пакетов заданий в конвейерных системах с промежуточными буферами ограниченных размеров // Информатика и автоматизация. 2023. Т. 22(6). С. 1415–1450.
5. Paradimitriou Ch.H., Kanellakis P.C. Flowshop scheduling with limited temporary storage // Journal of Association for Computing Machinery. 1980. vol. 27. no. 3. pp. 533–549.
6. Leisten R. Flowshop sequencing problems with limited buffer storage // International Journal of Production Research. 1990. vol. 28. no. 11. pp. 2085–2100.
7. Crowder B. Minimizing the makespan in a flexible flowshop with sequence dependent setup times, uniform machines and limited buffers // Graduate Theses, Dissertations and Problem Reports. Morgantown: West Virginia University, 2006. 145 p. DOI: 10.33915/etd.4220.
8. Han Zh., Zhang Q., Shi H., Qi Yu., Sun L. Research on limited buffer scheduling problems in flexible flow shops with setup times // International Journal of Modelling, Identification and Control. 2019. vol. 32. no. 2. pp. 93–104.
9. Eddaly M., Jarboui B., Siarry P., Rebai A. An Estimation of Distribution Algorithm for Flowshop Scheduling with Limited Buffers // Natural Intelligence for Scheduling, Planning and Packing Problems. Studies in Computational Intelligence. 2009. (SCI). vol. 250. pp. 89–110.
10. Frasc J.V., Krumke S.O., Westphal S. MIP Formulations for Flowshop Scheduling with Limited Buffers // Proceedings First International ICST Conference «Theory and Practice of Algorithms in (Computer) Systems» (TAPAS). 2011. vol. 6595. pp. 127–138.
11. Fu Q., Sivakumar A., Li K. Optimisation of flow-shop scheduling with batch processor and limited buffer // International Journal of Production Research. 2012. vol. 50. pp. 2267–2285.

12. Çakıcı M.K. Parallel flow shop scheduling with common workstations // A thesis submitted to the graduate school of natural and applied sciences of Middle East technical university. 2019. 138 p.
13. Кононова П.А., Кочетов Ю.А. Алгоритм локального поиска для построения расписаний работы одного станка с переналадкой оборудования и складом // Дискретный анализ и исследование операций. 2019. Т. 26. № 2. С. 60–78.
14. Lin Ch.-Ch., Liu W.-Y., Chen Y.-H. Considering Stockers in Reentrant Hybrid Flow Shop Scheduling with Limited Buffer Capacity // Computers & Industrial Engineering. 2020. vol. 139. DOI: 10.1016/j.cie.2019.106154.
15. Takano M.I., Nagano M.S. Solving the permutation flow shop problem with blocking and setup time constraints // International Journal of Industrial Engineering Computations. 2020. vol. 11. no. 3. pp. 469–480. DOI: 10.52677/ij.ijiec.2019.11.002.
16. Zhang Ch., Tan J., Peng K., Gao L., Shen W., Lian K. A discrete whale swarm algorithm for hybrid flow-shop scheduling problem with limited buffers // Robotics and Computer-Integrated Manufacturing. 2021. vol. 68. DOI: 10.1016/j.rcim.2020.102081.
17. Rooeinfar R., Raissi S., Ghezvati V.R. Stochastic flexible flow shop scheduling problem with limited buffers and fixed interval preventive maintenance: a hybrid approach of simulation and metaheuristic algorithms // Simulation. 2019. vol. 95(6). pp. 509–528. DOI: 10.1177/0037549718809542.
18. Ernst A., Fung J., Singh G., Zinder Ya. Flexible flow shop with dedicated buffers // Discrete Applied Mathematics. 2019. no. 261. pp. 148–163. DOI: 10.1016/j.dam.2018.07.002.
19. Berlinska J. Heuristics for scheduling data gathering with limited base station memory // Annals of Operations Research. 2020. no. 285(1). pp. 149–159. DOI: 10.1007/s10479-019-03185-3.
20. Berlinska J., Kononov A., Zinder Ya. Two-machine flow shop with dynamic storage space // Optimization Letters. 2021. vol. 15. pp. 2433–2454. DOI: 10.1007/s11590-020-01645-5.
21. Souiden S., Cerqueus A., Delorme X., Rascle J.-L. Retail order picking scheduling with missing operations and limited buffer // IFAC-Papers On Line, 21st IFAC World Congress. 2020. voo. 53(2). pp. 10767–10772. DOI: 10.1016/j.ifacol.2020.12.2859.
22. Esfeh M.K., Shojaei A.A., Javanshir H., Damghani K.K. Solving a bi-objective flexible flow shop problem with transporter preventive maintenance planning and limited buffers by NSGA-II and MOPSO // The International Journal of Nonlinear Analysis and Applications (IJNAA). 2022. vol. 13. no. 1. pp. 217–246. DOI: 10.22075/ijnaa.2021.24335.2719.
23. Koh J., Bodin B. K-periodic scheduling for throughput-buffering trade-off exploration of CSDF // ACM Transactions on Embedded Computing Systems. 2022. vol. 22. no. 1. DOI: 10.1145/3559760.
24. Honorat A., Dardaillon M., Miomandre H., Nezan J.-F. Automated buffer sizing of dataflow applications in a high-level synthesis workflow // ACM Transactions on Reconfigurable Technology and Systems (TRETS). 2024. no. 17(1). pp. 1–26. DOI: 10.1145/3626103.
25. Van Schilt I.M., Van Kalker J., Lefter I., Kwakkel J., Verbraeck A. Buffer scheduling for improving on-time performance and connectivity with a multi-objective simulation-optimization model: A proof of concept for the airline industry // Journal of Air Transport Management. 2024. no. 115(7). DOI: 10.1016/j.jairtraman.2024.102547.
26. Agnetis A., Pacciarelli D., Rossi F. Batch scheduling in a two-machine flow shop with limited buffer // Discrete Applied Mathematics. 1997. no. 72. pp. 243–260.

27. Pranzo M. Batch scheduling in a two-machine flow shop with limited buffer and sequence independent setup times and removal times // *European Journal of Operational Research*. 2004. no. 153(3). pp. 581–592.
28. Dai J. Batch Scheduling of Two-machine Limited-buffer Flow Shop with Setup and Removal Times // A Thesis for the Degree of Doctor of Philosophy in the School of Industrial and Systems Engineering. Georgia Institute of Technology. 2003. 108 p.
29. Belaid R., T'kindt V., Esswein C. Scheduling batches in flow shop with limited buffers in the shampoo industry // *European Journal of Operational Research*. 2012. vol. 223(2). pp. 560–572.
30. Ruhbakhsh R., Mehdizadeh E., Adibi M. A Mathematical Model for Lot-streaming Hybrid Flow Shop Scheduling Problem by Considering Learning Effect and Buffer Capacity // *Scientia Iranica*. 2022. DOI: 10.24200/SCI.2022.58131.5582.
31. Janeš G., Ištoković D., Jurković Z., Perinic M. Application of modified steady-state genetic algorithm for batch sizing and scheduling problem with limited buffers // *Applied Sciences*. 2022. no. 12(22). DOI: 10.3390/app122211512.

**Кротов Кирилл Викторович** — д-р техн. наук, доцент, профессор кафедры, кафедра «информационные системы», ФГАОУ ВО «Севастопольский государственный университет». Область научных интересов: теория расписаний, многоуровневая оптимизация, методы дискретной оптимизации. Число научных публикаций — 122. krotov\_k1@mail.ru; улица Университетская, 33, 299053, Севастополь, Россия; р.т.: +7(978)730-3819.

К. КРОТОВ

**TWO-LEVEL OPTIMIZATION OF TASK DISTRIBUTION INTO BATCHES AND SCHEDULING THEIR EXECUTION IN PIPELINE SYSTEMS WITH LIMITED BUFFERS**

*Krotov K. Two-Level Optimization of Task Distribution into Batches and Scheduling Their Execution in Pipeline Systems with Limited Buffers.*

**Abstract.** Currently, existing mathematical models and algorithms provide optimization of schedules for the execution of single tasks or fixed task packages on devices of conveyor systems containing buffers of limited sizes. These models and algorithms do not allow searching for optimal solutions for grouping the same type of tasks into packages and by sequence of packages to implement operations with them on devices of conveyor systems. Increasing the efficiency of using the resources of conveyor systems is achieved by optimizing solutions for grouping the same type of tasks into packages and by sequences of packages for performing operations with them. The solution to this problem is carried out in the work by using an approach that implements two-level optimization, which allows you to form a hierarchy of subtasks for finding effective solutions. The involvement of the mentioned approach involves the development of mathematical models of hierarchical games that allow identifying effective solutions of the type under consideration. Two mathematical models of hierarchical games have been constructed, the use of which makes it possible to optimize package compositions at the upper level by the leading player and optimize package execution schedules in pipeline systems at the lower level by the slave player. The method of determining the optimal solutions for each of the players provides for the order of moves set in the game and the exchange of solutions between them during the game. The first mathematical model of the hierarchical game implements the definition of effective solutions when taking into account the downtime of processing devices in the process of implementing operations with packages. The second mathematical model of the game implements the definition of effective solutions, taking into account the total waiting time for buffers to place tasks in them, with which operations on previous devices were completed. To do this, expressions have been formed that allow you to determine buffer downtime while waiting for tasks from packages to be ready for placement based on the time characteristics of the processes of performing operations with them on the devices of the systems under consideration. The algorithm for determining optimal solutions according to the order of operations with packages at the lower level in each of the hierarchical games is based on a developed mathematical model of the processes of implementing actions with packages in these systems and the corresponding modeling algorithm. The implementation of the optimization approach under consideration allowed us to obtain results that showed that the use of buffers can significantly increase the efficiency of the processes of performing operations with packets on the devices of the systems under consideration; increasing the size of intermediate buffers allows us to increase the efficiency of these processes to a greater extent with significant heterogeneities in the values of time parameters characterizing them; using the first model of a hierarchical game allows us to achieve a greater increase in the efficiency of processes in comparison with the second model.

**Keywords:** pipeline systems, schedules, buffers of limited size, task packages, two-level optimization.

**References**

1. Krotov K.V. [A complex method for determining effective solutions for data batch compositions and processing schedules in conveyor systems]. Vychislitel'nyye

- tehnologii, Izd-vo Instituta vychislitel'nyh tekhnologij SO RAN – Computational Technologies, Publishing House of the Institute of Computational Technologies SB RAS. 2018. vol. 23. no. 3. pp. 58–76. (In Russ.).
2. Krotov K.V., Skatkov A.V. [Construction of complex schedules for the execution of task packages when forming sets within the specified directive deadlines]. *Informatika i avtomatizacija – Informatics and Automation*. 2021. vol. 20(3). pp. 654–689. DOI: 10.15622/ia.2021.3.6. (In Russ.).
  3. Krotov K.V. [Optimization of task package execution planning in multistage systems under constraints and formation of sets]. *Komp'yuternye issledovaniya i modelirovanie – Computer research and modeling*. 2021. vol. 13. no. 5. pp. 917–946. (In Russ.).
  4. Krotov K.V. [Mathematical modeling of the processes of completing task packages in conveyor systems with intermediate buffers of limited size]. *Informatika i avtomatizacija – Informatics and Automation*. 2023. vol. 22. no. 6. pp. 1415–1450. (In Russ.).
  5. Papadimitriou Ch.H., Kanellakis P.C. Flowshop scheduling with limited temporary storage. *Journal of Association for Computing Machinery*. 1980. vol. 27. no. 3. pp. 533–549.
  6. Leisten R. Flowshop sequencing problems with limited buffer storage. *International Journal of Production Research*. 1990. vol. 28. no. 11. pp. 2085–2100.
  7. Crowder B. Minimizing the makespan in a flexible flowshop with sequence dependent setup times, uniform machines and limited buffers. Graduate Theses, Dissertations and Problem Reports. Morgantown: West Virginia University, 2006. 145 p. DOI: 10.33915/etd.4220.
  8. Han Zh., Zhang Q., Shi H., Qi Yu., Sun L. Research on limited buffer scheduling problems in flexible flow shops with setup times. *International Journal of Modelling, Identification and Control*. 2019. vol. 32. no. 2. pp. 93–104.
  9. Eddaly M., Jarboui B., Siarry P., Rebaï A. An Estimation of Distribution Algorithm for Flowshop Scheduling with Limited Buffers. *Natural Intelligence for Scheduling, Planning and Packing Problems. Studies in Computational Intelligence*. 2009. (SCI). vol. 250. pp. 89–110.
  10. Frasch J.V., Krumke S.O., Westphal S. MIP Formulations for Flowshop Scheduling with Limited Buffers. *Proceedings First International ICST Conference «Theory and Practice of Algorithms in (Computer) Systems» (TAPAS)*. 2011. vol. 6595. pp. 127–138.
  11. Fu Q., Sivakumar A., Li K. Optimisation of flow-shop scheduling with batch processor and limited buffer. *International Journal of Production Research*. 2012. vol. 50. pp. 2267–2285.
  12. Çakici M.K. Parallel flow shop scheduling with common workstations. A thesis submitted to the graduate school of natural and applied sciences of Middle East technical university. 2019. 138 p.
  13. Kononova P.A., Kochetov Yu.A. [Local search algorithm for building schedules for the operation of one machine with equipment changeover and warehouse]. *Diskretnyj analiz i issledovanie operacij – Discrete analysis and operations research*. 2019. vol. 26. no. 2. pp. 60–78. (In Russ.).
  14. Lin Ch.-Ch., Liu W.-Y., Chen Y.-H. Considering Stockers in Reentrant Hybrid Flow Shop Scheduling with Limited Buffer Capacity. *Computers & Industrial Engineering*. 2020. vol. 139. DOI: 10.1016/j.cie.2019.106154.
  15. Takano M.I., Nagano M.S. Solving the permutation flow shop problem with blocking and setup time constraints. *International Journal of Industrial Engineering Computations*. 2020. vol. 11. no. 3. pp. 469–480. DOI: 10.52677/ijieec.2019.11.002.



16. Zhang Ch., Tan J., Peng K., Gao L., Shen W., Lian K. A discrete whale swarm algorithm for hybrid flow-shop scheduling problem with limited buffers. *Robotics and Computer-Integrated Manufacturing*. 2021. vol. 68. DOI: 10.1016/j.rcim.2020.102081.
17. Rooeinfar R., Raissi S., Ghezvati V.R. Stochastic flexible flow shop scheduling problem with limited buffers and fixed interval preventive maintenance: a hybrid approach of simulation and metaheuristic algorithms. *Simulation*. 2019. vol. 95(6). pp. 509–528. DOI: 10.1177/0037549718809542.
18. Ernst A., Fung J., Singh G., Zinder Ya. Flexible flow shop with dedicated buffers. *Discrete Applied Mathematics*. 2019. no. 261. pp. 148–163. DOI: 10.1016/j.dam.2018.07.002.
19. Berlinska J. Heuristics for scheduling data gathering with limited base station memory. *Annals of Operations Research*. 2020. no. 285(1). pp. 149–159. DOI: 10.1007/s10479-019-03185-3.
20. Berlinska J., Kononov A., Zinder Ya. Two-machine flow shop with dynamic storage space. *Optimization Letters*. 2021. vol. 15. pp. 2433–2454. DOI: 10.1007/s11590-020-01645-5.
21. Souiden S., Cerqueus A., Delorme X., Rascle J.-L. Retail order picking scheduling with missing operations and limited buffer. *IFAC-Papers On Line, 21st IFAC World Congress*. 2020. vol. 53(2). pp. 10767–10772. DOI: 10.1016/j.ifacol.2020.12.2859.
22. Esfeh M.K., Shojaei A.A., Javanshir H., Damghani K.K. Solving a bi-objective flexible flow shop problem with transporter preventive maintenance planning and limited buffers by NSGA-II and MOPSO. *The International Journal of Nonlinear Analysis and Applications (IJNAA)*. 2022. vol. 13. no. 1. pp. 217–246. DOI: 10.22075/ijnna.2021.24335.2719.
23. Koh J., Bodin B. K-periodic scheduling for throughput-buffering trade-off exploration of CSDF. *ACM Transactions on Embedded Computing Systems*. 2022. vol. 22. no. 1. DOI: 10.1145/3559760.
24. Honorat A., Dardaillon M., Miomandre H., Nezan J.-F. Automated buffer sizing of dataflow applications in a high-level synthesis workflow. *ACM Transactions on Reconfigurable Technology and Systems (TRETS)*. 2024. no. 17(1). pp. 1–26. DOI: 10.1145/3626103.
25. Van Schilt I.M., Van Kalker J., Lefter I., Kwakkel J., Verbraeck A. Buffer scheduling for improving on-time performance and connectivity with a multi-objective simulation–optimization model: A proof of concept for the airline industry. *Journal of Air Transport Management*. 2024. no. 115(7). DOI: 10.1016/j.jairtraman.2024.102547.
26. Agnetis A., Pacciarelli D., Rossi F. Batch scheduling in a two-machine flow shop with limited buffer. *Discrete Applied Mathematics*. 1997. no. 72. pp. 243–260.
27. Pranzo M. Batch scheduling in a two-machine flow shop with limited buffer and sequence independent setup times and removal times. *European Journal of Operational Research*. 2004. no. 153(3). pp. 581–592.
28. Dai J. Batch Scheduling of Two-machine Limited-buffer Flow Shop with Setup and Removal Times. A Thesis for the Degree of Doctor of Philosophy in the School of Industrial and Systems Engineering. Georgia Institute of Technology. 2003. 108 p.
29. Belaid R., T'kindt V., Esswein C. Scheduling batches in flow shop with limited buffers in the shampoo industry. *European Journal of Operational Research*. 2012. vol. 223(2). pp. 560–572.
30. Rubbakhsh R., Mehdizadeh E., Adibi M. A Mathematical Model for Lot-streaming Hybrid Flow Shop Scheduling Problem by Considering Learning Effect and Buffer Capacity. *Scientia Iranica*. 2022. DOI: 10.24200/SCI.2022.58131.5582.

31. Janeš G., Ištoković D., Jurković Z., Perinić M. Application of modified steady-state genetic algorithm for batch sizing and scheduling problem with limited buffers. Applied Sciences. 2022. no. 12(22). DOI: 10.3390/app122211512.

**Krotov Kirill** — Ph.D., Dr.Sci., Associate Professor, Professor of the department, Department of "information systems", Sevastopol State University. Research interests: theory of schedules, multilevel optimization, methods of discrete optimization. The number of publications — 122. krotov\_k1@mail.ru; 33, Universitetskaya St., 299053, Sevastopol, Russia; office phone: +7(978)730-3819.

И.И. ЧИЖОВ  
**АУДИОКОДЕК НА ОСНОВЕ ПЕРЦЕПТУАЛЬНОГО  
РАВЕНСТВА ИСХОДНОГО И ВОССТАНОВЛЕННОГО  
ЗВУКОВОГО СИГНАЛА**

*Чижов И.И. Аудиокодек на основе перцептуального равенства исходного и восстановленного звукового сигнала.*

**Аннотация.** Представлен метод сжатия аудиоданных с потерями (аудиокодек), позволяющий улучшить объективное качества восстановленного аудиосигнала на 25% для битрейта 390 кбит/с и 55% для битрейта 64кбит/с по сравнению с форматом AAC MPEG-4. Предлагаемый метод сжатия аудиоданных, базируется на развитии положений теории сжатия аудиоданных с потерями (ТСАП). Повышение объективного качества восстановленного звукового сигнала (по стандартизованной мере PEAQ) достигается за счет того, что ТСАП устраняет несовершенство современных методов сжатия аудиоданных с потерями в части использования психоакустических принципов восприятия звука человеком, в том числе после преодоления «психоакустического предела сжатия» аудиосигнала (т.е. момента в перцептуальном кодировании, когда имеющегося бюджета бит недостаточно для кодирования всех спектральных компонент с необходимой с точки зрения психоакустики точностью), и позволяет достичь перцептуального равенства восприятия исходного и восстановленного аудиосигналов. В качестве анализа состояния вопроса рассмотрены решения по сжатию аудиоданных без потерь и с потерями, а также с использованием искусственного интеллекта. Во всех современных методах сжатия аудиоданных с потерями процедура выбора спектральных компонент, которые необходимо сохранить, а также допустимой погрешности квантования их по уровню выполняется путем ряда достаточно сложных процедур, носящих общее название «психоакустическая модель метода сжатия аудиоданных с потерями». В строгом смысле, перцептуальное равенство спектров исходного и восстановленного сигналов ни одна из групп исследователей не доказала и как следствие – не может его гарантировать. Независимые эксперты регулярно публикуют тесты, показывающие, что современные аудиокодеки имеют проблемы на ряде сигналов. В статье предложен аудиокодек на основе перцептуального равенства исходного и восстановленного звукового сигнала, который базируется на новых идеях теории сжатия аудиоданных с потерями (ТСАП). Эти идеи гарантируют достижение перцептуального равенства восприятия исходного и восстановленного аудиосигналов на различных битовых скоростях, поэтому построенный на ее основе аудиокодек свободен от указанных выше недостатков и, как следствие, существенно превосходит современные кодеки в смысле объективного качества восстановленного аудиосигнала по мере PEAQ.

**Ключевые слова:** сжатие аудиоданных, психоакустическая модель, спектр, субполосное разделение, перцептуальное равенство сигналов.

**1. Введение.** Обычно, сжатие аудиосигнала с потерями, т.е. процедура выбора спектральных компонент, которые необходимо сохранить, а также допустимой погрешности квантования их по уровню, выполняется путем ряда последовательных, достаточно сложных процедур, носящих общее название «психоакустическая модель метода сжатия аудиоданных с потерями» [1 – 5].

В строгом смысле, перцептуальное равенство спектров исходного и восстановленного сигналов ни одна из групп исследователей не доказала, поэтому ни один из современных коммерческих кодеков и не может его гарантировать. Другими словами, современные методы сжатия аудиоданных с потерями не могут гарантировать, что все возможные аудиосигналы будут сжаты корректно с точки зрения соблюдения принципов психоакустики. В качестве доказательства этого утверждения, стоит напомнить, что независимые эксперты регулярно проводят тесты на сложных для аудиокодеков файлах (“codec killer test”), на которых у всех методов сжатия аудиоданных, возникают проблемы на ряде сигналов.

Из всего вышеизложенного следует, что построение новой Теории сжатия аудиоданных с потерями (ТСАП), обеспечивающей достижение перцептуального равенства восприятия исходного и восстановленного аудиосигналов на любых битовых скоростях, представляется важной научной проблемой, существенной для такой области знания как кодирование звуковых сигналов.

Отметим, что метод сжатия аудиоданных, построенный на основе ТСАП, будет свободен от указанных выше недостатков и поэтому должен существенно превосходить современные кодеки в смысле объективного качества восстановленного аудиосигнала по мере PEAQ (“Objective Measurements of Perceived Audio Quality” – стандартизована Международным союзом электросвязи) [6].

**2. Аналитический обзор аудиокодеков.** В настоящее время существует два различных по сути подхода к сжатию аудиоинформации: сжатие аудиосигнала без потерь и сжатие с потерями.

**2.1. Методы сжатия аудиоданных без потерь.** Данные методы основываются на удалении статистической избыточности исходного сигнала и не вносят изменений в восстановленный сигнал по отношению к исходному. Они представляют слабый интерес с точки зрения значительного уменьшения битового представления исходного аудиосигнала. К основным алгоритмам сжатия аудиоданных (и вообще любых данных) без потерь относятся кодирование по методу Хаффмана [7, 8] и арифметическое кодирование [9].

На современном рынке представлен ряд подобных кодеков:

- FLAC [10] (Free Lossless Audio Codec) позволяет осуществить сжатие аудиоданных до 50%.
- Кодек WavPack [11], разработанный Дэвидом Бриантом, позволяет осуществлять компрессию на 30-70%.

- Кодек Monkey's Audio [12] в настоящее время активно дорабатывается. Последнее обновление было опубликовано на официальном сайте 29 сентября 2023 года. Сжимает исходный файл на 40-60%.

- Кодек Apple Lossless Audio Codec (ALAC) [13], разработанный компанией Apple, впервые был представлен в 2004 году и позволяет осуществить сжатие аудиоданных на 40-60%.

Основными методами уменьшения битового представления аудиоданных в настоящее время являются методы и алгоритмы сжатия с потерями. В основе данных методов сжатия аудиоданных лежит учет особенностей слуха человека, в частности неспособность человека различать звуки при определенных условиях. Например, невозможность различать тихие аудиосигналы после громких, неспособность слышать очень высокие и очень низкие звуки и т.д. Таким образом, психоакустическая модель, учитывающая особенности слуха человека, положена в основу методов и алгоритмов сжатия аудиоданных с потерями, которые реализованы в виде различных кодеков.

Большинство кодеков, использующихся для сжатия аудиоданных, в своей основе ориентированы на следующий алгоритм:

1. Разбиение исходного аудиосигнала на отдельные фреймы (кадры). Длина фрейма может быть различной, но, как правило, исчисляется в миллисекундах. При дальнейшем анализе входящей аудиоинформации зачастую используется не только текущий фрейм, но и предыдущий, либо фреймы берутся с перекрытием, чтобы обеспечить более полный и адекватный анализ сигнала.

2. Спектральный анализ аудиосигнала по каждому из фреймов. Преимущественно для получения информации о спектральных составляющих фрейма аудиосигнала используется модифицированное дискретное косинус-преобразование (МДКП). Также перед реализацией МДКП в некоторых алгоритмах осуществляется разделение спектра исходного сигнала на спектральные полосы, а вычисление непосредственно МДКП осуществляется по каждой отдельной полосе (как например в MPEG-1 Layer III [1]).

3. Использование психоакустической модели для принятия решения о том, какие из составляющих необходимо оставить. Использование психоакустической модели позволяет обеспечить приемлемое качество полученного аудиосигнала, основываясь на несовершенстве слуха человека, эффектах временной и частотной маскировки одних звуков другими. Таким образом, данные действия

позволяют в значительной мере уменьшить объём исходной аудиоинформации.

4. Квантование по уровню спектральных коэффициентов. Методы, используемые для квантования, а также количество уровней квантования влияют на объём и качество аудиосигнала.

5. В заключение для получения еще меньшего представления аудиосигнала используется какой-либо метод сжатия данных без потерь (метод Хаффмана или арифметическое кодирование).

**2.2. Методы сжатия аудиоданных с потерями.** Сжатие с потерями является на сегодняшний день «де-факто» стандартом в отрасли и большинство стандартизованных методов сжатия работают именно в этой парадигме.

Рассмотрим основные стандартизованные методы сжатия аудиоданных с потерями, в основе которых лежит психоакустический подход к сжатию:

**Формат MP3** [1, 2, 4]. Довольно старым, но тем не менее популярным в настоящее время является формат MP3 (полное название MPEG-1 Audio Layer III, MPEG-2 Audio Layer III и MPEG-2.5 Audio Layer III). Исторически он является лидером по распространенности и имеет хорошую совместимость с различными программами и аудиоустройствами. По техническим параметрам формат MP3 уступает большинству современных форматов.

**Формат WMA (Windows Media Audio)** [14]. Данный формат сжатия аудиоданных был создан компанией Microsoft в качестве кодека, конкурирующего с MP3. Данный формат довольно распространен, но так как он является закрытым, то, соответственно, возникают проблемы с его использованием вне Windows систем.

**Формат Vorbis** [15]. Данный формат был создан Xiph в качестве свободного формата. Vorbis является похожим на MP3, однако имеются отличия как в психоакустической модели формата, так и в математических методах и их практической реализации в используемой модели.

**Аудиокодек Opus** [16]. Opus разработан Internet Engineering Task Force (IETF). При создании аудиокодека Opus была использована комбинация технологий, а именно кодека CELT (Constrained Energy Lapped Transform) от организации Xiph и кодека SILK от компании Skype. На сегодняшний день широко внедрен и является лидером среди бесплатных кодеков. Также сравнение результатов работы кодеков Opus и OGG Vorbis приводятся в [17].

**MPEG Advanced Audio Coding (AAC).** В настоящее время именно данный алгоритм является основным стандартом в отрасли.

Этот кодекс базируется на подходах, апробированных в формате МР3, но использует значительно более эффективные технологии сжатия, чем его предшественник.

Спецификации AAC претерпели ряд изменений с момента подачи первого предложения. Для определения системы AAC, был выбран модульный подход, при котором вся система разбивается на ряд автономных модулей или инструментов, где инструмент определяется как модуль кодирования, который можно использовать как отдельный модуль, компонент общей системы.

Некоторые из этих инструментов взяты из более раннего стандарта MPEG Audio, в то время как другие являются новыми.

Если говорить очень обобщенно и оставить в стороне модули, отвечающие за стереокодирование (такие как кодирование суммарного и разностного каналов) и улучшение качества сигнала (такие как уменьшение размера используемого окна анализа в зависимости от наличия в сигнале резких всплесков), то схема работы метода AAC для каждого фрейма (1024 отсчета аудиосигнала) выглядит достаточно типично для современных методов сжатия с потерями:

- Психоакустическая модель определяет пороги маскирования, т.е. то на сколько можно изменить каждое значение спектрального компонента, чтобы эти искажения были не заметны на слух.

- Параллельно фрейм переводится в спектральную область путем применения Модифицированного дискретного косинусного преобразования (МДКП) и далее разбивается на субполосы в соответствии с Барк-шкалой.

- В каждой субполосе производится нормировка и квантование по уровню каждого спектрального компонента, таким образом, чтобы не нарушить полученные от психоакустической модели пороги маскировки (маскирования), применяя кантователи и масштабные коэффициенты (для нормировки) в соответствии со стандартом – данные значения берутся из соответствующих таблиц стандарта. Здесь стоит выделить применение нормировки, т.к. другие методы сжатия с потерями ее не используют.

- Далее применяется энтропийное кодирование (по методу Хаффмана или арифметическое).

Описанную выше часть метода сжатия с потерями принято называть перцептуальным сжатием аудиосигнала, в то время как остальные модули являются вспомогательными, однако именно они занимают значительную часть стандарта.

Более подробно с описанными процедурами можно ознакомиться, например, в рекомендациях ITU-R BS.1196-8 [18].

Здесь стоит отметить, что в указанной схеме есть несколько явных нарушений ТСАП (подробное описание будет дано в разделе 4.2).

В настоящее время используется современная версия стандарта AAC MPEG-4 [18]. Важно отметить, что существуют более современные версии кодека AAC – это AAC High Efficiency v1/2 (AAC HE v1/2) [18].

Однако, с точки зрения перцептуального сжатия аудиосигнала эти форматы не имеют отличий от AAC, а повышение качества достигается в случае AAC HE v1 за счет технологии восстановления высоких частот на основе имеющихся нижних и средних, именуемой SBR (Spectral Band Replication), в AAC HE v2 еще дополнительно появляется технология Параметрического кодирования стереопанорамы – PS (Parametric Stereo).

Достаточно очевидно, что эти технологии имеют крайне отдаленные отношения к перцептуальному сжатию аудио с потерями и поэтому не являются частью настоящего исследования.

**2.3. Методы сжатия аудиоданных с использованием нейронных сетей.** В настоящее время многие группы разработчиков, в первую очередь за рубежом, пытаются применить технологии нейросетей для компрессии звука и речи [19].

Так же вызывает большой интерес прогресс в смежных областях, таких как генерация речи из его мел-представления (что может быть использовано в качестве Декодера в системе сжатия речи) [20, 21], а так же попытка применения генеративных моделей на основе диффузии (“Diffusion”) [22] для улучшения качества речи, что может быть применено как пост-обработка в системах сжатия речи (и потенциально аудио). Надо отметить, что в сжатии речи ими были достигнуты определенные успехи, однако в сжатии аудиосигналов пока применение указанных технологий не дало существенного результата.

Основными кодеками, в основе которых лежит применение искусственного интеллекта, являются SoundStream и LPCNet.

**LPCNet.** LPCNet является гибридным кодеком [23, 24], использующим в своем составе как традиционные подходы к сжатию (в данном случае это сжатие на основе линейного предсказания значений речевого сигнала – LPC (linear predictive coding), широко применяемое в традиционных речевых кодексах), так и блок восстановления спектра речевого сигнала на основе нейросетевого



подхода (в случае LPCNet применена многослойная рекуррентная архитектура нейросети) [25, 26].

**SoundStream** [27, 28]. Основным компонентом аудиокодека SoundStream является нейронная сеть, состоящая из энкодера, декодера и квантователя, каждый из которых обучен сквозным способом (end-to-end). Кодер преобразует входной аудиопоток в закодированный сигнал, который сжимается с помощью квантователя, а затем преобразуется обратно в аудио с помощью декодера. Для повышения субъективного качества восстановленного сигнала SoundStream использует генеративно-состязательное обучение.

Дискриминатор вычисляет комбинацию функций состязательных и генеративных искажений, что повышает субъективное качество сигнала. После обучения энкодер и декодер можно запускать на отдельных машинах для эффективной передачи высококачественного звука по сети.

Кодер SoundStream создает латентные векторы, которые могут принимать нефиксированное количество значений. Чтобы передать их приемнику, используя ограниченное количество бит, необходимо заменить их близкими векторами из конечного набора (называемого кодовой книгой). Данный процесс известен как векторное квантование. Для повышения эффективности векторного квантования в SoundStream используется векторное квантование остаточных векторов – RVQ (Residual vector quantizer). Данный подход сейчас применяется большинством методов сжатия речи на основе нейронных сетей.

Стоит отметить, что существует более эффективный, чем SoundStream, метод сжатия, называющийся **EnCodec** (компания Meta) [29], но так как оба метода достаточно конструктивно схожи, приводить его подробно не имеет смысла. Отметим только, что в настоящее время именно EnCodec демонстрирует наиболее высокое качество восстановленного аудиосигнала из всех методов подобного класса. Достигается это за счет применения новых функций потерь и автоматического «балансера» весовой функции в суммарной функции потерь.

Однако, все указанные методы сжатия на основе искусственного интеллекта в настоящее время хотя и дают хорошие результаты для сжатия речи, но при сжатии аудиофайлов (например, музыки) не дают приемлемых результатов. Уместно предположить, что указанные различия в результатах обусловлены тем, что в речи есть явно выраженные паттерны (гармоники, обертона и т.д.), а в музыке либо

их вообще нет, либо они слишком разнообразны, что не позволяет нейронной сети эффективно обучиться для их генерации.

Кроме того, с увеличением битрейта качество у Нейросетевых кодеков не растет (а иногда даже немного снижается [29]). Это связано с отсутствием в их составе психоакустической модели, которая определяет, какие спектральные компоненты необходимо передать, а какие будут не слышны человеческим ухом. Таким образом, рост битрейта обеспечивает более точное квантование латентного вектора, что вообще-то не ведет непосредственно к увеличению субъективного качества восстановленного сигнала, так как далее латентный вектор проходит через «глубокий» декодер, который в достаточной мере способен компенсировать погрешность квантования латентного вектора, и точность его восстановления после квантования уже почти не играет никакой роли.

**2.4. Постановка проблемы.** Сжатие без потерь – интересная технология, но имеет маленький рынок, т.к. степень сжатия не превышает 2-3 раз, что явно недостаточно. И хотя многие стриминговые сервисы добавили их как возможные, но большой популярностью они, по оценкам самих сервисов, в настоящее время не пользуются. Кроме того, в них не используется перцептуальное сжатие, поэтому их рассмотрение выходит за рамки настоящей статьи. Сжатие на основе нейронных сетей – перспективное направление, но для сжатия аудио пока эти технологии применимы слабо (хотя и эффективны при сжатии речи). Различия на слух исходного и восстановленного файлов весьма значительны и нет способа их устранить, т.е. повышение битрейта не ведет к достижению перцептуальной близости восстановленного сигнала к исходному. В этом легко убедиться, прослушав примеры на демо-сайте EnCodec [30]. При этом стоит помнить, что авторы выбрали удачные примеры, но даже в них искажения легко заметны на слух. MP3, Vorbis: в целом эффективные, но устаревшие форматы сжатия аудио с потерями. Исходный и восстановленный файл достаточно легко различить на слух. Opus – перспективный формат, составлявший серьезную конкуренцию AAC, но в настоящее время не развивается и поэтому отстал от AAC. Различить на слух исходный и восстановленный файл возможно (особенно на низких битрейтах). AAC – наиболее совершенный формат сжатия с потерями на сегодняшний день. Исходный и восстановленный сигнал на слух достаточно близки, но объективная оценка (например, по мере PEAQ) сразу выявляет существенные искажения в восстановленном сигнале. Кроме того, в ряде случаев нарушаются психоакустические принципы

восприятия звука человеком с целью снижения необходимого бюджета бит (например, когда удаляются верхние частоты на низких битрейтах). Теоретического доказательства перцептуального равенства исходного и восстановленного сигнала авторы данного метода также не приводят.

Из приведенного анализа следует, что, несмотря на высокую эффективность современных методов сжатия аудиоданных с потерями, ни один из них не гарантирует перцептуального равенства исходного и восстановленного звукового сигнала.

Таким образом, возникает необходимость развития теории сжатия аудиоданных с потерями, позволяющей осуществлять высокоэффективное сжатие аудиоданных с потерями на основе психоакустических принципов восприятия звука человеком и гарантировать перцептуальное равенство исходного и восстановленного аудиосигнала на любых битовых скоростях, поддерживаемых современными стандартизованными аудиокодеками.

**3. Уточненная теория сжатия аудиоданных с потерями.** Достаточно очевидно, что целью развития теории сжатия аудиоданных с потерями является достижение равенства исходного и восстановленного сигналов. Так как речь идет о сжатии с потерями, то равенство в строгом математическом смысле невозможно и можно говорить лишь о перцептуальном (психоакустическом) равенстве исходного и восстановленного сигналов.

Рассмотрим следующую теорему.

**3.1. Теорема о перцептуальном равенстве исходного и восстановленного аудиосигналов.**

**Теорема:** При сжатии аудиоданных с потерями достижение перцептуального равенства исходного и восстановленного аудиосигналов достижимо при разбиении спектра сжимаемого аудиосигнала на субполосы, соответствующие требованиям психоакустики, и выполнении условия неизменности соотношения энергии и фазы исходного и обработанного (сжатого) сигнала в каждой субполосе.

В данном случае фазу необходимо рассматривать как величину, определяющую уровень звукового давления известной звуковой волны в текущий момент времени в измеряемой точке пространства. Другими словами, фаза определяет (совместно с амплитудой) энергию сигнала для каждой частотной группы в данный момент времени, а т.к. длительность фрейма не превышает 20 мс. (что существенно меньше времени реакции слухового аппарата человека на раздражитель), то и среднюю энергию по фрейму допустимо считать мгновенной с точки

зрения перцептуального равенства исходного и восстановленного аудиосигналов.

Стоит отметить, что при переходе в спектральную область при помощи ДКП или МДКП в каждый коэффициент такого разложения включена как энергия, так и фаза сигнала. Это является одной из основных причин, почему в сжатии аудиосигналов МДКП является на сегодняшний день наиболее широко применяемым разложением.

Второй причиной является то, что МДКП оперирует в действительной области, а не в комплексной (как, например Дискретное преобразование Фурье (ДПФ)), что существенно упрощает машинные вычисления с применением данного разложения.

Применение МДКП (или любого другого преобразования сохраняющего информацию о фазе сигнала) для перевода аудиосигнала в спектральную область при его сжатии является необходимым условием, т.к. восстановление сигнала без информации о его фазе внесет значительные искажения в восстановленный сигнал и является недопустимым с точки зрения предлагаемой ТСАП.

Именно поэтому методы, использовавшие в качестве разложения ДПФ с последующим вычислением энергетического спектра (и как следствие – потерей фазы), впоследствии были заменены на методы, сохраняющие информацию о фазе, в том числе это касается и методов на основе нейронных сетей. Более того, для восстановления фазы сигнала был разработан специальный метод, именуемый PHASEN [31], и ряд более современных его аналогов, получивший широкое применение в сфере улучшения качества речи с применением нейросетевых подходов.

Таким образом, сохранение фазы сигнала так же является важной научной задачей, обеспечивающей перцептуальное равенство исходного и восстановленного сигналов, а в случае многоканального сигнала – еще и все бинауральные эффекты человеческого слуха, например позиционирование источника сигнала в пространстве.

Доказательство:

Исходя из взаимной однозначности представления сигнала в спектральном и временном доменах

$$X(n)_{ucx} = X(n)_{восм}, \quad (1)$$

если

$$F(n)_{ucx} = F(n)_{восм}, \quad (2)$$

где  $X$  – цифровой сигнал, а  $F$  – спектр этого сигнала (здесь и далее большими буквами обозначены вектора вида  $x(t)$ , т.к. дискретный аудиосигнал – это вектор амплитуды от времени),  $n = 1, 2, \dots, N$ , где  $N$  – количество временных/спектральных отсчетов.

$$F = (f_1, f_2, \dots, f_N), \quad (3)$$

где  $f$  – спектральные коэффициенты (в нашем случае – МДКП),  $N$  – количество спектральных коэффициентов.

Для достижения (2) воспользуемся тем, что в пределах критической полосы слух интегрирует возбуждение по частоте и не различает тонкой структуры возбуждения. На этом, в частности, строится эффект маскирования в спектральной области. Другими словами, спектр звукового сигнала может рассматриваться как набор сумм значений коэффициентов МДКП (включающих энергию и фазу сигнала) в частотных группах, соответствующих критическим полосам слуха.

Напомним, что при восприятии звука слуховой аппарат человека разделяет его на частотные группы, называемые критическими полосами слуха. В диапазоне от 20 до 22050 Гц число критических полос равно 25. Ширина этих полос меняется от низких к высоким частотам нерегулярным образом.

Таким образом, спектр сигнала можно рассматривать как 25 значений энергии сигнала в соответствующих критических полосах слуха. Для получения указанных значений необходимо применить соответствующее преобразование (например, МДКП) с последующей группировкой частот в соответствии с неравнополосной шкалой БАРК (Bark-scale).

На основании вышеизложенного сформулируем первый Принцип теории сжатия аудиоданных с потерями.

**Принцип 1.** Необходимо использовать субполосное разделение спектра сжимаемого аудиосигнала в соответствии с критическими полосами слуха (желательно в соответствии с Барк-шкалой), а не на субполосы произвольной ширины.

Стоит отметить, что данный подход применяется во многих современных аудиокодеках, что только подтверждает корректность предлагаемой теории сжатия аудиоданных с потерями.

При переходе к рассмотрению спектра сигнала в субполосном представлении (3) можно записать как

$$F = (Z_1, Z_2, \dots, Z_R), \quad (4)$$

где  $Z_i = (f_1, f_2, \dots, f_L)$ , где  $L$  – количество частотных компонент в субполосе,  $R$  – количество субполос.

На основе Принципа 1 формулу (4) можно записать как

$$F = (z_1, z_2, \dots, z_R), \quad (5)$$

где  $z_i = \sum_{j=1}^L f_j^2$ , т.е. сумма спектральных компонент в субполосе.

Как уже отмечалось выше, в данном случае речь идет об энергии сигнала в конкретный момент времени, а она определяется амплитудой и фазой сигнала. Таким образом, фаза сигнала всегда присутствует в энергии каждой субполосы в неявном виде, поэтому не требуется ее оценивать дополнительно. Это объясняет тот факт, почему во всех источниках, относящихся к перцептуальному сжатию аудиосигналов (например, [5]) речь всегда идет только об энергии сигнала и не упоминается о его фазе.

Так же отметим, что т.к. фаза не вычисляется в явном виде (а только влияет на энергию), то возведение спектрального компонента в квадрат не может на нее повлиять. Кроме того, МДКП (в отличие от ДПФ) позволяет легко восстановить значение спектрального компонента после его возведения в квадрат – необходимо тривиально взять квадратный корень (т.к. это действительное число) и получить знак из дополнительного потока знаков. Достаточно очевидно, что знаки МДКП коэффициентов необходимо сохранять и включать в сжатый аудиопоток (файл).

В качестве альтернативы МДКП уместно рассмотреть пакетное дискретное вейвлетное преобразование (ПДВП), изучению которого посвящено достаточное количество работ. Указанный подход к переходу в спектральную область позволяет получить частотные группы в соответствии с Барк-шкалой и поэтому в полной мере соответствует Принципу 1.

Перцептуальное сжатие аудиоданных с применением ПДВП так же достаточно изученная область знаний. Значительный вклад в ее развитие внесли д.т.н. Петровский А.А. [17], Фадеев Д.Р., д.т.н. Рогозинский Г.Г. [32], и другие авторы. В упомянутых работах доказано, что ПДВП применимо в качестве базиса для использования в психоакустической модели (ПАМ) и авторы успешно строят ПАМ на его основе.

Вместе с тем, несмотря на полное соответствие Принципу 1 и доказанную возможность строить ПАМ с применением ПДВП, именно вопрос точного сохранения энергии и фазы сигнала при применении ПДВП (например, во время квантования по уровню ПДВП коэффициентов) авторами не рассматривался, и более того в работе [32] идет прямое противопоставление пространства время-частота (используемому, например, в МДКП) пространству время-масштаб при использовании ПДВП.

Данное обстоятельство не позволяет прямо сейчас рекомендовать ПДВП как базис перехода в спектральную область в соответствии с Принципом 1 (о причинах этого будет сказано ниже в формуле (6)), но в случае проведения новых исследований, доказывающих его применимость, планируется вернуться к его рассмотрению.

Если потребовать при осуществлении любых изменений спектра при сжатии аудиосигнала с потерями выполнения условия

$$\frac{z_i}{\tilde{z}_i} = const = a, \quad (6)$$

где  $\tilde{z}_i$  – сумма спектральных компонент в субполосе сжатого (искаженного) аудиосигнала, то спектр сжатого аудиосигнала в соответствии с (5) будет равен

$$F_{сж} = (\tilde{z}_1, \tilde{z}_2, \dots, \tilde{z}_R), \quad (7)$$

где  $\tilde{z}_i = \frac{z_i}{a}$ , где  $a=const$  (одинаковая для всех субполос), следовательно, спектр восстановленного сигнала будет равен

$$F_{восм} = (a\tilde{z}_1, a\tilde{z}_2, \dots, a\tilde{z}_R), \quad (8)$$

т.е.

$$F_{восм} = \left( a \frac{z_1}{a}, a \frac{z_2}{a}, \dots, a \frac{z_R}{a} \right) = (z_1, z_2, \dots, z_R) \Rightarrow, \quad (9)$$

$$F_{восм} = F_{исх}. \quad (10)$$

Таким образом, соблюдение требования (6) во время сжатия аудиосигнала с потерями обеспечивает выполнение (2), т.е. исходный сигнал будет равен восстановленному в перцептуальном смысле.

Более того, так как энергия сигнала в субполосе определяется в том числе его фазой (точнее это суперпозиция сигналов соответствующих частот входящих в указанную субполосу, каждый из которых обладает своей фазой), то требование (6) позволяет так же сохранять фазовые соотношения между субполосами, что позволяет нам говорить не только о равенстве энергий исходного и восстановленного сигналов (с точностью до множителя), но и о равенстве их фаз, что однозначно определяет исполнение (1).

Стоит отметить, что когда речь идет о сжатии с потерями, то равенство исходного и восстановленного сигналов (2) принято понимать как приближенное (даже в перцептуальном смысле мы можем говорить только о «прозрачности» – т.е. о неотличимости сигналов на слух)

$$X_{восст} \approx X_{исх}, \quad (11)$$

поэтому тот факт, что в общем случае на реальных аудиоданных точного выполнения (6) достичь не удастся, говорит о том, что (6) более уместно записать как

$$\frac{z_i}{\tilde{z}_i} = a \pm \varepsilon, \text{ где } \varepsilon \gg a. \quad (12)$$

что не идет в разрез с общепринятыми в науке определениями.

Другими словами, т.к. речь идет о сжатии с потерями, то выполнение (6) в смысле выполнения (12) является достаточным, а поиск точного решения для (6) представляется избыточным и противоречащим самой идее сжатия с потерями, т.к. смысл сжатия с потерями состоит в передаче более важных с психоакустической точки зрения компонент сигнала (временных или спектральных) и отбрасывании менее важных в соответствии с требуемым субъективным качеством восстановленного сигнала, что подразумевает возможность достижения только приближенного равенства исходного и восстановленного сигналов в смысле (11). Вместе с тем, нельзя не отметить тот факт, что ни один из описанных в литературе современных методов сжатия не выполняет (6) в своей



работе (даже в смысле (12)), а следовательно, и (2) (даже в смысле (11)) также не выполняется.

Предлагаемая теория сжатия аудиоданных с потерями призвана устранить указанные несовершенства современных методов сжатия аудиоданных с потерями и позволить достигать (2), в смысле (11). В дальнейшем, для упрощения вычислений, решения будут приведены в смысле (6) и (2), т.к. это не влияет на результат вычислений.

Таким образом, на основании вышеизложенного доказательства сформулируем второй Принцип теории сжатия аудиоданных с потерями.

**Принцип 2.** На всех этапах процедуры сжатия-восстановления аудиосигнала необходимо выполнение (6), т.к. это в свою очередь является необходимым и достаточным условием выполнения (2).

Теорема доказана.

**3.2. Уменьшение битовых представлений аудиоданных с соблюдением перцептуального равенства исходного и восстановленного аудиосигналов.** Стоит отметить, что уменьшение битовых представлений аудиоданных (аудиосигналов) возможно исключительно за счет максимально возможного снижения энтропии сжатой (обработанной) последовательности и последующего применение статистического сжатия (например, арифметического кодирования).

Снижения энтропии сжатого аудиофайла возможно за счет квантования по уровню отсчетов исходного сигнала или его спектральных компонент (в том числе и квантования к нулю).

**3.2.1. Разделение на тональные и шумоподобные компоненты.** С точки зрения психоакустики удаление или обнуление произвольных отсчетов сигнала или его спектральных компонент совершенно неверно. Исключением являются спектральные компоненты, лежащие ниже абсолютного порога слышимости, которые не воспринимаются человеком в принципе ни при каких условиях – поэтому их можно сразу отнести к шумоподобным, и это не скажется на достижении перцептуального равенства исходного и восстановленного аудиосигналов, особенно учитывая тот факт, что энергия подобных компонент крайне незначительна и подобные компоненты есть в каждой субполосе, и их энергия в достаточной мере уравновешивает друг друга.

Для всех остальных спектральных компонент звукового сигнала, допустимо только квантовать отсчеты сигнала или спектральные компоненты в пределах, которые будут незаметны для

слушателя восстановленного аудиосигнала, базируясь на «эффекте маскировки».

Во всех современных методах сжатия аудиоданных с потерями для использования данного эффекта строят так называемые кривые маскировки, позволяющие определить, маскируется ли данный отсчет или спектральный коэффициент соседними или нет и насколько сильно возможно его проквантовать.

Однако, в реальной практике современных аудиокодеков в случае недостатка бюджета бит в процессе квантования по уровню при достижении кривыми маскировки уровня, который не позволяет присвоить квантованному значению никакого уровня, кроме нуля, возникает необходимость зануления части спектральных компонент.

Это является важным нарушением принципов психоакустики и подлежит устранению.

Кроме того, важным является тот факт, что принятие решения об отбрасывании (обнулении) текущего компонента осуществляется на основе эвристически полученных коэффициентов (используемых при построении кривых маскировки), значения которых вызывают существенные разночтения (проще говоря – они у каждой психоакустической модели свои, и их значения различаются значительно, например, у психоакустических моделей MPEG AAC и Dolby AC-4 обе из которых являются современными и эффективными решениями), что позволяет делать вывод о не точном соответствии психоакустической теории.

Прежде чем говорить о способах преодоления зануления спектральных компонент на низких битрейтах, поговорим о корректной с точки зрения психоакустики маскировке спектральных компонент.

Если в процессе проведения процедуры сжатия разделять субполосы в соответствии с критическими полосами слуха, то решение задачи маскировки достаточно очевидно – любой спектральный компонент в пределах субполосы маскируется спектральными компонентами с большей энергией.

Это обусловлено самим определением критической полосы слуха, т.к. в ней кривая маскировки любого маскера линейна (находится на одном уровне, без убывания), а убывание возникает за пределами критической полосы слуха (собственно, так и возникло понятие «критической полосы слуха» – ширины спектра, где кривая маскировки имеет линейный вид). Наличие линейной части в кривой маскирования представляется крайне важным, т.к. полное описание эффекта маскировки в критической полосе слуха позволяет построить

перцептуально обоснованную теорию сжатия аудиоданных с потерями.

Таким образом, каждый компонент спектра в субполосе маскируется главным маскером, и все, что лежит ниже кривой маскирования главного маскера, можно рассматривать как шумоподобные компоненты.

Важным моментом является то, что при таком подходе не происходит зануления спектральных компонент, т.к. все шумоподобные компоненты будут закодированы одним значением, вычисляемым как

$$Z_n = \frac{\sum f_{nz}}{N}, \quad (13)$$

где  $N$  – количество шумоподобных компонент в субполосе,  $f_{nz}$  – шумоподобные спектральные компоненты в субполосе

Данный подход представляется обоснованным, т.к. согласно, например, [5] в пределах критической полосы слух реагирует лишь на общую энергию шума, не учитывая его распределение.

Таким образом, представление спектральных компонент с низкой энергией в пределах критической полосы слуха как шумоподобных и дальнейшая замена их на среднее значение их энергии полностью согласуется с принципами психоакустики. На этой основе сформулируем Принцип 3 теории сжатия аудиоданных с потерями.

**Принцип 3.** В целях выполнения (6) и, как следствие, (2), необходимо не допускать зануления спектральных компонент с малой энергией, а вместо этого, в случае нехватки бюджета бит для кодирования всех спектральных компонент с необходимой с точки зрения психоакустики точностью (т.е. после преодоления психоакустического предела сжатия), кодировать их как шумоподобные, т.е. одним значением (средним значением их энергии), вычисляемой при помощи (13) с сохранением знака.

Остается показать, каким образом осуществить указанную операцию, чтобы не был нарушен принцип (6).

С точки зрения психоакустической теории, указанные спектральные компоненты с малой энергией допустимо отнести к шумоподобным, и, как следствие, закодировать только средним значением, если их энергия ниже кривой маскирования главного маскера.

Уровень кривой маскирования главного маскира ( $M_i$ ) в децибелах по отношению к значению самого маскира можно вычислить по следующей формуле (14), позволяющей рассчитать уровень маскирования шума тоном [1]:

$$M_i = -1,525 - 0,175 * z_i - 0,5, \quad (14)$$

где  $z_i$  – номер текущей субполосы.

Как уже отмечалось ранее, в качестве дополнительного шага можно сразу отнести к шумоподобным спектральные компоненты, лежащие ниже абсолютного порога слышимости, как не воспринимаемые человеческим слухом.

В общих чертах, в соответствии с предлагаемой теорией сжатия с потерями, процедура выделения шумоподобных спектральных компонент должна выглядеть следующим образом:

В случае нехватки бюджета бит выполняется следующее:

1. Ранжирование спектральных компонент в субполосе по модулю.
2. Помечание части спектральных компонент, лежащих ниже кривой маскирования главного маскира  $M_i$  (в соответствии с (14)) как шумоподобных, начиная с наименьшего, до тех пор, пока битовая скорость потока не будет соблюдена.
3. Замена помеченных шумоподобных спектральных компонент на среднее значение их энергии и (если это предусматривает метод сжатия) квантование по уровню полученного значения.

Следовательно, т.к. после применения процедуры выделения шумоподобных спектральных компонент во всех субполосах будет выполнено (6), то это позволяет говорить о выполнении (2).

Количество шумоподобных спектральных компонент определяется целевым битрейтом, т.е. тем бюджетом бит, которым обладает энкодер: чем ниже битрейт, тем их берем больше, т.к. это очевидно позволяет уменьшить расход бит за счет передачи не точных значений конкретных спектральных коэффициентов, а только одного значения  $Z_n$  для всех. Напомним, что современные аудиокодеки эти компоненты просто зануляют.

Кстати говоря, степень сжатия при предлагаемом подходе уменьшится незначительно, т.к. нужно дополнительно передавать только значение  $Z_n$  для каждой субполосы. При этом перерасход бит компенсируется за счет более высокого субъективного качества получаемого сигнала, что позволит, в свою очередь, применить

квантование по уровню спектральных компонент на меньшее число уровней квантования.

Важно, что, т.к. имеет место замена спектральных компонент на среднее значение их энергии, то эта процедура не влияет на корректность процедуры сжатия в смысле выполнения (6).

Однако, (6) будет необратимо нарушено, если в шумоподобные компоненты попадут маскиры, т.е. спектральные компоненты, лежащие выше кривой маскирования главного маскира и не относящиеся к шуму. Это грубое нарушение теории и его необходимо избегать.

Для реализации указанной процедуры прореживания предлагается использовать итерационную процедуру:

Вектор  $Z_i$ , отсортированный по возрастанию, будем называть  $Z_{i\_corm}$

$$Z_{i\_corm} = (f_{corm1}, f_{corm2}, \dots, f_{cormL}), \quad (15)$$

где  $L$  – количество  $f$  в субполосе.

Достаточно очевидно, что

$$z_i = z_{i\_corm} = \sum_{j=1}^L f_j^2. \quad (16)$$

Для определения шумоподобных компонент в соответствии с (14) вычисляем реальное значение кривой маскирования с использованием обратной формулы для вычисления децибелов:

$$\tilde{M}_i = f_{max}^2 * \sqrt[20]{M_i}, \quad (17)$$

Следовательно, для выполнения (6) выполняем итерации по  $j$  пока

$$f_{cormj}^2 \leq \tilde{M}_i, \quad (18)$$

Если

$$f_{cormj}^2 > \tilde{M}_i, \quad (19)$$

то  $j = j - 1$ , т.е. отмена пометки текущего спектрального коэффициента как шумоподобного и конец итерационной процедуры.

Следовательно, после (19) получаем вектор  $Z_{\text{сорт}}$

$$Z_{\text{сорт}} = (f_{\text{сорт}k}, f_{\text{сорт}k+1}, \dots, f_{\text{сорт}L}), \quad (20)$$

где  $k = j$ , т.е. последнему успешному шагу итерационной процедуры.

Проводим обратную сортировку, т.е. возвращаем  $f$  на исходные места и завершаем итерационную процедуру.

Далее, как уже говорилось, применяем (13) и проводим замену всех шумоподобных спектральных компонент на среднее значение их энергии ( $Z_n$ ).

После выполнения процедуры получаем вектор  $Z_i$  с энергией, равной исходному значению в строгом соответствии с (6), т.к. замена значений энергии каждого из шумоподобных спектральных компонент на среднее значение их энергии не вносит изменения в энергию в субполосе.

Еще раз отметим тот факт, что указанную процедуру кодирования шумоподобных компонент стоит проводить только в случае нехватки бюджета бит (когда современные кодеки начинают занулять спектральные компоненты).

Это обусловлено тем, что хотя с точки зрения психоакустики маскирование шума тоном и реакция слуха человека на общую энергию шума и не вызывают сомнений, но сама структура музыкального (звукового) сигнала не всегда включает реальную шумовую составляющую, для которой было рассчитано соотношение (14), например, когда в сигнале присутствует только один чистый тон, то никакого реального шума в нем нет и Принцип 3 неприменим. При этом достаточно очевидно, что для кодирования такого простого сигнала и не может потребоваться Принцип 3 – бюджет бит позволяет ограничиться остальными Принципами.

**3.2.2. Квантование по уровню.** Если проводится квантование по уровню  $Z_n$ , то необходимо руководствоваться тем же Принципом, как и при квантовании маскеров (тональных компонент):

**Принцип 4.** При квантовании по уровню спектральных компонент аудиосигнала необходимо, чтобы погрешность квантования не превышала едва заметных изменений звука в смысле изложенном Цвайкером [33] (это значение лежит в диапазоне 1-2 дБ в зависимости от энергии спектрального компонента) и чтобы квантованная последовательность соответствовала (6). Метод квантования при этом

может быть любым – это не влияет на предлагаемую ТСАП, главное – это выполнение (6).

В статье применен метод квантования по уровню с адаптивным шагом квантования, позволяющий естественным образом контролировать погрешность квантования на каждом уровне и держать ее в диапазоне 1-2 дБ, что гарантирует выполнение (6). Разумеется, что значения шагов квантования необходимо так же передавать на приемную сторону. Детали метода квантования по уровню не имеют важного значения для предлагаемой Теории, поэтому выходят за рамки настоящей статьи.

**3.3. Общая схема Аудиокодека на основе ТСАП.** Аудиокодек на основе перцептуального равенства исходного и восстановленного звукового сигнала представляет собой программно-алгоритмическую реализацию указанной ТСАП и использует все 4 указанных Принципа в явном виде. Разработанный аудиокодек в настоящее время работает на битовых скоростях 32, 64, 128, 195 кбит/с (для моносигнала), в моно, стерео и многоканальных режимах. Сжатие аудиосигнала с потерями на любой битовой скорости может осуществляться как независимо для каждого канала, так и с учетом общего битрейта в многоканальном сигнале, что позволяет более эффективно использовать доступный бюджет бит и кодировать более сложные каналы большим числом бит.

Общая схема работы предлагаемого метода сжатия аудиоданных с потерями (аудиокодека) на основе ТСАП представлена на рисунке 1.

Легко видеть, что в структуре кодека отсутствует модуль психоакустической модели. Как было показано выше, аудиокодеки, построенные на основе ТСАП, не нуждаются в данном модуле, т.к. перцептуальное равенство исходного и восстановленного сигналов обеспечено выполнением (12) и Принципов ТСАП, которые в свою очередь обеспечивают выполнение психоакустических основ восприятия звука человеком и позволяют выполнить (12).

Указанное обстоятельство существенно снижает вычислительную сложность разработанного кодека по сравнению с современными кодеками (такими как AAC) и при этом (как будет показано далее) позволяют получить более высокое качество восстановленного сигнала. Все это является следствием применения ТСАП.





– Кодирование шумоподобных спектральных компонент средним значением их энергии (Принцип 3).

– Контроль выполнения соотношения (12) для каждой из субполос. Стоит отметить, что  $\epsilon$  выбирается в соответствии с битрейтом и должно быть существенно меньше «а». Разумно определить  $\epsilon$  как долю от «а» (Принцип 2).

Точная оценка вычислительной сложности разработанного аудиокодека будет выполнена в ходе дальнейших исследований, но даже приблизительная оценка этапов сжатия показывает потенциальное снижение вычислительной сложности по сравнению с современными аудиокодеками, использующими полноценную «психоакустическую модель».

Отметим, что кодек несимметричен, т.к. на стороне энкодера возможно несколько итераций по подбору параметра «а» и возможен поиск шумоподобных компонент сигнала.

**4. Вычислительные эксперименты по сравнению объективного качества сжатых аудиофайлов.** Для подтверждения работоспособности предложенного в работе Аудиокодека на основе перцептуального равенства исходного и восстановленного звукового сигнала были проведены вычислительные эксперименты по сравнению результатов сжатия с потерями исходного звукового сигнала:

- по технологии Advanced Audio Coding (AAC);
- на основе предлагаемого Аудиокодека на основе перцептуального равенства исходного и восстановленного звукового сигнала.

**4.1. Алгоритм вычислительного эксперимента.** В качестве исходных сигналов использовались музыкальные файлы различных жанров, направлений и времени издания (около 1000 файлов), подобранные таким образом, чтобы максимально полно представить всю генеральную совокупность музыкальных произведений, записанных и используемых в настоящее время.

В качестве сравнения использовался стандартизованный формат сжатия аудиоданных с потерями (по Рекомендации Международного союза электросвязи ITU-R BS.1196-8 [18], более известный как Advanced Audio Coding (AAC MPEG-4), который в настоящее время является стандартом в отрасли как де-юре, так и де-факто.

Сравнение полученных результатов с результатами стандартизованного кодека Advanced Audio Coding (AAC) осуществлялось в смысле объективного качества

(по стандартизированной мере PEAQ) и субъективного качества восстановленного сигнала, оцениваемого на слух группой экспертов.

Мера PEAQ (Objective Measurements of Perceived Audio Quality) стандартизована Международным союзом электросвязи и является общепризнанной на сегодняшний день как мера оценки объективного качества звукового сигнала. PEAQ измеряется в условных единицах, которые соответствуют MOS (Mean opinion score – усредненное мнение экспертов) и лежит в диапазоне от 5,0 – качество оригинала, до 1,0 – качество ужасно (невозможно или очень трудно узнать оригинальный сигнал) – Рекомендация Международного союза электросвязи ITU-R BS.1284-2 [34].

Необходимо отметить, что в реальной практике для оценки объективного качества часто используется шкала ODG (Objective Difference Grade) [6].

В таблице 1 приведены эквивалентные значения качества по шкалам MOS и ODG, а также описание этого качества, как это приводится в Рекомендациях Международного союза электросвязи [34] и Рекомендации ITU-R BS.1387-2 [6].

Таблица 1. Эквивалентные значения качества по шкале MOS и ODG

Тип вносимых искажений	MOS (ITU)	ODG
Незаметные искажения	5,0	0,0
Заметные, но не раздражающие	4,0	-1,0
Немного раздражающие	3,0	-2,0
Раздражающие	2,0	-3,0
Крайне раздражающие	1,0	-4,0

**4.2. Анализ метода AAC на соответствие ТСАП.** Прежде чем привести результаты вычислительных экспериментов, проведем теоретический анализ технологии AAC на соответствие ТСАП.

Как несложно заметить из [18] и краткого описания, приведенного в разделе 2.2, в целом AAC в достаточной степени соответствует принципам психоакустики, что вполне ожидаемо для лучшего современного коммерческого аудиокодека.

Однако, есть ряд тонких моментов, в которых AAC нарушает ТСАП:

1. Отсутствует контроль выполнения основного соотношения (12), что не позволяет говорить о гарантированном выполнении (2), подробно указанные соотношения рассмотрены в разделе 3.1.

2. Начиная с битрейта 64 кбит/с для каждого канала, метод AAC начинает занулять часть верхних частот, чтобы сэкономить

бюджет бит для нижних и средних частот. Это очевидное нарушение (2), т.к. часть спектра сигнала просто отсутствует, и если для средних битрейтов (например, 64 кбит/с) зануляются только частоты выше 18кГц, (что, разумеется, тоже нарушение, но разницу услышат только молодые люди), то для 32 кбит/с зануляются частоты выше 14кГц, что будет слышно практически всем. Отметим, что аудиокодек на основе ТСАП не имеет указанного недостатка – он всегда кодирует полный спектр сигнала.

3. Несмотря на то, что общая схема ААС в целом соответствует психоакустическим принципам, проблемы начинаются после преодоления психоакустического порога сжатия, т.е. когда имеющегося бюджета бит начинает не хватать для кодирования всех спектральных компонент в соответствии с порогами маскирования. В этот момент авторы ААС вынуждены идти на осознанное нарушение психоакустических принципов и понижать пороги маскирования. Часто это приводит к тому, что многие спектральные компоненты не могут быть проквантованы ни к какому другому уровню кроме нулевого и поэтому зануляются, что является грубым нарушением Принципа 3 ТСАП (подробное описание приведено в разделе 3.2.1).

Таким образом, несложно заключить, что, несмотря на свою очевидную эффективность, метод ААС нарушает ТСАП и не может гарантировать перцептуального равенства исходного и восстановленного сигнала.

**4.3. Вычислительный эксперимент.** Проведенный вычислительный эксперимент подтвердил справедливость теоретических выводов, приведенных выше.

Сжатие аудиоданных осуществлялось с битовой скоростью 390 кбит/с для стереосигнала или 195 кбит/с для каждого канала кодируемого независимо и 64 кбит/с для стереосигнала или 32 кбит/с для каждого канала кодируемого независимо.

В качестве примера приведем результаты 10 специально выбранных тестовых файлов, которые были определены группой MPEG как рекомендуемые для тестирования аудиокодексов.

**4.3.1. Вычислительный эксперимент на битовой скорости 390 кбит/с.** В Таблице 2 представлены результаты ODG для каждого сжатого звукового файла для битовой скорости 390 кбит/с, сравнение с результатами ААС на равном битрейте, разность между ААС и предлагаемым методом сжатия с потерями в абсолютных значениях и процентах, а также усредненная оценка по всем файлам.

Таблица 2. Результаты сопоставительного анализа результатов сжатия аудиофайлов при помощи кодека AAC MPEG-4 и предлагаемого аудиокодека по шкале ODG для битрейта 390кбит/с

№	Имя файла	AAC, [ODG]	Предлагаемый метод, [ODG]	Разность (Предлагаемый – AAC), [ODG]	Разность, [%]
1	Adel	-0,076	-0,034	0,042	55,26
2	Avril_Lavil	-0,062	-0,024	0,038	61,29
3	Elvis_Prestly	-0,124	-0,12	0,004	3,23
4	Evanescence	-0,064	-0,06	0,004	6,25
5	Hotel_California	-0,148	-0,146	0,002	1,35
6	Jump_in_my_car	-0,118	-0,094	0,024	20,34
7	NewYorkCity	-0,074	-0,065	0,009	12,16
8	Poison	-0,069	-0,041	0,028	40,58
9	SlavesOfFear	-0,054	-0,026	0,028	51,85
10	You_are_yong	-0,075	-0,073	0,002	2,67
<b>Среднее по всем:</b>		<b>-0,086</b>	<b>-0,068</b>	<b>0,018</b>	<b>25,50</b>

Несложно заметить, что, по усредненной оценке, преимущество предлагаемого аудиокодека составило 25%, а для ряда файлов превысило 50%, при этом важно помнить, что речь идет о высоких битрейтах, где у современных методов сжатия аудиоданных с потерями нет проблем с бюджетом бит для кодирования сжатого звукового потока.

Таким образом, доказано, что применение Аудиокодека на основе перцептуального равенства исходного и восстановленного звукового сигнала позволяет превзойти результаты современных аудиокодеков даже на высоких битрейтах, что в современной науке считается малодостижимой задачей, т.к. при наличии достаточного бюджета бит AAC нарушает ТСАП только в части отсутствия контроля выполнения (12), а остальных нарушений в явном виде нет, но даже этого вполне достаточно, чтобы предложенный аудиокодек превзошел AAC.

На рисунке 2 представлены спектрограммы исходного аудиофайла и восстановленных файлов, сжатых при помощи кодека AAC и предлагаемого аудиокодека.

Стоит отметить, что на подобном высоком битрейте отличия восстановленного аудиосигнала от оригинала должны быть крайне незначительны – тогда можно говорить о корректной работе аудиокодека. В нашем случае отличия составляют 0,068 ODG, что крайне мало, и увидеть такое отличие на спектрограмме крайне сложно.

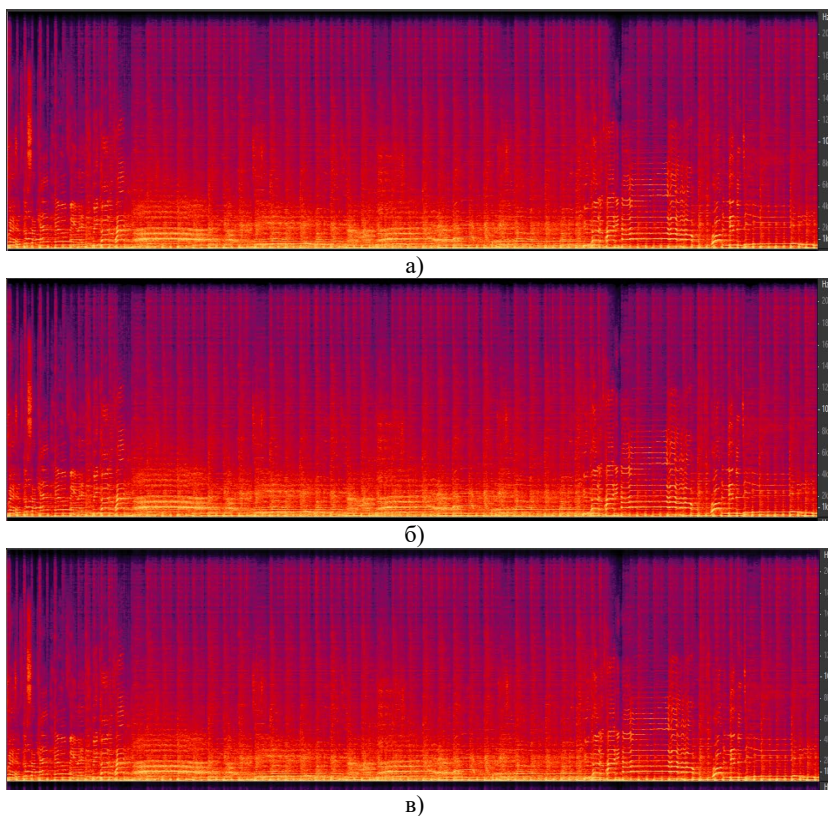


Рис. 2. Сравнение спектрограмм: а) Оригинал, б) AAC, в) Предлагаемый метод (файл №1)

Приведенные спектрограммы зрительно практически неотличимы друг от друга, что подтверждает корректность работы предлагаемого кодека и достижения им качества оригинала.

Таким образом, для корректного сравнения результатов предложенного аудиокодека и формата Advanced Audio Coding (AAC MPEG-4) необходим инструмент более точный, чем зрительное сравнение спектрограмм – именно поэтому мы опираемся на результат объективного оценивания по мере PEAQ, которая показывает явное превосходство предложенного аудиокодека.

Также отметим, что оценка субъективного качества восстановленного звукового сигнала по шкале MOS носила в данном исследовании вспомогательный характер и была направлена на

контроль ошибок со стороны объективной меры оценки качества восстановленного звукового сигнала PEAQ и контроля качества воспроизводимого сигнала на слух.

Для битрейта 390 кбит/с MOS был оценен всеми экспертами как близкий к оригиналу, но вместе с тем большее число экспертов признали результат, полученный на основе предлагаемой теории, более насыщенным и близким к оригиналу.

**4.3.2. Вычислительный эксперимент на битовой скорости 64 кбит/с.** В Таблице 3 представлены результаты ODG для каждого сжатого звукового файла для битовой скорости 64 кбит/с, сравнение с результатами AAC на равном битрейте, разность между AAC и предлагаемым методом сжатия с потерями в абсолютных значениях и процентах, а также усредненная оценка по всем файлам.

Таблица 3. Результаты сопоставительного анализа результатов сжатия аудиофайлов при помощи кодека AAC MPEG-4 и предлагаемого аудиокодека по шкале ODG для битрейта 64кбит/с

№	Имя файла	AAC, [ODG]	Предлагаемый метод, [ODG]	Разность (Предлагаемый – AAC), [ODG]	Разность, [%]
1	Adel	-3,127	-1,215	1,912	61,14%
2	Avril_Lavil	-2,585	-0,949	1,636	63,29%
3	Elvis_Prestly	-3,139	-2,080	1,059	33,74%
4	Evanescence	-2,895	-1,114	1,781	61,52%
5	Hotel_California	-3,137	-1,726	1,411	44,98%
6	Jump_in_my_car	-3,019	-1,410	1,609	53,30%
7	NewYorkCity	-2,831	-1,260	1,571	55,49%
8	Poison	-2,934	-0,973	1,961	66,84%
9	SlavesOfFear	-2,500	-0,837	1,663	66,52%
10	You_are_yong	-2,838	-1,525	1,313	46,26%
<b>Среднее по всем:</b>		<b>-2,901</b>	<b>-1,309</b>	<b>1,592</b>	<b>54,87%</b>

Несложно заметить, что, на низком битрейте, преимущество предлагаемого аудиокодека составило уже 55%, а для ряда файлов превысило 65%. Это значительно превосходит результат на высоких битрейтах и является следствием нарушений ТСАП, о которых говорилось выше.

Указанный результат представляется крайне важным, т.к. повышение перцептуального качества звуковых сигналов на низких и ультранизких битрейтах является важнейшей задачей, стоящей перед современным кодированием аудиоданных с потерями.

На рисунке 3 представлены спектрограммы исходного аудиофайла и восстановленных файлов, сжатых при помощи кодека AAC и предлагаемого аудиокодека.

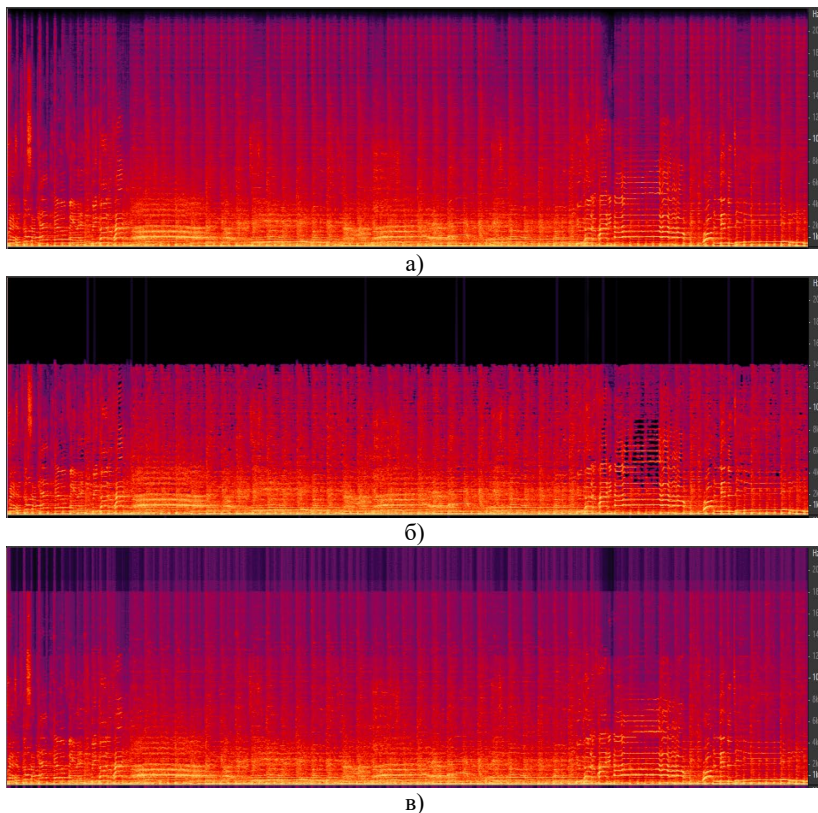


Рис. 3. Сравнение спектрограмм: а) Оригинал, б) AAC, в) Предлагаемый метод (файл №1)

Отметим, что на низких битовых скоростях искажения на спектрограмме метода AAC легко заметить невооруженным глазом – это полностью черные области, относящиеся к зануленным спектральным компонентам, чего, как показано в ТСАП, быть не должно.

Таким образом, как и было показано теоретически, после преодоления перцептуального предела сжатия, качество современных коммерческих кодеков быстро падает и на битрейте 64кбит/с

становиться уже достаточно низким. Причины подобных проблем подробно описаны в разделе 4.2.

Для битрейта 64 кбит/с субъективная оценка на слух исходного и полученных при помощи предлагаемого аудиокодека аудиосигналов, показала, что большинство экспертов не могут с уверенностью определить оригинал в слепом безреференсном тесте, а в референсном – MOS был оценен экспертами достаточно высоко, как не отличающийся от оригинала и не имеющий существенных искажений.

В свою очередь, результаты формата Advanced Audio Coding (AAC) на данном битрейте уже вызвали нарекания экспертов и, в первую очередь, вызвало недовольство сужение полосы пропускания, т.е. обрезание верхних частот, которое является важной частью алгоритма AAC для низких битрейтов, что обусловлено нехваткой имеющегося бюджета битов для кодирования всей слышимой полосы звукового сигнала.

**5. Заключение.** Доказано, что предложенный в статье Аудиокодек на основе перцептуального равенства исходного и восстановленного звукового сигнала, базирующийся на новых положениях ТСАП, позволяет достичь перцептуального равенства исходного и восстановленного аудиосигнала, что невозможно для современных методов сжатия аудиоданных с потерями. Кроме того, эти положения ТСАП позволяют создавать понятные и теоретически обоснованные методы сжатия аудиоданных с потерями, соответствующие (6) и как следствие (2). В качестве примера было проведено сжатие аудиоданных с потерями, в ходе которого Аудиокодек на основе перцептуального равенства исходного и восстановленного звукового сигнала показал объективное качество восстановленного сигнала на 25% выше (для битрейта 390 кбит/с, стерео) и 55% выше (для битрейта 64 кбит/с, стерео) по сравнению с наиболее прогрессивным на сегодняшний день форматом AAC MPEG-4.

Стоит отметить, что прогресса в объективном качестве восстановленного аудиосигнала на указанных высоких битовых скоростях не наблюдалось последние 10 лет и высокие битрейты считаются наиболее сложными для улучшения результатов сжатия аудиоданных с потерями. Как показано в статье, современные форматы AAC HE (v1/2) не могут помочь в повышении качества на высоких битрейтах, т.к. используют тот же алгоритм перцептуального сжатия, как и AAC MPEG-4.



Проведенные эксперименты подтверждают тот факт, что метод сжатия аудиоданных, построенный на основе развитой теории сжатия аудиоданных с потерями, позволяет получать более высокие, в смысле объективного качества (по стандартизированной мере PEAQ), результаты в сравнении с лидирующим на сегодняшний день на рынке форматом AAC даже на высоких битрейтах, где качество восстановленных аудиосигналов приближается к качеству оригинала.

### Литература

1. Ковалгин Ю.А., Вологдин Э.И. Цифровое кодирование звуковых сигналов // М.: КОРОНА-принт, 2015. 240 с.
2. Журавлёва Л.В., Шишурин А.И. Сравнительный анализ аудиоформатов // Технологии инженерных и информационных систем. 2022. № 2. С. 67–78.
3. Каргин Р.И., Стаценко Л.Г. Форматы сжатия аудиоданных. Анализ и сравнение // Известия СПбГЭТУ ЛЭТИ. 2019. № 9. С. 31–37.
4. Koops N.V., Micchi G., Quinton E. Robust lossy audio compression identification. 2024. arxiv preprint arxiv:2407.21545.
5. Ковалгин Ю.А., Фадеева Д.Р. Исследование психоакустических моделей кодеков с компрессией цифровых аудиоданных // Современная наука: актуальные проблемы теории и практики. Серия: Естественные и технические науки. 2016. № 7. С. 29–38.
6. Официальный сайт ITU. Method for objective measurements of perceived audio quality. Recommendation ITU-R BS.1387-2 (05/2023). URL: [https://www.itu.int/dms\\_pubrec/itu-r/rec/bs/R-REC-BS.1387-2-202305-1!!PDF-E.pdf](https://www.itu.int/dms_pubrec/itu-r/rec/bs/R-REC-BS.1387-2-202305-1!!PDF-E.pdf) (дата обращения: 05.01.2025).
7. Cormen T., Leiserson C., Rivest R., Stein C. Introduction to Algorithms 4th Edition. Cambridge, Massachusetts: The MIT Press, 2022. 1312 p.
8. Чугунова О.В., Буслова В.Е. Архивации данных методами Шеннон-Фано и Хаффмана // Актуальные проблемы науки и техники: Сборник статей по материалам международной научно-практической конференции. Уфа: Изд. НИЦ Вестник науки, 2020. С. 58–65.
9. Сергеев И.С., Балакирев Н.Е. Сравнение алгоритмов сжатия звуковой информации алгоритмом Хаффмана и арифметическим кодированием // Наукосфера. 2022. № 8-2. С. 31–35.
10. Официальная страница проекта FLAC. URL: <https://sourceforge.net/projects/flac/> (дата обращения: 05.01.2025).
11. Salomon D. Data compression: the complete reference 4th Edition // London: Springer-Verlag, 2007. 1117 с.
12. Официальный сайт Monkey's Audio (Спецификация Monkey's Audio). URL: <https://www.monkeysaudio.com/index.html> (дата обращения: 05.01.2025).
13. Официальный сайт проекта ALAC (Спецификация Apple Lossless Audio Codec). URL: <https://macosforge.github.io/alac/> (дата обращения: 05.01.2025).
14. Официальный сайт Microsoft (Windows Media Player). URL: <https://apps.microsoft.com/detail/9WZDNCRFJ3PT?hl=en-us&gl=US> (дата обращения: 05.01.2025).
15. Официальный сайт Xiph.Org фонда (Спецификация Vorbis I Xiph.Org). URL: [https://xiph.org/vorbis/doc/Vorbis\\_I\\_spec.html](https://xiph.org/vorbis/doc/Vorbis_I_spec.html) (дата обращения: 05.01.2025).
16. Официальный сайт Opus Interactive Audio Codec. URL: <https://opus-codec.org/> (дата обращения: 05.01.2025).

17. Петровский Ал.А., Петровский А.А. Масштабируемые аудиоречевые кодеры на основе адаптивного частотно-временного анализа звуковых сигналов // Труды СПИИРАН. 2017. № 1(50). С. 55–92. DOI: 10.15622/sp.50.3.
18. Официальный сайт ITU. Audio coding for digital broadcasting. Recommendation ITU-R BS.1196-8 (10/2019). URL: [https://www.itu.int/dms\\_pubrec/itu-r/rec/bs/R-REC-BS.1196-8-201910-1!PDF-E.pdf](https://www.itu.int/dms_pubrec/itu-r/rec/bs/R-REC-BS.1196-8-201910-1!PDF-E.pdf) (дата обращения: 05.01.2025).
19. Jenrungrot T., Chinen M., Kleijn W.B., Skoglund J., Borsos Z., Zeghidour N., Tagliasacchi M. LMcodec: a Low Bitrate Speech Codec With Causal Transformer Models // Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). 2023. pp. 1–5. DOI: 10.1109/ICASSP49357.2023.10095442.
20. Shi H., Shimada K., Hirano M., Shibuya T., Koyama Y., Zhong Z., Takahashi S., Kawahara T., Mitsufoji Y. Diffusion-Based Speech Enhancement with Joint Generative and Predictive Decoders // IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). 2024. pp. 12951–12955. DOI: 10.1109/ICASSP48485.2024.10448429.
21. Kong J., Kim J., Bae J. HiFi-GAN: Generative Adversarial Networks for Efficient and High Fidelity Speech Synthesis // Proceedings of the 34th Conference on Neural Information Processing Systems (NIPS). 2020. vol. 33. pp. 17022–17033.
22. Kaneko T., Tanaka K., Kameoka H., Seki S. Istftnet: Fast and Lightweight Mel-spectrogram Vocoder Incorporating Inverse Short-Time Fourier Transform. 2022. arxiv preprint arXiv:2203.02395v1.
23. Subramani K., Valin J.-M., Isik U., Smaragdis P., Krishnaswamy A. End-to-end LPCNet: A Neural Vocoder With Fully-Differentiable LPC Estimation // Proceedings of the Annual Conference of the International Speech Communication Association (INTERSPEECH). 2022. pp. 818–822.
24. Valin J.-M., Skoglund J. LPCNet: Improving Neural Speech Synthesis Through Linear Prediction // Proceedings of the International Conference on Acoustics, Speech and Signal Processing (ICASSP). 2019. pp. 5891–5895. DOI: 10.1109/ICASSP.2019.8682804.
25. Valin J.-M., Isik U., Smaragdis P., Krishnaswamy A. Neural Speech Synthesis on a Shoestring: Improving the Efficiency of LPCNet // Proceedings of the International Conference on Acoustics, Speech and Signal Processing (ICASSP). 2022. pp. 8437–8441.
26. Valin J.-M., Buthe J., Mustafa A. Low-Bitrate Redundancy Coding of Speech Using a Rate-distortion-optimized Variational Autoencoder // Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). 2023. pp. 1–5. DOI: 10.1109/ICASSP49357.2023.10096528.
27. Zeghidour N., Luebs A., Omran A., Skoglund J., Tagliasacchi M. SoundStream: An End-to-End Neural Audio Codec // Proceedings of the IEEE/ACM Transactions on Audio, Speech, and Language Processing. 2021. vol. 30. pp. 495–507.
28. Du Z., Zhang S., Hu K., Zheng S. FunCodec: A Fundamental, Reproducible and Integrable Open-Source Toolkit for Neural Speech Codec. 2023. arxiv preprint arXiv:2309.07405v1.
29. Defossez A., Copet J., Synnaeve G., Adi Y. Funcodec: High Fidelity Neural Audio Compression. 2022. arxiv preprint arXiv:2210.13438v1.
30. Демо-сайт проекта High Fidelity Neural Audio Compression (EnCodec). URL: <https://ai.honu.io/papers/encodec/samples.html> (дата обращения: 18.10.2024).
31. Yin D., Luo C., Xiong Z., Zeng W. PHASEN: A Phase-and-Harmonics-Aware Speech Enhancement Network. 2019. arxiv preprint arXiv:1911.04697v1.
32. Рогозинский Г.Г. Перцепционное сжатие звука с использованием вейвлетных пакетов // Диссертация СПбГУКиТ. 2010.

33. Zwicker E., Fastl H. Psychoacoustics: Facts and Models // Springer-Verlag, Berlin Heidelberg. 1990.
34. Официальный сайт ITU. General methods for the subjective assessment of sound quality. Recommendation ITU-R BS. 1284-2 (01/2019). URL: [https://www.itu.int/dms\\_pubrec/itu-r/rec/bs/R-REC-BS.1284-2-201901-1!!PDF-E.pdf](https://www.itu.int/dms_pubrec/itu-r/rec/bs/R-REC-BS.1284-2-201901-1!!PDF-E.pdf) (дата обращения: 05.01.2025).

**Чижов Илья Игоревич** — канд. техн. наук, ведущий инженер ключевых проектов, руководитель команды речевого и аудиокодирования, Российский исследовательский институт Huawei. Область научных интересов: сжатие аудиоданных, цифровая обработка сигналов, машинное/глубокое обучение. Число научных публикаций — 22. [argroximation18@yandex.ru](mailto:argroximation18@yandex.ru); улица Крылатская, 17/2, 121614, Москва, Россия; р.т.: +7(495)234-0686.

I. CHIZHOV

**AN AUDIOCODEC BASED ON THE PERCEPTUAL EQUALITY BETWEEN THE ORIGINAL AND RESTORED AUDIO SIGNALS**

*Chizhov I. An AudioCodec Based on the Perceptual Equality between the Original and Restored Audio Signals.*

**Abstract.** A method for lossy audio data compression (AudioCodec) is presented. It allows for improving objective quality of the restored audio signal by 25% at a bitrate of 390 kbps and 55% at a bitrate of 64 kbps compared to the AAC MPEG-4 format. The proposed method of audio data compression is based on an advanced theory of lossy audio data compression (TLAC), which is also introduced in the article. The improvement in the objective quality of the reconstructed audio signal (according to the standardized PEAQ measure) is achieved because the TLAC overcomes issues in modern lossy audio data compression methods related to the use of psychoacoustic principles of human sound perception, including after overcoming the "psychoacoustic compression limit" of the audio signal (i.e. the moment in perceptual coding when the available bit budget is insufficient to encode all spectral components with the accuracy required from a psychoacoustic perspective). This allows for achieving perceptual equality between the original and reconstructed audio signals. As an analysis of the state of the art, solutions for both lossless and lossy audio data compression, as well as those using artificial intelligence, are considered. In all modern lossy audio data compression methods, the procedure for selecting the spectral components to be preserved, as well as the permissible quantization error, is carried out through a series of highly complex procedures collectively referred to as the "psychoacoustic model of the lossy audio compression method". In a strict sense, perceptual equality between the spectra of the original and restored signals has not been proven by any research group and, therefore, cannot be guaranteed by them. Independent experts regularly publish tests demonstrating that modern audio codecs have issues with certain audio signals. The article proposes an AudioCodec based on the perceptual equality between the original and restored audio signals, which is based on the new ideas of the theory of lossy audio compression (TLAC). These ideas guarantee the achievement of perceptual equality between the original and restored audio signals at different bitrates, therefore, the AudioCodec built on its basis is free from the above-mentioned issues and, as a result, significantly outperforms modern AudioCodecs in terms of the objective quality of the restored audio signal, as measured by PEAQ.

**Keywords:** audio data compression, psychoacoustic model, spectrum, subband division, perceptual equality of the spectra.

**References**

1. Kovalgin Yu.A., Vologdin E.I. *Cifrovoe kodirovanie zvukovyh signalov* [Digital coding of sound signals]. M.: KORONA-print, 2015. 240 p. (In Russ.).
2. Zhuravleva L.V., Shishurin A.I. [Comparative analysis of audio formats]. *Tehnologii inzhenernyh i informatsionnyh sistem – Technologies of engineering and information systems*. 2022. no. 2. pp. 67–78. (In Russ.).
3. Kargin R.I., Statsenko L.G. [Audio data compression formats. Analysis and comparison]. *Izvestija SPbGJeTU LJeTI – Bulletin of ETU LETI*. 2019. no. 9. pp. 31–37.
4. Koops H.V., Micchi G., Quinton E. Robust lossy audio compression identification. 2024. arxiv preprint arxiv:2407.21545.

5. Kovalgin Yu.A., Fadeeva D.R. [Study of psychoacoustic models of codecs with digital audio data compression]. *Sovremennaja nauka: aktual'nye problemy teorii i praktiki*. Serija: Estestvennye i tehniceskije nauki – Modern science: current problems of theory and practice. Series: Natural and Technical Sciences. 2016. no. 7. pp. 29–38. (In Russ.).
6. Official website of the ITU. Method for objective measurements of perceived audio quality. Recommendation ITU-R BS.1387-2 (05/2023). Available at: [https://www.itu.int/dms\\_pubrec/itu-r/rec/bs/R-REC-BS.1387-2-202305-!!PDF-E.pdf](https://www.itu.int/dms_pubrec/itu-r/rec/bs/R-REC-BS.1387-2-202305-!!PDF-E.pdf). (accessed 05.01.2025).
7. Cormen T., Leiserson C., Rivest R., Stein C. *Introduction to Algorithms* 4th Edition. Cambridge, Massachusetts: The MIT Press, 2022. 1312 p.
8. Chugunova O.V., Buslova V.E. [Data archiving using Shannon-Fano and Huffman methods] *Aktual'nye problemy nauki i tehniki: Sbornik statej po materialam mezhdunarodnoj nauchno-prakticheskoj konferencii* [Actual problems of science and technology: Collection of articles based on the materials of the international scientific and practical conference]. Ufa: Izd. NIC Vestnik nauki, 2020. pp. 58–65. (In Russ.).
9. Sergeev I.S., Balakirev N.E. [Comparison of audio compression algorithms using the Huffman algorithm and arithmetic coding]. *Naukosfera – Naukosphere*. 2022. no. 8-2. pp. 31–35. (In Russ.).
10. Official page of the FLAC project. Available at: <https://sourceforge.net/projects/flac/>. (accessed 05.01.2025).
11. Salomon D. *Data compression: the complete reference* 4th Edition. London: Springer-Verlag, 2007. 1117 c.
12. Official website of Monkey's Audio (Monkey's Audio Specification). Available at: <https://www.monkeysaudio.com/index.html>. (accessed 05.01.2025).
13. Official website of the ALAC project (Apple Lossless Audio Codec Specification). Available at: <https://macosforge.github.io/alac/>. (accessed 05.01.2025).
14. Official website of Microsoft (Windows Media Player). Available at: <https://apps.microsoft.com/detail/9WZDNCRFJ3PT?hl=en-us&gl=US>. (accessed 05.01.2025).
15. Official website of the Xiph.Org Foundation (Vorbis I Specification Xiph.Org). Available at: [https://xiph.org/vorbis/doc/Vorbis\\_I\\_spec.html](https://xiph.org/vorbis/doc/Vorbis_I_spec.html). (accessed 05.01.2025).
16. Official website of Opus Interactive Audio Codec. Available at: <https://opus-codec.org/>. (accessed 05.01.2025).
17. Petrovsky A.A., Petrovsky A.A. [A scalable speech and audio coders based on adaptive time-frequency signal analysis. *Trudy SPIIRAN – SPIIRAS Proceedings*. 2017. no. 1(50). pp. 55–92. DOI: 10.15622/sp.50.3. (In Russ.).
18. Official website of ITU. Audio coding for digital broadcasting. Recommendation ITU-R BS.1196-8 (10/2019). Available at: [https://www.itu.int/dms\\_pubrec/itu-r/rec/bs/R-REC-BS.1196-8-201910-!!PDF-E.pdf](https://www.itu.int/dms_pubrec/itu-r/rec/bs/R-REC-BS.1196-8-201910-!!PDF-E.pdf). (accessed 05.01.2025).
19. Jenrungrot T., Chinen M., Kleijn W.B., Skoglund J., Borsos Z., Zeghidour N., Tagliasacchi M. LMcodec: a Low Bitrate Speech Codec With Causal Transformer Models. *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. 2023. pp. 1–5. DOI: 10.1109/ICASSP49357.2023.10095442.
20. Shi H., Shimada K., Hirano M., Shibuya T., Koyama Y., Zhong Z., Takahashi S., Kawahara T., Mitsuji Y. Diffusion-Based Speech Enhancement with Joint Generative and Predictive Decoders. *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. 2024. pp. 12951–12955. DOI: 10.1109/ICASSP48485.2024.10448429.

21. Kong J., Kim J., Bae J. HiFi-GAN: Generative Adversarial Networks for Efficient and High Fidelity Speech Synthesis. Proceedings of the 34th Conference on Neural Information Processing Systems (NIPS). 2020. vol. 33. pp. 17022–17033.
22. Kaneko T, Tanaka K., Kameoka H., Seki S. Istfnet: Fast and Lightweight Mel-spectrogram Vocoder Incorporating Inverse Short-Time Fourier Transform. 2022. arxiv preprint arXiv:2203.02395v1.
23. Subramani K., Valin J.-M., Isik U., Smaragdis P., Krishnaswamy A. End-to-end LPCNet: A Neural Vocoder With Fully-Differentiable LPC Estimation. Proceedings of the Annual Conference of the International Speech Communication Association (INTERSPEECH). 2022. pp. 818–822.
24. Valin J.-M., Skoglund J. LPCNet: Improving Neural Speech Synthesis Through Linear Prediction. Proceedings of the International Conference on Acoustics, Speech and Signal Processing (ICASSP). 2019. pp. 5891–5895. DOI: 10.1109/ICASSP.2019.8682804.
25. Valin J.-M., Isik U., Smaragdis P., Krishnaswamy A. Neural Speech Synthesis on a Shoestring: Improving the Efficiency of LPCNet. Proceedings of the International Conference on Acoustics, Speech and Signal Processing (ICASSP). 2022. pp. 8437–8441.
26. Valin J.-M., Buthe J., Mustafa A. Low-Bitrate Redundancy Coding of Speech Using a Rate-distortion-optimized Variational Autoencoder. Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). 2023. pp. 1–5. DOI: 10.1109/ICASSP49357.2023.10096528.
27. Zeghidour N., Luebs A., Omran A., Skoglund J., Tagliasacchi M. SoundStream: An End-to-End Neural Audio Codec. Proceedings of the IEEE/ACM Transactions on Audio, Speech, and Language Processing. 2021. vol. 30. pp. 495–507.
28. Du Z., Zhang S., Hu K., Zheng S. FunCodec: A Fundamental, Reproducible and Integrable Open-Source Toolkit for Neural Speech Codec. 2023. arxiv preprint arXiv:2309.07405v1.
29. Defossez A., Copet J., Synnaeve G., Adi Y. Funcodec: High Fidelity Neural Audio Compression. 2022. arxiv preprint arXiv:2210.13438v1.
30. Demo site of the High Fidelity Neural Audio Compression (EnCodec). Available at: <https://ai.honu.io/papers/encodec/samples.html>. (accessed 18.10.2024).
31. Yin D., Luo C., Xiong Z., Zeng W. PHASEN: A Phase-and-Harmonics-Aware Speech Enhancement Network. 2019. arxiv preprint arXiv:1911.04697v1.
32. Rogozinsky G.G. [Perceptual sound compression using wavelet packets]. Dissertation of St. Petersburg State University of Culture and Technology. 2010. (In Russ.).
33. Zwicker E., Fastl H. Psychoacoustics: Facts and Models. Springer-Verlag, Berlin Heidelberg. 1990.
34. Official website of ITU. General methods for the subjective assessment of sound quality. Recommendation ITU-R BS. 1284-2 (01/2019). Available at: [https://www.itu.int/dms\\_pubrec/itu-r/rec/bs/R-REC-BS.1284-2-201901-!#!PDF-E.pdf](https://www.itu.int/dms_pubrec/itu-r/rec/bs/R-REC-BS.1284-2-201901-!#!PDF-E.pdf). (accessed 05.01.2025).

**Chizhov Ilya** — Ph.D., Principal engineer, team lead of AI speech/audio coding team, Huawei Russian Research Institute. Research interests: audio compression, digital signal processing, ML/DL, AI-based speech enhancement, personalized noise reduction, text-to-speech. The number of publications — 22. [aproximation18@yandex.ru](mailto:aproximation18@yandex.ru); 17/2, Krylatskaya St., 121614, Moscow, Russia; office phone: +7(495)234-0686.

В.И. ЗИНОВ, В.М. КАРТАК, Ю.И. ВАЛИАХМЕТОВА  
**РЕШЕНИЕ МНОГОКРИТЕРИАЛЬНОЙ ЗАДАЧИ  
РАЦИОНАЛЬНОГО РАЗМЕЩЕНИЯ НЕСУЩИХ СТЕН  
С ПОМОЩЬЮ ГЕНЕТИЧЕСКОГО АЛГОРИТМА**

*Зинов В.И., Картак В.М., Валиахметова Ю.И. Решение многокритериальной задачи рационального размещения несущих стен с помощью генетического алгоритма.*

**Аннотация.** Рациональное размещение несущих стен остается сложной и малоизученной задачей, несмотря на большое количество алгоритмов и моделей решения схожей задачи размещения колонн. Основными факторами сложности являются большое количество вариантов решения, большое время, требуемое для оценки полученного плана размещения на деформации, а также – многокритериальный характер. Кроме нелинейного критерия оценки деформаций требуется минимизировать протяженность несущих стен и количество различных уникальных типоразмеров. В статье предложена модель задачи размещения несущих стен, разбивающая стены на функциональные участки с некоторым шагом и учитывающая перечисленные критерии. При этом соседствующие участки стен одной функциональности объединяются в сегменты. Комбинаторный вид разбиения стен в модели задачи позволяет применить генетические алгоритмы для её решения. Поэтому была предложена новая схема многокритериального генетического алгоритма, содержащая метрики расчёта разнообразия фенотипа и генотипа популяции. Представлены модификации операторов скрещивания, мутации и селекции, учитывающие сегментальный вид генотипа стен. Анализ сравнения разработанного алгоритма с другими многокритериальными генетическими алгоритмами показал, что, несмотря на в 2 раза более длительное время выполнения, разработанный алгоритм находит в среднем в 3 раза больше недоминируемых решений на общем множестве, особенно с меньшим значением оценки деформаций. Предложенная модель разительно отличается от предлагаемых ранее с точки зрения оперирования с деформациями в системах опоры-плита, лишь сравнивая между собой планы размещения, а не рассчитывая точные оценки армирования, что зачастую бывает излишне на ранних этапах. Предложенная схема генетического алгоритма увеличивает количество найденных недоминируемых решений без потери их разнообразия, жертвуя временем выполнения, и может быть использована для решения других многокритериальных задач с учётом указанных особенностей. Разработанный алгоритм был легко интегрирован в программное средство поддержки принятия решений на базе САПР и может быть использован на практике специалистами.

**Ключевые слова:** оптимизация в проектировании зданий, задача размещения стен, многокритериальный генетический алгоритм, комбинаторная оптимизация, сравнительная оценка деформаций плит.

**1. Введение.** Расположение несущих стен – одна из важнейших характеристик проекта здания. Несущие стены воспринимают нагрузки от плит перекрытий, распределяя их между собой. Плиты перекрытий прогибаются, что определяет их требуемое арматурное усиление. Расчет деформаций производится на программно-расчётных комплексах (ПРК). В силу необходимости учёта множества условий и факторов такие расчеты сложны и выполняются долго.

Сложность подбора расположения стен складывается из ряда причин. Во-первых, из сложности расчета явной функциональной связи между планом стен и деформациями, что в том числе приводит к интуитивному выбору расположения стен на основе опыта. Во-вторых, из-за большого числа вариантов планов стен и отсутствия автоматизированной системы их анализа, что не позволяет в полной мере оценить планы по различным критериям и выбрать наилучший вариант среди всего множества. В-третьих, из-за больших временных затрат на этапе расчета деформаций в ПРК, что исключает возможность автоматизированного просчёта множества вариантов. Таким образом, задача рационального подбора плана несущих стен является сложной задачей с большим числом вариантов и, по крайней мере, одним нелинейным целевым критерием – оценкой деформаций плит. Был проведен анализ научных работ, посвященных проблемам оптимизации расположения структурных конструкций здания.

В работе [1] предлагается модель размещения колонн и балок на основе приоритетного подхода. Для оптимизации их расположения по критерию стоимости применены генетические алгоритмы и алгоритмы роя частиц. Оценка прочности конструкций производится по модели балок Эйлера. Предложенная модель учитывает множество ограничений и позволяет получить допустимые планы колонн, однако подходит только для малоэтажных зданий и приспособлена для размещения колонн, но не стен. Похожая модель на основе генетических алгоритмов применяется в работе [2].

Один из исследователей проблемы оптимизации размещения конструкций – Пежман Шарафи (Pezhman Sharafi). В его докторской работе [3] предложена методология оптимизации конструкций, в т.ч. балок, колонн и плит, на основе популяционных алгоритмов. В другой работе [4] представлен эвристический алгоритм поиска оптимального регулярного размещения колонн. В качестве целевой функции принята стоимость армирования плиты, которая зависит от пролетов между колоннами и определяется по расчёту изгибающих моментов методом заменяющих рам. Недостатком подхода является ограничение метода заменяющих рам, который подходит только для регулярных схем опор, что ограничивает область применения. В статье [5] разработана модель поиска оптимального размещения, размеров и армирования колонн. Целевой критерий – стоимость конструкций, а ограничения – условие рассчитываемой прочности структур на изгиб. Тем не менее, область применения также ограничивается регулярными планами колонн.



В работе [6] предложен метод оптимизации колонн с помощью множителей Лагранжа. Целевой критерий – жесткость конструкций, рассчитываемая по модулю Юнга. Позиции для колонн заданы дискретно, неизвестное – количество колонн. В работе [7] предложена модель оптимизации размещения колонн на основе расчёта градиента объёма бетона в плите перекрытия над колоннами при ограничении прогиба плиты, напряжения и изгибающего момента, что повышает вычислительную эффективность, однако сильно усложняет модель.

В работе [8] колонны и балки представлены в виде планарного графа. Для расчёта нагрузок на конструкции предложен эвристический метод, распределяющий вес плиты на колонны и балки. Реализованная программная система позволяет сравнивать автоматизированный план и план, разработанный специалистами. В статье [9] рассмотрена задача оптимизации размещения несущих стен. Для анализа геометрии плиты решается задача о разбиении многосвязного ортогонального полигона. Толщина стен константна. Для решения предлагается эвристический алгоритм поиска рационального плана размещения стен с правилами, ограничивающими поиск в пределах области допустимости. Однако модель не содержит практико-ориентированного инструмента анализа деформаций и не приспособлена для многокритериального решения.

В работе [10] описывается интегрированная в BIM-процесс проектирования процедура задания зон дополнительного армирования в перекрытии на основе расчётных значений моментов напряжений, полученных с помощью метода конечных элементов. Несмотря на простоту практического применения, ключевой недостаток подхода – отсутствие достаточной гибкости при изменении исходных данных: изменив плиту перекрытия или её опоры, весь процесс расчётов придется повторить заново. Подобный подход оптимизации каркасов с применением средств глубокого обучения применен в [11]. Кроме общепринятых методов оценки прочности, например, упомянутым методом заменяющих рам, разрабатываются также оценки на основе нейросетевого подхода [12]. ИНС также могут применяться в качестве инструмента генерации всего плана конструкций [13].

В большинстве работ решается задача оптимизации размещения колонн, но не стен. Причина этого – в особенностях проектирования: колонны зачастую располагают по сетке с регулярным шагом, что упрощает задачу, но ограничивает область применения предлагаемых методов. Кроме того, чаще всего для оценки деформаций в плитах применяются рассчитываемые значения, что усложняет структуру модели и увеличивает вычислительную сложность. Однако на этапе планирования опорных конструкций точные оценки не обязательны;

подход оценки, направленный на сравнение планов по деформациям между собой, может быть более быстрым и рациональным. Поэтому цель данной работы состоит в разработке модели и алгоритма решения задачи рационального размещения несущих стен с применением сравнительного подхода оценки деформаций в плитах по планам стен.

В данной статье описываются научные результаты реализации этой цели. В главе 2 представлена новая математическая модель задачи рационального размещения несущих стен в комбинаторной многокритериальной постановке. Критерии задачи – протяженность несущих стен, количество различных длин несущих стен и оценка деформаций, которая выражается через разработанную ранее модель [14]. В главе 3 описана схема многокритериального генетического алгоритма (МГА), решающая задачу рационального размещения несущих стен на основе предлагаемой модели. Т.к. стены в модели представлены в комбинаторном виде с однородными сегментами, то в предлагаемом МГА гены также рассматриваются сегментно. Поэтому предлагаются новые алгоритмы мутации и скрещивания, которые оперируют не с отдельными генами, а со связными цепочками генов. Для того чтобы учесть особенности и условия задачи, был разработан новый алгоритм оператора селекции. В главе 4 проведено сравнение эффективности решения задачи рационального размещения несущих стен разработанным МГА и аналогами на основе предложенной модели. Разработанные модель и алгоритмы реализуются в виде автоматизированной системы поддержки принятия решений.

**2. Модель задачи рационального размещения стен.** Дано множество несущих стен  $T = \{T_1, \dots, T_N\}$ , где  $N \in \mathbb{Z}$  – количество стен. Каждая  $i$ -ая стена характеризуется как  $T_i = \langle (x_i, y_i), w_i, l_i, n_i \rangle$ , где  $x_i \in \mathbb{Z}$ ,  $y_i \in \mathbb{Z}$  – координаты стены;  $w_i \in \mathbb{Z}$  – толщина стены;  $l_i \in \mathbb{Z}$  – длина стены;  $n_i \in \mathbb{Z}$  – количество участков разбиения в стене;  $i = 1, \dots, N$ . Участки разбиения в стене определены как  $t_{ij} = \langle l_{ij} \rangle$ , где  $l_{ij} \in \mathbb{Z}$  – длина  $j$ -го участка;  $j = 1, \dots, n_i$ ,  $i = 1, \dots, N$ .

Введём переменные  $X_{ij} \in \{0, 1\}$ , отражающие функциональное свойство (ФС)  $j$ -го участка  $i$ -ой стены: если  $X_{ij} = 0$ , то участок является ненесущим; если  $X_{ij} = 1$ , то участок является несущим.

Соседствующие участки стены с одинаковым ФС объединяются в сегменты. Планом опор назовём множество планов сегментов  $Q = \{Q_1, \dots, Q_N\}$ . Пусть для каждой стены задано количество сегментов  $m_i \in \mathbb{Z}$ ,  $1 \leq m_i \leq n_i$ . Тогда каждый  $i$ -ый план сегментов задан как множество сегментов  $Q_i = \{Q_{ik}: k = 1, \dots, m_i\}$ . Пусть для каждого

сегмента задано количество участков  $\eta_{ik} \in \mathbb{Z}$ ,  $1 \leq \eta_{ik} \leq n_i$ . Тогда каждый  $k$ -ый сегмент  $i$ -ой стены задан множеством индексов его участков:

$$\forall k = \overline{1, \dots, m_i} : Q_{ik} = \{j : j = \sum_{k^*=1}^{k-1} \eta_{ik^*} + 1, \dots, \sum_{k^*=1}^{k-1} \eta_{ik^*} + \eta_{ik}\}.$$

Зададим начальное ФС в  $i$ -ой стене как  $\chi_i \in \{0, 1\}$ , а ФС отдельного сегмента –  $\chi_{ik} \in \{0, 1\}$ . Принимается, что сегменты чередуются по ФС. Тогда ФС любого сегмента может быть определено как  $\chi_{ik} = (\chi_i + k + 1) \bmod 2$ ,  $\forall k = 1, \dots, m_i$ . Обозначим значение ФС  $j$ -ого участка  $i$ -ой стены по плану опор  $Q$  как:

$$\forall j = \overline{1, \dots, n_i} : X_{ij}(Q) = \chi_{ik} \Leftrightarrow j \in Q_{ik}. \quad (1)$$

Длина  $k$ -ого сегмента в  $i$ -ой стене определяется как:

$$length(Q_{ik}) = \sum_{j \in Q_{ik}} l_{ij}.$$

Длины всех сегментов, выстроенных по плану опор  $Q$ , должны быть не меньше минимумов длин по соответствующему ФС:

$$\forall i = 1, \dots, N, \forall k = 1, \dots, m_i : length(Q_{ik}) \geq \begin{cases} l^0, & \text{если } \chi_{ik} = 0 \\ l^1, & \text{если } \chi_{ik} = 1 \end{cases}, \quad (2)$$

где  $l^0$  и  $l^1$  – минимальные длины несущих и не несущих сегментов.

Требуется, чтобы значения трёх целевых функций по построенному плану опор  $Q$  стремились к минимуму:

$$f_L(Q) = \frac{\sum_{i=1}^N \sum_{k=1}^{m_i} (length(Q_{ik}) \cdot \chi_{ik})}{\sum_{i=1}^N l_i} \rightarrow \min, \quad (3)$$

где  $f_L(Q)$  – функция протяженности несущих сегментов, нормированная относительно общей длины стен;

$$f_U(Q) = \frac{|\{length(Q_{ik}) : \chi_{ik} = 1; k = \overline{1, \dots, m_i}, i = \overline{1, \dots, N}\}|}{|\{Q_{ik} : \chi_{ik} = 1; k = \overline{1, \dots, m_i}, i = \overline{1, \dots, N}\}|} \rightarrow \min, \quad (4)$$

где  $f_U(Q)$  – функция количества типов длин несущих сегментов, нормированная относительно общего количества несущих сегментов;

$$f_{Sp}(Q) = V(\overline{Sp}) = \frac{\sum_{Sp \in \overline{Sp}} V(Sp)}{|\overline{Sp}|} \rightarrow \min, \quad (5)$$

где  $f_{Sp}(Q)$  – функция оценки деформаций плит по плану опор;  $V(\overline{Sp}) \in [0, 1]$  – оценка деформации плиты, описанная в работе [14];  $\overline{Sp} = \{Sp_1, \dots, Sp_u\}$  – множество пролетов,  $u \in \mathbb{Z}$  – количество пролетов;  $Sp = \langle (x_{Sp}, y_{Sp}), (w_{Sp}, l_{Sp}) \rangle$  – пролет в плите,  $(x_{Sp}, y_{Sp})$  – координаты пролета,  $(w_{Sp}, l_{Sp})$  – размеры пролета;  $V(Sp) \in [0, 1]$  – оценка деформаций плиты в пролете  $Sp \in \overline{Sp}$ . Оценка деформации плиты в пролете задана как:

$$V(Sp) = \hat{V}(Sp) + Penalty(Sp),$$

где  $\hat{V}(Sp) \in [0, 0.5]$  – основная оценка пролета,  $Penalty(Sp) \in [0, 0.5]$  – штрафное слагаемое пролета. Штрафное слагаемое определяется как:

$$Penalty(Sp) = \begin{cases} 0.5, & \text{если } size(Sp) > \max_{size} \\ 0, & \text{иначе} \end{cases},$$

где  $size(Sp)$  – размер пролета,  $\max_{size}$  – допустимый размер пролета. Основная оценка пролета  $\hat{V}(Sp)$  определена следующим образом:

$$\hat{V}(Sp) = \frac{(1 + |\phi - 13/18|)^{-(Pr+2/3)^2} \cdot S(Sp)^{1+Pr^2} \cdot (\log_4 S(FSI))^2}{2 \cdot (\log_4 S(Sp))^{1+Pr^2} \cdot S(FSI)^2},$$

где  $\phi \in (0, 1]$  – соотношение сторон пролета,  $S(Sp) \in \mathbb{Z}$  – площадь пролета  $Sp$ ,  $S(FSI) \in \mathbb{Z}$  – площадь плиты,  $FSI$  – прямоугольная плита,  $Pr$  – эвристический параметр расположения опор (максимальное расстояние между двумя опорами вдоль периметра MPD [14]).

Определим переход между моделью рационального размещения стен (план опор  $Q$ ) и моделью определения пролётов перекрытия [14] (препятствия  $Ob$ ). Пусть позиция сегмента  $k$  в стене  $i$  задана как:

$$pos(Q_{ik}) = (x_i, y_i) + dir(i) \cdot \sum_{k^*=1}^{k-1} (length(Q_{ik^*})),$$

где  $dir(i) \in \{(0, 1), (1, 0)\}$  – направление стены  $i$ . Определим множество препятствий по стенам как  $Ob = \{Ob_i\}$ , где множество препятствий по  $i$ -ой стене определяется как:

$$Ob_i = \{ \langle pos(Q_{ik}), (length(Q_{ik}), w_i) \rangle \mid \chi_{ik} = 1, k = \overline{1, \dots, m_i} \}.$$

Тогда модели могут быть использованы совместно: результаты модели детерминации пролётов используются в целевом критерии оценки деформации в модели рационального размещения стен (рисунок 1):

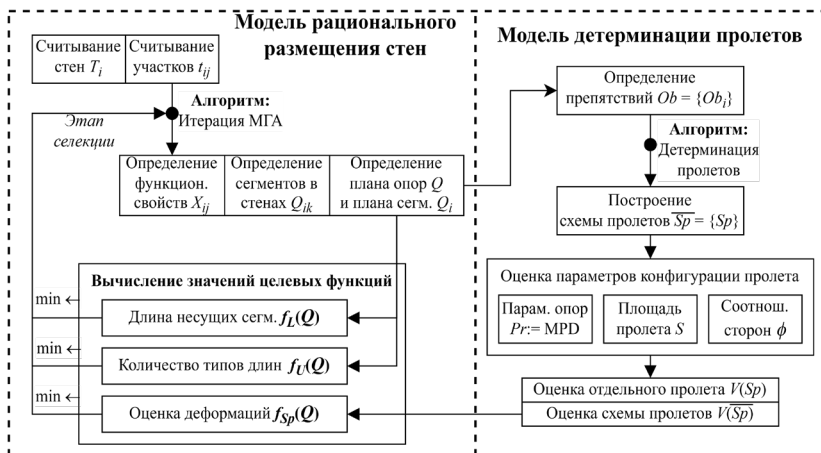


Рис. 1. Схема связи моделей детерминации пролётов и размещения стен

Описанная математическая модель позволяет формализовать параметры задачи и, в дальнейшем, разработать алгоритмы решения. Таким образом, представленная математическая модель объединяет подзадачи проблемы рационального размещения несущих стен и системно структурирует задачу. Далее представлены алгоритмы многокритериального анализа, основанные на предлагаемой модели.

### 3. Многокритериальный генетический алгоритм.

Многокритериальный генетический алгоритм (МГА) – это способ многокритериальной оптимизации, применяющий для поиска решения

методы случайного подбора, изменения и комбинирования вариантов решений. МГА применяются в задачах поддержки принятия решений.

Пусть на итерации МГА была получена популяция  $\Lambda = \{Q^1, \dots, Q^\Omega\}$ , где  $\Omega$  – мощность популяции,  $Q^\zeta$  – план опор, полученный  $\zeta$ -ой особью,  $\zeta = 1, \dots, \Omega$ . Для решения задачи рационального размещения несущих стен предлагается новая схема МГА (рисунок 2). Схема содержит три метрики для оценки особей популяции: метрику различности  $M_{div}$  – оценивает расстояние между особями в пространстве критериев; метрику схожести  $M_{sim}$  – оценивает схожесть между генотипами особей по расстоянию Хэмминга; и метрику элитарности  $M_{eli}$  – оценивает степень доминирования по Парето между особями. Также используются ранг по разнообразию  $R_{div}$  для мутации и скрещивания и дискриминированный ранг по элитарности  $R_{dis}$  для селекции. Далее перечисленные метрики и ранги рассматриваются подробнее.

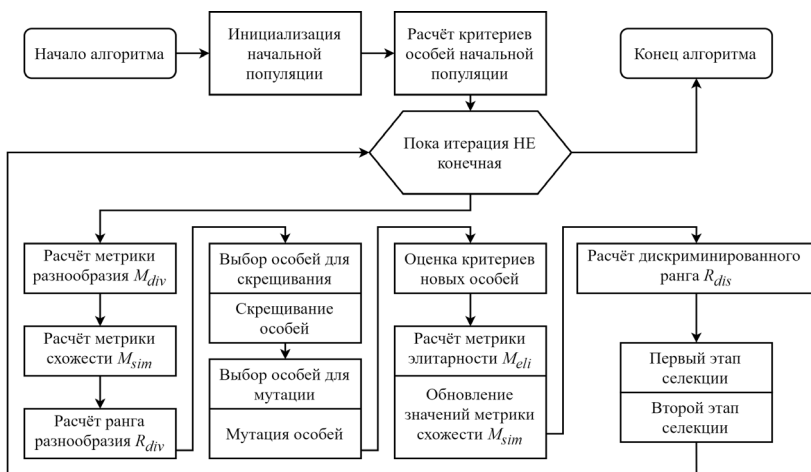


Рис. 2. Схема разработанного многокритериального генетического алгоритма

**3.1. Метрики оценки и ранжирования популяции.** Расчет метрики различности  $M_{div} \in \mathbb{R}$  оценивает расстояние между особями в пространстве критериев. Наиболее подходящий алгоритм расчёта метрики различности описан в работе [15]. Для особи, находящейся на наибольшем суммарном отдалении от других особей, оценка будет  $\zeta$  наименьшей. Метрика различности  $M_{div}$  для особи  $Q^\zeta \in \Lambda$  задана как:

$$\forall Q^\zeta \in \Lambda : d(Q^\zeta, Q^\theta) = \sqrt{\sum_{f \in \{f_L, f_U, f_{Sp}\}} (f(Q^\zeta) - f(Q^\theta))^2};$$

$$\beta(Q^\zeta, Q^\theta) = 1 - \frac{d(Q^\zeta, Q^\theta)}{\min(\{d(Q^\zeta, Q^{\theta*}) : \forall Q^{\theta*} \in \Lambda \setminus \{Q^\zeta\}\})};$$

$$M_{div}(Q^\zeta) = \sum_{\forall Q^\theta \in \Lambda \setminus \{Q^\zeta\}} \beta(Q^\zeta, Q^\theta). \quad (6)$$

**Метрика схожести**  $M_{sim} \in (0, 1]$  применена в работе [16] и оценивает количество схожих по генотипу особей. Для каждой особи рассчитывается расстояние Хэмминга до генотипов других особей. Тогда если расстояние до одной из особей меньше порога, то дискриминация ранга по метрике схожести увеличивается. Обозначим расстояние Хэмминга между генотипами особей  $Q^\zeta \in \Lambda$  и  $Q^\theta \in \Lambda$  в соответствии с формулой индекса сегмента для участка (1) как:

$$d_H(Q^\zeta, Q^\theta) = \sum_{i=1}^N \sum_{j=1}^{n_i} |X_{ij}(Q^\zeta) - X_{ij}(Q^\theta)|.$$

Пусть  $sr \in (0, 1]$  – параметр дискриминации особи за схожесть,  $cr \in \mathbb{Z}$  – параметр максимальной различности двух схожих планов. Тогда определим метрику схожести  $M_{sim}$  для некоторой особи  $Q^\zeta$  как:

$$M_{sim}(Q^\zeta) = \prod_{Q^\theta \in \Lambda \setminus \{Q^\zeta\}} \begin{cases} 1, & \text{если } d_H(Q^\zeta, Q^\theta) > cr \\ sr + (1 - sr) \cdot d_H(Q^\zeta, Q^\theta) / cr, & \text{иначе} \end{cases}. \quad (7)$$

**Метрика элитарности**  $M_{eli} \in \mathbb{Z}$  оценивает особей по степени Парето-доминируемости: если особь не доминируема в популяции, то ей присваивается ранг 1; если особь доминируема только особями с рангом 1, то ей присваивается ранг 2, и т.д. Для этого вычисляются показатели доминирования между парами особей ( $Q^\zeta$  и  $Q^\theta \in \Lambda$ ):

$$\alpha(Q^\zeta, Q^\theta) = \begin{cases} 1, & \text{если } f(Q^\zeta) \leq f(Q^\theta) \\ -1, & \text{если } f(Q^\zeta) \geq f(Q^\theta), \forall f \in \{f_L, f_U, f_{Sp}\}. \\ 0, & \text{иначе} \end{cases}$$

Алгоритм расчёта метрики элитарности  $M_{eli}$  для особей популяции  $\Lambda$  представлен на листинге 1.

1.  $v := 1$  – ранг;  $\omega := 1$  – индекс подмножества;  $\Lambda^* := \Lambda$  – временный список;
2. **while**  $\Lambda^* \neq \emptyset$  **do**
3.      $\Lambda^\omega := \emptyset$ ;
4.     **for**  $\zeta \in (1 : |\Lambda^*|)$  **do**
5.         **if**  $\exists Q^\theta \in \Lambda^* : a(Q^\zeta, Q^\theta) = -1$  **then**
6.              $M_{eli}(Q^\zeta) := v$ ;
7.              $\Lambda^\omega.add(Q^\zeta)$ ;
8.         **end if**;
9.     **end for**;
10.      $\Lambda^* := \Lambda^* \setminus \Lambda^\omega$ ;
11.      $v := v + 1$ ;  $\omega := \omega + 1$ ;
12. **end while**;

Листинг 1. Псевдокод алгоритма метрики элитарности.

Таким образом, указанный алгоритм позволяет получить ранг по метрике элитарности для всех особей популяции.

Цель *ранжирования по разнообразию*  $R_{div} \in \mathbb{Z}$  состоит в выборе для мутации и скрещивания особей с наименьшей дискриминацией по схожести и наибольшим значением метрики различности. В таком случае популяция будет наполняться наиболее различными особями. Алгоритм оценки ранга разнообразия  $R_{div}$  для особей популяции  $\Lambda$  представлен на листинге 2.

1.  $\rho := 1$  – ранг;  $\tau := 1$  – индекс по схожести (7);  $\Lambda^* := \Lambda$  – временный список;
2. **while**  $\Lambda^* \neq \emptyset$  **do**
3.      $\Lambda^\tau := \emptyset$ ;
4.     **for**  $\zeta \in (1 : |\Lambda^*|)$  **do**
5.         **if**  $\exists Q^\theta \in \Lambda^* : M_{sim}(Q^\zeta) > M_{sim}(Q^\theta)$  **then**  $\Lambda^\tau.add(Q^\zeta)$ ;
6.     **end for**;
7.      $\Lambda^{\tau^*} := \Lambda^\tau$ ;
8.     **while**  $\Lambda^{\tau^*} \neq \emptyset$  **do**
9.          $\mu := 1$  – индекс по различности (6);  $\Lambda^{\tau,\mu} := \emptyset$ ;
10.         **for**  $\zeta \in (1 : |\Lambda^{\tau^*}|)$  **do**
11.             **if**  $\exists Q^\theta \in \Lambda^{\tau^*} : M_{div}(Q^\zeta) < M_{div}(Q^\theta)$  **then**
12.                  $R_{div}(Q^\zeta) := \rho$ ;
13.                  $\Lambda^{\tau,\mu}.add(Q^\zeta)$ ;
14.             **end if**;
15.         **end for**;
16.          $\Lambda^{\tau^*} := \Lambda^{\tau^*} \setminus \Lambda^{\tau,\mu}$ ;
17.          $\rho := \rho + 1$ ;  $\mu := \mu + 1$ ;



18. **end while**;
19.  $\Lambda^* := \Lambda^* \setminus \Lambda^\tau$ ;
20.  $\tau := \tau + 1$ ;
21. **end while**;

Листинг 2. Псевдокод алгоритма ранжирования по разнообразию

Таким образом, при выборе особей для скрещивания и мутации по рангам  $R_{div}$  популяция будет наполняться более разнообразными потомками, в генотипическом и фенотипическом смыслах.

**Ранжирование по элитарности с дискриминацией по схожести**  $R_{dis} \in \mathbb{R}$  использован в операторе селекции для улучшения разнообразия в формируемой популяции. Чем меньше ранг по метрике элитарности  $M_{eli}$  и чем меньше в популяции схожих особей по метрике  $M_{sim}$  (7), тем выше шанс особи быть выбранной в новую популяцию. Дискриминированный ранг элитарности для особи  $Q^c \in \Lambda$  задан как:

$$R_{dis}(Q^c) = \frac{M_{eli}(Q^c)}{M_{sim}(Q^c)^2}. \quad (8)$$

Таким образом, дискриминированный ранг будет равен рангу по элитарности, если в популяции отсутствуют схожие по генотипу особи, и будет больше в противном случае.

### 3.2. Скрещивание – алгоритм сегментального обобщения.

Оператор скрещивания обеспечивает итеративное улучшение решения, создавая новые особи комбинированием особей-предков. Скрещивание применяется для особей с наилучшими показателями разнообразия  $R_{div}$ . Для обработки сегментов стен требуется алгоритм скрещивания не на уровне отдельных генов, а на уровне генных цепочек. Такой алгоритм назовем алгоритмом сегментального обобщения.

Алгоритм сегментального обобщения состоит в построении планов ФС новой особи при выполнении нескольких правил.

Во-первых, разница плана стены дочерней особи от планов стен родительских особей по расстоянию Хэмминга должна составлять не более отношения среднего минимальных протяженностей несущих и ненесущих сегментов по ограничению (2) к шагу разбиения стен.

Во-вторых, разница между протяженностью несущих сегментов дочерней особи и средним протяженностей несущих сегментов родительских особей должна быть наименьшей из допустимых и проходящих первое правило планов.

В-третьих, оценка плана стены должна быть наименьшей из допустимых и проходящих первое и второе правила. Не должны

нарушаться ограничения математической модели на минимально допустимые протяженности несущих и ненесущих сегментов (2).

Алгоритм сегментального обобщения достигает выполнения указанных правил благодаря операциям балансировки и обработки нарушений, а покрытия всевозможных допустимых вариантов – благодаря механизму критических точек (КТ). Алгоритм проходится вдоль стены, расставляя ФС участкам в соответствии с выполнением ограничений или балансом. Если ограничения не выполняются, то следующий участок будет установлен с ФС аналогичным нынешнему. Иначе, ФС устанавливается по той родительской стене, баланс для которой нарушен (рисунок 3). Далее алгоритм описывается подробнее.

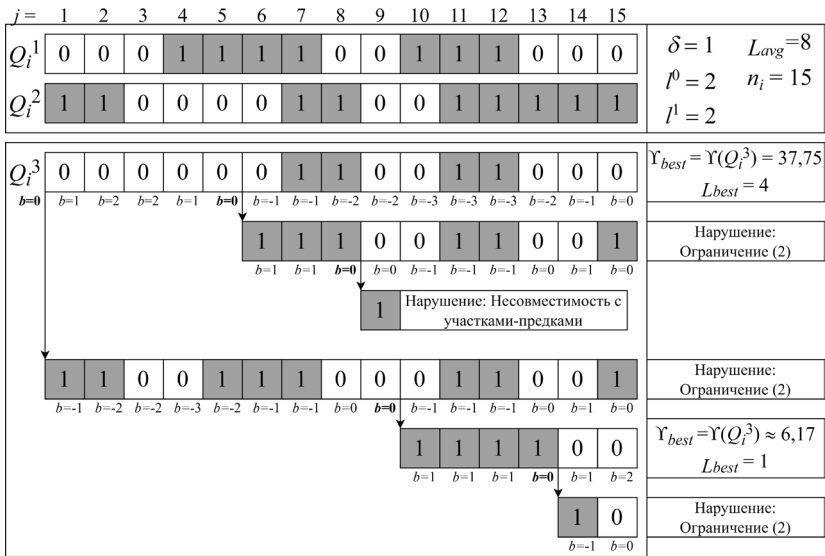


Рис. 3. Пример работы алгоритма сегментального обобщения

Даны два плана опор  $Q^1$  и  $Q^2 \in \Lambda$ . Рассмотрим для некоторой стены  $i \in \{1, \dots, N\}$  их планы сегментов  $Q_i^1 = \{Q_{ik}^1: \forall k = 1, \dots, m_i^1\}$  и  $Q_i^2 = \{Q_{ik}^2: \forall k = 1, \dots, m_i^2\}$ ;  $Q_i^1 \in Q^1$ ,  $Q_i^2 \in Q^2$ . Определим по ним планы ФС в  $i$ -ой стене по формуле (1) как  $X_i^1 = \{X_{ij}^1: X_{ij}^1 = X_{ij}(Q^1), \forall j = 1, \dots, n_i\}$  и  $X_i^2 = \{X_{ij}^2: X_{ij}^2 = X_{ij}(Q^2), \forall j = 1, \dots, n_i\}$ . Сформируем по ним новый план  $X_i^3 = \{X_{ij}^3: X_{ij}^3 \in \{0, 1\}, \forall j = 1, \dots, n_i\}$ . Введем параметр баланса плана-потомка между планами-предками  $b \in \mathbb{Z}$ , текущий тип ФС  $v \in \{0, 1\}$  и стек КТ  $J \subseteq \{j | \forall j = 1, \dots, n_i\}$ , хранящий

индексы участков с равновесным состоянием  $b = 0$ . Описание алгоритма сегментального обобщения представлено на листинге 3.

```

1.  $j := 1$  – индекс участка;  $b := 0$  – баланс;  $v := 0$  – тип ФС;  $J := \{1\}$  – стек КТ;
2. while  $J \neq \emptyset$  do
3.   if  $j = n_i$  then
4.      $j := J.pop()$ ;  $v := 1$ ;
5.   end if;
6.   while  $j \leq n_i$  do
7.      $j^* := \min(\{j^{**}: X_{ij^{**}}^3 = v, j^{**} = 1, \dots, j\})$ ;
8.     if  $j \neq 1$  и  $\sum_{j=j^*, \dots, j} (I_{ij}) \geq I^v$  (ограничение (2)) then
9.       if  $b = 0$  then
10.         $v := 0$ ;  $J.push(\{j\})$  – новая КТ;
11.       end if;
12.     else
13.       if  $b < 0$  then  $v := X_{ij}^1$ ;
14.       else if  $b > 0$  then  $v := X_{ij}^2$ ;
15.     end else;
16.      $X_{ij}^3 := v$ ;
17.     if  $X_{ij}^3 \neq X_{ij}^1$  и  $X_{ij}^3 \neq X_{ij}^2$  (нарушение) then
18.        $j := J.pop()$ ;  $v := 1$ ;
19.     else
20.       if  $X_{ij}^3 \neq X_{ij}^1$  then  $b := b - 1$ ;
21.       else if  $X_{ij}^3 \neq X_{ij}^2$  then  $b := b + 1$ ;
22.        $j := j + 1$ ;
23.     end else;
24.   end while;
25.    $j^* := \min(\{j^{**}: X_{ij^{**}}^3 = v, j^{**} = 1, \dots, j\})$ ;
26.   if  $\sum_{j=j^*, \dots, j} (I_{ij}) > I^v$  (ограничение (2)) then
27.     Оценка построенного плана;
28.   end if;
29. end while;

```

Листинг 3. Псевдокод алгоритма сегментального обобщения

Суть оценки плана (шаг 27) состоит в расчёте суммы отношений квадрата протяженности прилегающих несущих сегментов к каждому несущему сегменту; при этом ситуации, когда несущие сегменты расположены по краям стены, оцениваются хуже. Наилучший план сегментов должен иметь наименьшую оценку.

Для сравнения различных планов ФС построим полученный план сегментов в  $i$ -ой стене. Для этого определим количество сегментов в новом плане как  $m_i^3 = 1 + |\{j: X_{ij}^3 \neq X_{ij+1}^3, \forall j = 1, \dots, n_i - 1\}|$ . Определим количество участков в первом сегменте нового плана как

$\eta_{i1}^3 = \min(\{j: X_{ij}^3 \neq X_{ij+1}^3, j = 1, \dots, n_i - 1\} \cup \{n_i\})$ . Тогда количество участков сегментов  $\forall k = 2, \dots, m_i^3$  определяется следующим образом:

$$\eta_{ik}^3 = \min(\{j - \eta(k) : X_{ij}^3 \neq X_{ij+1}^3, j = \overline{\eta(k), \dots, n_i}\} \cup \{n_i - \eta(k)\}),$$

где  $\eta(k) = \sum_{k^*=1, \dots, k-1} (\eta_{ik^*}^3)$  – количество участков до  $k$ -го сегмента. План сегментов по новому плану ФС задан как  $Q_i^3 = \{Q_{ik}^3: k = 1, \dots, m_i^3\}$ , где  $k$ -ый сегмент задается как  $Q_{ik}^3 = \{j: j = \eta(k) + 1, \dots, \eta(k) + \eta_{ik}^3\}$ .

Определим число несущих участков по плану  $Q$  как  $L(Q_i)$ :

$$L(Q_i) = \sum_{j=1}^{n_i} X_{ij}(Q), Q_i \in Q.$$

Определим лучший план как  $Q_i^{best}$ , среднее количество несущих участков родительских планов как  $L_{avg} \in \mathbb{R}$ , разницу длин несущих участков между лучшим планом и  $L_{avg}$  как  $L_{best} \in \mathbb{R}$ , оценку наилучшего плана как  $Y_{best} \in \mathbb{R}$ . Изначально  $Q_i^{best} = \emptyset$ ,  $L_{best} = \infty$ ,  $Y_{best} = \infty$ . Описание алгоритма оценки плана представлено на листинге 4.

1. **Begin**
2.  $Q_i^{best} = \emptyset; L_{best} = \infty; Y_{best} = \infty; \delta \in \mathbb{Z}$  – шаг разбиения стен на участки;
3.  $L_{avg} = (L(Q_i^1) + L(Q_i^2)) / 2$ ;
4. **if**  $|b| > |l| + l^0 / (2 \cdot \delta)$  **then End** (по первому правилу);
5. **if**  $|L(Q_i^3) - L_{avg}| > L_{best}$  **then End** (по второму правилу);
6. **if**  $m_i^3 = 1$  **then** (стена из одного типа ФС)
7.     **if**  $X_{ij}^3 = 1$  **then** (полностью несущая стена)
8.         **if**  $L_{avg} \leq 3 \cdot n_i / 5$  **then**  $Y(Q_i^3) = \infty$ ;
9.         **else**  $Y(Q_i^3) = n_i / (L_{avg} - 3 \cdot n_i / 5)^3$ ;
10.     **else** (полностью ненесущая стена)
11.         **if**  $L_{avg} \geq 2 \cdot n_i / 5$  **then**  $Y(Q_i^3) = \infty$ ;
12.         **else**  $Y(Q_i^3) = n_i / (2 \cdot n_i / 5 - L_{avg})^3$ ;
13.     **end else**;
14. **else** (стена из разных типов ФС)
15.      $Y_k \in \mathbb{R}$  – слагаемое оценки;  $\gamma_k \in \{1, 2, 3\}$  – коэффициент по стыку;
16.      $k_s \in \{1, \dots, m_i^3\}$  – индекс несущего;  $k_{ns} \in \{1, \dots, m_i^3\}$  – индекс ненесущего;
17.     **for**  $k = \{1, \dots, m_i^3 - 1\}$  **do**
18.         **if**  $\chi_{ik}^3 = 0$  **then**  $k_{ns} = k; k_s = k + 1$ ;
19.         **else**  $k_{ns} = k + 1; k_s = k$ ;
20.         **if**  $k_{ns} = 1$  или  $k_{ns} = m_i^3$  **then**  $\gamma_k = 3$ ;
21.         **else if**  $k_s \neq 1$  и  $k_s \neq m_i^3$  **then**  $\gamma_k = 2$ ;
22.         **else**  $\gamma_k = 1$ ;

23. Определим слагаемое по стыку сегментов:

$$Y_k = \gamma_k \cdot \frac{(\text{length}(Q_{ik_{ns}}^3))^2}{2 \cdot \text{length}(Q_{ik_s}^3)};$$

24. **end for**;  
25. Оценка полученного плана опор  $Q_i^3$ :

$$Y(Q_i^3) = \left( \sum_{k=1}^{m_i^3-1} Y_k \right) \sum_{j=1}^{n_i} X_{ij} / n_i;$$

26. **end else**;  
27. **if**  $Y(Q_i^3) < Y_{best}$  **then**  
28.  $Y_{best} = Y(Q_i^3)$ ;  
29.  $L_{best} = |L(Q_i^3) - L_{avg}|$ ;  
30.  $Q_i^{best} = Q_i^3$ ;  
31. **end if**;  
32. **End**;

Листинг 4. Псевдокод алгоритма оценки построенного плана

Таким образом, результатом работы алгоритма сегментального обобщения является наилучший план распределения  $Q_i^{best}$ .

### 3.3. Модифицированный алгоритм оператора мутации.

Операция мутации применяется для избегания чрезмерной сходимости популяции к однотипным решениям и позволяет увеличивать разнообразие популяции точечными изменениями в генотипе особи. Однако при разработке алгоритма мутации необходимо следить за нарушениями допустимости решения, т.е. ограничений задачи.

В связи с сегментальным представлением генотипа в задаче рационального размещения несущих стен было принято решение разработать алгоритм изменения генов в цельных сегментах. Суть алгоритма состоит в попытке уменьшить протяженность наибольших сегментов в различных стенах некоторого плана. Для этого обозначим множество наибольших сегментов по каждой стене как  $Sg$ :

$$Sg = \{(i, k) : \text{length}(Q_{ik}) = \max(\text{length}(Q_{ik^*})) \\ \forall k, k^* = \{1, \dots, m_i\}, \forall i = \{1, \dots, N\}\}$$

Обозначим отдельный сегмент за  $sg = (i, k)$ , компоненты кортежа обозначим за  $sg.i$  – индекс стены,  $sg.k$  – индекс сегмента. Упорядочим сегменты из  $Sg$  по их длине. Пусть мутации подвергнуты первые

несколько из них  $Sg' \subseteq Sg$ . Рассмотрим один из этих сегментов  $sg' \in Sg'$ ,  $X' = \chi_{sg'.i \ sg'.k}$  – ФС сегмента. Длина уменьшения сегмента равна минимальной протяженности по противоположному ФС  $L_{\downarrow} = l^{1-X'}$ . Если протяженность сегмента после уменьшения меньше допустимой  $length(sg') - L_{\downarrow} < l^X$ , то алгоритм обрабатывает следующий сегмент.

Если сегмент – единственный в стене  $m_{sg'.i} = 1$ , то создается новый сегмент по случайному направлению минимальной длины  $sg'_{new} = (i, 1 \text{ или } 2)$ ,  $length(sg'_{new}) = L_{\downarrow}$ , ФС нового сегмента  $X'_{new} = 1 - X'$ . Если сегмент не единственный, то из направлений  $-sg' = (sg'.i, \max(sg'.k - 1, 1))$  (предыдущего) и  $+sg' = (sg'.i, sg'.k + 1)$  (последующего) выбирается направление с меньшим сегментом и увеличивается на минимальную протяженность  $\min(length(-sg'), length(+sg')) + L_{\downarrow}$ . Протяженность исходного сегмента в обоих случаях уменьшается на соответствующую длину  $length(sg') - L_{\downarrow}$ . Таким образом, удастся изменить генотип особи, не нарушая установленных ограничений.

**3.4. Схема оператора селекции по дискриминированному рангу.** Алгоритм оператора селекции основан на дискриминированном ранге по элитарности  $R_{dis}$  (8), описанном ранее. Селекция направлена на отбор особей с наилучшими рангами элитарности и наибольшими показателями разнообразия. Пусть  $\Lambda = \{Q^1, \dots, Q^{\Omega}\}$ , – исходная популяция, тогда  $\Omega$  – исходная мощность популяции,  $\Lambda_{new} = \emptyset$  – новая популяция. Алгоритм селекции разделен на три этапа (рисунок 4).

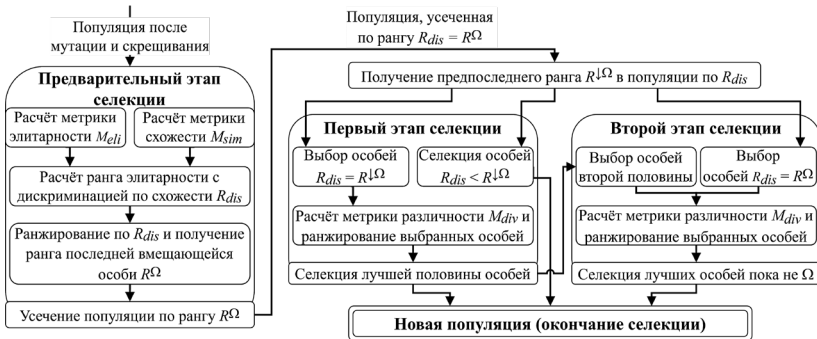


Рис. 4. Схема разработанного алгоритма оператора селекции

На предварительном этапе особи, полученные при мутации  $\Lambda_{mut}$  и скрещивании  $\Lambda_{scr}$ , объединяются с исходной популяцией:

$$\Lambda^+ = \Lambda \cup \Lambda_{\text{мут}} \cup \Lambda_{\text{скр}}, \quad \Omega^+ = \left| \Lambda^+ \right|,$$

где  $\Omega^+$  – мощность итоговой популяции. Полученная популяция  $\Lambda^+$  упорядочивается по  $R_{\text{dis}}$  (8) и усекается рангом  $\Omega$ -ой особи:

$$\Lambda_{\text{dis}}^+ = \{Q^\zeta \in \Lambda^+ \mid R_{\text{dis}}(Q^\zeta) \leq R_{\text{dis}}(Q^{\zeta+1}), \forall \zeta = \{1, \dots, \Omega^+\}\};$$

$$R^\Omega = R_{\text{dis}}(Q^\zeta), Q^\zeta \in \Lambda_{\text{dis}}^+, k = \Omega;$$

$$\Lambda_{\text{dis}}^* = \{Q^\zeta \in \Lambda_{\text{dis}}^+ \mid R_{\text{dis}}(Q^\zeta) \leq R^\Omega, \forall \zeta = \{1, \dots, \Omega^+\}\}, \quad \Omega^* = \left| \Lambda_{\text{dis}}^* \right|, \quad (9)$$

где  $\Omega^*$  – мощность усеченной популяции  $\Lambda_{\text{dis}}^*$ . Таким образом, к селекции остаются лишь особи с лучшими рангами  $R_{\text{dis}}$ . Если в усеченной популяции (9) остались особи с разными рангами, то проводится первый этап селекции: в новую популяцию добавляются особи с рангом меньшим предпоследнего ранга  $R^{\Omega}$  и выделяется часть популяции, ранг которых равен предпоследнему рангу:

$$R^{\Omega\downarrow} = \max(\{Q^\zeta \in \Lambda_{\text{dis}}^* \mid R_{\text{dis}}(Q^\zeta) < R^\Omega, \forall \zeta = \{1, \dots, \Omega^*\}\});$$

$$\Lambda_{\text{new}} = \Lambda_{\text{new}} \cup \{Q^\zeta \in \Lambda_{\text{dis}}^* \mid R_{\text{dis}}(Q^\zeta) < R^{\Omega\downarrow}, \forall \zeta = \{1, \dots, \Omega^*\}\};$$

$$\Lambda^{\Omega\downarrow} = \{Q^\zeta \in \Lambda_{\text{dis}}^* \mid R_{\text{dis}}(Q^\zeta) = R^{\Omega\downarrow}, \forall \zeta = \{1, \dots, \Omega^*\}\}, \quad \Omega^\downarrow = \left| \Lambda^{\Omega\downarrow} \right|, \quad (10)$$

где  $\Omega^\downarrow$  – количество особей с рангом, равным  $R^{\Omega\downarrow}$ . Тогда особи с предпоследним рангом (10) упорядочиваются по метрике различности  $M_{\text{div}}$  (6), лучшая половина из них добавляется в новую популяцию и удаляется из усеченной популяции:

$$\Lambda_{\text{div}}^{\Omega\downarrow} = \{Q^\zeta \in \Lambda^{\Omega\downarrow} \mid M_{\text{div}}(Q^\zeta) \leq M_{\text{div}}(Q^{\zeta+1}), \forall \zeta = \{1, \dots, \Omega^\downarrow\}\}.$$

$$\Lambda_{\text{new}} = \Lambda_{\text{new}} \cup \{Q^\zeta \in \Lambda_{\text{div}}^{\Omega\downarrow} \mid \forall \zeta = \{1, \dots, \Omega^\downarrow/2\}\};$$

$$\Lambda_{\text{dis}}^* = \Lambda_{\text{dis}}^* \setminus \{Q^\zeta \in \Lambda_{\text{div}}^{\Omega\downarrow} \mid \forall \zeta = \{1, \dots, \Omega^\downarrow/2\}\}.$$

После первого этапа в усеченной популяции  $\Lambda_{dis}^*$  остаются только особи с последним рангом и половина особей с предпоследним рангом элитарности с большими значениями метрики различности. На втором этапе оставшиеся в усеченной популяции особи ранжируются по метрике различности  $M_{div}$  (6) и дозаполняют новую популяцию. На обоих этапах оценка метрики различности ищется в том числе и на множестве уже добавленных в новую популяцию особей.

$$\Lambda_{div}^* = \{Q^\zeta \in \Lambda_{dis}^* \mid M_{div}(Q^\zeta) \leq M_{div}(Q^{\zeta+1}), \forall \zeta = \{1, \dots, \Omega^*\}\},$$

$$\Lambda_{new} = \Lambda_{new} \cup \{Q^\zeta \in \Lambda_{div}^* \mid \forall \zeta = \{1, \dots, \Omega - |\Lambda_{new}|\}\}.$$

Представленные метрики и алгоритмы генетических операторов позволяют полноценно применить многокритериальные генетические алгоритмы для задачи рационального размещения стен. Описанные метрики различности, схожести и элитарности и основанные на них ранг разнообразия и дискриминированный ранг элитарности позволяют оценивать особи популяции как с точки зрения доминируемости, так и с точки зрения разнообразия генотипа и фенотипа.

Разработанные алгоритмы скрещивания и мутации позволяют создать новые планы без нарушения ограничений задачи. Предложенный алгоритм селекции использует описанные метрики, поддерживая как Парето-сходимость, так и разнообразие популяции.

**4. Численный эксперимент: сравнение эффективности работы многокритериального генетического алгоритма.** Для проведения численных экспериментов было разработано программное средство, реализующее описанные математическую модель и алгоритмы. Программа выполнена в виде надстройки над САПР Revit Autodesk версии 2021 года. Численные эксперименты были проведены на машине со следующими характеристиками: 4-х ядерный Intel Core i7-7700 3.60GHz; RAM 32Gb. Проект для эксперимента предоставлен компанией ООО «Конструктив», г. Уфа.

Проект содержит 31 стену, из которых изменяемых стен  $N = 21$ . Общая протяженность стен – 98.670 м, протяженность изменяемых стен – 61.480 м. Шаг разбиения стен на участки  $\delta = 125$  мм, суммарное количество участков стен – 485. Ограничения протяженности  $l^0$  и  $l^1 = 500$  мм. Эксперимент направлен на сравнение разработанного алгоритма с другими МГА при решении задачи рационального



размещения стен. Для сравнения были выбраны известные модели МГА [17], их описание содержится в таблице 1. Настройки МГА: мощность популяции – 350, количество итераций – 375, доля скрещивания – 20% от мощности популяции; доля обобщаемых сегментов – 20% стен, доля мутации – 15% от популяции, шанс мутации для случайной стены – 25%.

Таблица 1. Многокритериальные генетические алгоритмы для сравнения

	Наименование	Описание
[15]	Optimum Order Multiobjective Genetic Algorithm	выстраивается новая многокритериальная задача по рангу недоминируемости и метрике схожести; строится общая метрика, аналогичная рангам Парето-доминирования
[16]		модификация; вводится дискриминация за однообразие особей; применяется в численном эксперименте
[18]	Non-dominated Sorting Genetic Algorithm	упорядочивание по рангам недоминируемости; для равномерности покрытия фронта Парето применяется метрика разреженности на основе метрики Манхэттена
[19]	Non-dominated Sorting Genetic Algorithm-II	для улучшения качества Парето-аппроксимации используется инструмент сохранения лучших индивидов в отдельном архиве
[20]	Strength Pareto Evolutionary Algorithm 2	применяется селекция по Парето-доминированию, нишевание, архивирование; ранжирование по слабости – суммарной Парето-силе доминирующих особей, где Парето-сила – количество доминируемых особей
[21]	Dynamical Multiobjective Evolutionary Algorithm	основан на принципе минимальной свободной энергии термодинамики; применяется агрегирующая функция, сочетающая ранжирование с энтропией и плотностью
[22]	Knee Point Driven Evolutionary Algorithm	для увеличения разнообразия в популяции использует точки перегиба – решения, для которых улучшение одной цели приведет к сильному ухудшению другой

Эффективность работы алгоритмов оценивается по времени выполнения алгоритма, по количеству найденных недоминируемых решений в популяции алгоритма и на общем множестве среди всех алгоритмов, а также по оценке разнообразия популяций, полученных алгоритмами, в индивидуальной популяции и на общем множестве. Для оценки разнообразия использовался алгоритм, оценивающий разнообразие популяций по количеству получаемых по особям кластеров относительно общего количества особей. Для построения кластеров использован алгоритм DBSCAN [23].

Для сравнения решений алгоритмов было проведено несколько испытаний с разным ограничением минимальной общей протяженности несущих стен, т.е. ограничение снизу для  $f_L$ . Испытания разделены на две группы: 1-я – с ограничением в 30%;

и 2-я – с ограничениями выше. В первой группе было произведено три испытания. Во второй группе проведено по одному испытанию для каждого ограничения (35%, 45%, ..., 85%). Результаты оценки времени выполнения представлены в таблице 2, оценки количества получаемых недоминируемых решений – в таблице 3 и таблице 4, оценки разнообразия полученных популяций особей – в таблице 5 и таблице 6 ((PA) – разработанный алгоритм).

Таблица 2. Время выполнения алгоритмов (минуты : секунды)

1 гр.	(PA)	OOMOGA	NSGA	NSGA-II	SPEA2	DMOEA	KNEA
< 30%	27:52	18:29	13:39	06:55	21:42	33:01	05:58
	25:57	24:12	15:32	14:29	32:35	50:17	05:41
	26:18	19:48	14:02	07:54	23:31	34:36	06:49
2 гр.	(PA)	OOMOGA	NSGA	NSGA-II	SPEA2	DMOEA	KNEA
35%	30:11	18:28	21:33	11:00	22:58	36:36	07:52
45%	28:47	19:56	30:40	15:39	26:07	49:00	09:36
55%	30:35	23:08	28:19	19:52	32:07	67:25	09:37
65%	32:42	22:20	14:46	10:51	22:23	40:47	06:54
75%	25:27	21:39	14:05	11:19	29:26	44:28	07:52
85%	24:23	19:48	10:00	09:25	16:12	18:03	09:11

По времени решения разработанный алгоритм работает быстрее DMOEA, сопоставим со SPEA2, и медленнее других. Следует отметить, что время работы разработанного алгоритма практически не уменьшается при повышении ограничения на протяженность, в отличие от других МГА. Это в том числе является следствием большого количества упорядочиваний в генетических операторах. Поэтому одним из способов увеличения скорости работы алгоритма может быть применение более быстрых алгоритмов упорядочивания.

Таблица 3. Количество недоминируемых решений в своей популяции (шт.)

1 гр.	(PA)	OOMOGA	NSGA	NSGA-II	SPEA2	DMOEA	KNEA
< 30%	174	95	49	38	54	97	23
	143	90	49	174	55	195	13
	173	100	42	70	48	92	25
2 гр.	(PA)	OOMOGA	NSGA	NSGA-II	SPEA2	DMOEA	KNEA
35%	147	90	43	143	44	148	15
45%	139	73	54	94	38	128	14
55%	66	52	43	33	54	90	8
65%	33	46	15	13	29	27	10
75%	48	21	18	35	18	18	11
85%	19	15	13	28	16	18	3

Таблица 4. Количество недоминируемых решений в общей популяции (шт.)

1 гр.	(PA)	OOMOGA	NSGA	NSGA-II	SPEA2	DMOEA	KNEA
< 30%	96	1	20	22	28	44	0
	68	5	9	12	47	26	1
	75	1	4	39	34	66	1
2 гр.	(PA)	OOMOGA	NSGA	NSGA-II	SPEA2	DMOEA	KNEA
35%	129	2	7	32	40	6	1
45%	126	1	30	2	0	31	0
55%	65	1	0	0	7	11	0
65%	22	3	5	0	0	7	0
75%	48	2	0	0	0	0	0
85%	19	0	0	0	0	7	0

Таблица 5. Разнообразие решений в своей популяции (% участия в кластерах)

1 гр.	(PA)	OOMOGA	NSGA	NSGA-II	SPEA2	DMOEA	KNEA
< 30%	6,6	6,79	3,21	2,64	4,53	3,96	2,64
	3,76	3,89	2,09	5,7	3,89	5,29	1,67
	6,18	8,18	2,91	3,82	2,91	4,18	3,64
2 гр.	(PA)	OOMOGA	NSGA	NSGA-II	SPEA2	DMOEA	KNEA
35%	3,65	4,76	3,65	6,03	2,38	4,6	2,06
45%	3,15	3,33	1,11	5,37	1,85	4,26	1,48
55%	2,31	2,02	5,2	4,05	3,18	2,89	1,73
65%	4,05	5,78	4,05	2,31	3,47	3,47	2,89
75%	5,92	5,33	4,14	5,33	2,96	2,96	4,14
85%	4,46	6,25	4,46	7,14	4,46	7,14	2,68

Таблица 6. Разнообразие решений в общей популяции (% участия в кластерах)

1 гр.	(PA)	OOMOGA	NSGA	NSGA-II	SPEA2	DMOEA	KNEA
< 30%	6,16	0,47	4,27	2,84	7,11	4,74	0
	5,36	1,79	3,57	5,95	13,1	5,95	0,6
	6,36	0,45	1,36	3,18	5,45	3,64	0,45
2 гр.	(PA)	OOMOGA	NSGA	NSGA-II	SPEA2	DMOEA	KNEA
35%	7,37	0,92	1,38	5,07	5,99	0,92	0,46
45%	8,95	0,53	2,63	0,53	0	6,84	0
55%	7,14	1,19	0	0	3,57	4,76	0
65%	16,22	8,11	5,41	0	0	5,41	0
75%	20	4	0	0	0	0	0
85%	23,08	0	0	0	0	15,38	0

По количеству недоминируемых решений на индивидуальной популяции разработанный алгоритм в среднем находит в 1.28 раз больше по сравнению с DMOEA, в 1.58 раз – с OOMOGA, в 1.89 раз – с NSGA-II, и больше – с остальными. На общем множестве особей разработанный алгоритм находит в среднем в 3 раза больше

недоминируемых решений в сравнении с DMOEA, в 4 раза больше в сравнении с NSGA-II, и больше в сравнении с другими МГА.

По разнообразию найденных решений на индивидуальной популяции разработанный алгоритм справляется в среднем в 1.11 раз хуже OOMOGA, наравне с NSGA-II и DMOEA, и в 1.5 раза лучше других алгоритмов. Однако каждый из указанных выше алгоритмов получает разнообразие популяции с меньшей мощностью Парето-оптимального множества решений, что оставляет потенциал для дальнейшего улучшения описанного в статье генетического алгоритма в сторону увеличения разнообразия получаемой популяции.

Дополнительно, представлены графики распределения особей разных алгоритмов в пространстве критериев на рисунке 5. Каждая ось отображает оценку по критерию (3), (4) или (5).

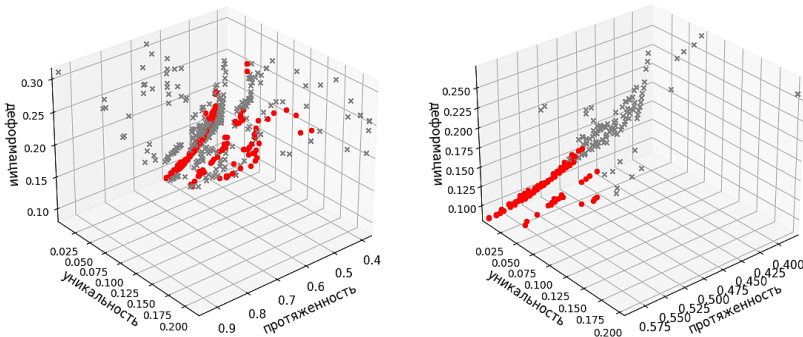


Рис. 5. Распределение особей; слева – по популяции; справа – по общему множеству; круги – особи разработанного МГА, кресты – особи других МГА

По графикам видно, что разработанный алгоритм лучше находит планы с большим количеством несущих стен и меньшими деформациями, жертвуя разнообразием популяции решений в пользу их недоминируемости.

Таким образом, по проведенным экспериментам была подтверждена эффективность работы разработанного алгоритма. Преимуществом описанного алгоритма является нахождение большего числа недоминируемых решений по сравнению с другими МГА, а также сопоставимое разнообразие получаемых решений. Недостатком алгоритма является большое время выполнения. Кроме того, продемонстрированы возможности разработанной модели задачи, в т.ч. при применении многокритериальных генетических алгоритмов.

**5. Заключение.** Представлена новая модель задачи рационального размещения несущих стен в многокритериальной комбинаторной постановке. По результатам численного эксперимента подтверждается, что предлагаемая модель обладает достаточной гибкостью и репрезентативностью проблематики поставленной задачи для нахождения разнообразных решений. Модель учитывает три критерия задачи: протяженности стен, количества различных длин стен и оценки деформаций по разработанной модели [14].

Был разработан новый многокритериальный генетический алгоритм, учитывающий особенности задачи рационального размещения стен. Был разработан алгоритм сегментального обобщения для операции скрещивания, находящий наилучший план стены, наиболее схожий с родительскими планами. Был предложен алгоритм мутации, обрабатывающий гены сегментно. Кроме того была разработана схема селекции, направленная на поддержку разнообразия в популяции как в фенотипическом, так и в генотипическом смысле.

Были проведены эксперименты анализа эффективности разработанного генетического алгоритма в сравнении с другими МГА. По количеству найденных недоминируемых решений разработанный алгоритм в сравнении с ближайшим конкурентным МГА ДМОЕА получил в 1.3 раз и в 3 раза больше Парето-оптимальных решений по популяции и общему множеству соответственно. Алгоритм показал равную оценку разнообразия популяции в сравнении с не специализированными МГА в своей популяции, уступив лишь модификации алгоритма ООМОГА. Одним из недостатков представленного алгоритма является высокое время работы, в среднем – в 2 раза дольше других алгоритмов, что может быть вызвано неоптимизированным подходом к упорядочиванию популяции на разных шагах алгоритма. Тем не менее, разработанные модель и генетический алгоритм были реализованы в виде программного продукта и могут быть использованы на практике.

### Литература

1. Petprakob W. Beam-slab floor optimization using genetic and particle swarm optimization algorithms // A Thesis for degree of master of science in engineering and technology. Thailand: Thammasat University, Sirindhorn International Institute of Technology. 2014. 90 p. DOI: 10.14457/TU.the.2014.490.
2. Nimitawat A., Nanakorn P. Automated layout design of beam-slab floors using a genetic algorithm // Computers & Structures. 2009. vol. 87(21-22). pp. 1308–1330. DOI: 10.1016/j.compstruc.2009.06.007.
3. Sharafi P. Cost optimization of the preliminary design layout of reinforced concrete framed buildings // A Thesis for degree of Doctor of Philosophy in civil engineering. Australia: University of Wollongong, School of Civil, Mining and Environmental Engineering. 2013. 286 p.

4. Hadi M.H., Sharafi P., Teh L.H. A new formulation for the geometric layout optimisation of flat slab floor systems // *Proceeding of the Australasian Structural Engineering Conference 2012: The Past, Present and Future of Structural Engineering*. Perth, Western Australia: Engineers Australia, 2012. pp. 122–129. DOI: 10.3316/informit.026275045488938.
5. Sahab M.G., Ashour A.F., Toropov V.V. Cost optimisation of reinforced concrete flat slab buildings // *Engineering Structures*. 2005. vol. 27(3). pp. 313–322. DOI: 10.1016/j.engstruct.2004.10.002.
6. Meng X., Lee T.U., Xiong Y., Huang X., Xie Y.M. Optimizing support locations in the roof–column structural system // *Applied Science*. 2021. vol. 11(6). DOI: 10.3390/app11062775.
7. Zelickman Y., Amir O. Optimization of column layouts in buildings considering structural and architectural constraints // *Engineering Archive (engrXiv)*. 2024. 45 p. DOI: 10.31224/2723.
8. Steiner B., Mousavian E., Mehdizadeh Saradj F., Wimmer M., Musialski P. Integrated structural-architectural design for interactive planning // *Computer Graphics Forum*. 2016. vol. 36(8). pp. 80–94. DOI: 10.1111/cgf.12996.
9. Валиахметова Ю.И., Васильева Л.И., Зинов В.И. Алгоритм решения задачи определения рационального плана размещения несущих конструкций при строительстве многоэтажных зданий // *Современные проблемы и перспективы развития естествознания: Сб. научн. тр. Национальной научно-практической конференции (г. Уфа, 8–9 июня 2020 г.)*, Т. 2. Уфа: БГПУ, 2020. С. 46–54.
10. Eleftheriadi S., Duffour P., Stephenson B., Mumovic D. Automated specification of steel reinforcement to support the optimisation of RC floors // *Automation in Construction*. 2018. vol. 96. pp. 366–377. DOI: 10.1016/j.autcon.2018.10.005.
11. Wang J., Chen K., Yang H., Zhang L. Ensemble deep learning enabled multi-condition generative design of aerial building machine considering uncertainties // *Automation in Construction*. 2024. vol. 157. DOI: 10.1016/j.autcon.2023.105134.
12. Лешкевич О.Н. Использование искусственных нейронных сетей для оценки армирования железобетонных плит перекрытия // *Проблемы современного бетона и железобетона: Сб. научн. тр. № 11*. Минск: Ин-т БелНИИС, 2019. С. 51–62. DOI: 10.35579/2076-6033-2019-11-04.
13. Liao W.J., Lu X.Z., Fei Y.F., Gu Y., Huang Y.L. Generative AI design for building structures // *Automation in Construction*. 2024. vol. 157. DOI: 10.1016/j.autcon.2023.105187.
14. Зинов В.И., Картак В.М., Валиахметова Ю.И. Алгоритм оценки деформации плит перекрытий по пролетно-опорным схемам здания // *Системы анализа и обработки данных*. 2023. № 4(92). С. 35–54. DOI: 10.17212/2782-2001-2023-4-35-54.
15. Long Q., Wu C.Z., Wang X.Y., Jiang L., Li J. A multiobjective genetic algorithm based on a discrete selection procedure // *Mathematical Problems in Engineering*. 2015. vol. 2015. 17 p. DOI: 10.1155/2015/349781.
16. Валиахметова Ю.И., Прокудина Е.И., Зинов В.И. Многокритериальный генетический алгоритм решения задачи рационального размещения опорных конструкций на плане зданий // *Современные физика, математика, цифровые и нанотехнологии в науке и образовании: Сб. научн. тр. II Всероссийской молодежной школы-конференции (г. Уфа, 18–20 апреля 2023 г.)*. Уфа: БГПУ, 2023. С. 95–100.
17. Rahimi I., Gandomi A.H., Nikoo M.R., Chen F. A comparative study on evolutionary multi-objective algorithms for next release problem // *Applied Soft Computing*. 2023. vol. 144. DOI: 10.1016/j.asoc.2023.110472.

18. Srinivas N., Deb K. Multiobjective optimization using nondominated sorting in genetic algorithms // *Evolutionary Computation*. 1994. vol. 2(3). pp. 221–248. DOI: 10.1162/evco.1994.2.3.221.
19. Deb K., Pratap A., Agarwal S., Meyarivan T. A fast and elitist multiobjective genetic algorithm: NSGAII // *IEEE Transactions on Evolutionary Computation*. 2002. vol. 6(2). pp. 182–197. DOI: 10.1109/4235.996017.
20. Zitzler E., Laumanns M., Thiele L. SPEA2: Improving the strength Pareto evolutionary algorithm // *TIK-Report*. 2001. vol. 103. DOI: 10.3929/ethz-a-004284029.
21. Liu M., Zou X., Chen Y., Wu Z. Performance assessment of DMOEA-DD with CEC 2009 MOEA Competition Test Instances // *Proceeding of the IEEE Congress on Evolutionary Computation (CEC 2009)*. Norway: Trondheim, 2009. pp. 2913–2918. DOI: 10.1109/CEC.2009.4983309.
22. He M., Wang X., Chen H., Li X. A Knee Point-Driven Many-Objective Evolutionary Algorithm with Adaptive Switching Mechanism // *Journal of Applied Mathematics*. 2024. vol. 2024(1). 43 p. DOI: 10.1155/2024/4737604.
23. Ester M., Kriegel H.-P., Sander J., Xu X. A density-base algorithm for discovering clusters in large spatial databases with noise // *Proceeding of 2nd International Conference on Knowledge Discovery and Data Mining (KDD-96)*. Germany: Munich, 1996. pp. 226–231.

**Зинов Владислав Игоревич** — аспирант, кафедра технической кибернетики, Уфимский университет науки и технологий. Область научных интересов: оптимизация в сфере проектирования зданий, эвристические и метаэвристические методы, системы поддержки принятия решений. Число научных публикаций — 6. zinovvladislavufa@gmail.com; улица Заки Валиди, 32, 450076, Уфа, Россия; р.т.: +7(917)435-5560.

**Картак Вадим Михайлович** — д-р физ.-мат. наук, профессор, заведующий кафедрой, кафедра вычислительной техники и защиты информации, Уфимский университет науки и технологий. Область научных интересов: большие данные, методы оптимизации, информационная безопасность. Число научных публикаций — 80. kvmail@mail.ru; улица Заки Валиди, 32, 450076, Уфа, Россия; р.т.: +7(917)433-3919.

**Валнахметова Юлия Ильясовна** — канд. техн. наук, доцент кафедры, кафедра вычислительной математики и кибернетики, Уфимский университет науки и технологий. Область научных интересов: эвристические алгоритмы решения задач дискретной оптимизации, математическое моделирование, численные методы, комплексы программ. Число научных публикаций — 60. julikas@inbox.ru; улица Заки Валиди, 32, 450076, Уфа, Россия; р.т.: +7(917)348-1947.

V. ZINOV, V. KARTAK, Y. VALIAKHMETOVA  
**SOLVING MULTI-OBJECTIVE RATIONAL PLACEMENT  
OF LOAD-BEARING WALLS PROBLEM VIA GENETIC  
ALGORITHM**

*Zinov V., Kartak V., Valiakhetova Y. Solving Multi-Objective Rational Placement of Load-Bearing Walls Problem via Genetic Algorithm.*

**Abstract.** The rational placement of load-bearing walls remains a complex and poorly studied problem, despite the existence of numerous algorithms and models for solving the similar problem of column placement. The main complexity factors are the large number of alternative solutions, the significant time required to calculate deformations for a given wall placement, and the multi-objective nature of the problem. In addition to the nonlinear criterion for estimating deformations, it is necessary to minimize the length of load-bearing walls and the number of their unique lengths. A model for the rational placement of load-bearing walls is proposed, which divides the walls into functional parts with a specific step and considers all the required target criteria. Adjacent wall parts with the same functionality are combined into segments. The combinatorial formulation applied in the model of the problem allows the use of genetic algorithms as a solution tool. Therefore, a new approach to multi-objective genetic algorithm is proposed, containing metrics for calculating population diversity at the phenotype and genotype levels. Modifications of crossover, mutation, and selection operators, considering the segmental structure of the wall's genotype, are presented. A comparative analysis of the developed algorithm with other known multi-objective genetic algorithms showed that the developed algorithm finds, on average, three times more non-dominated solutions, particularly more plans with a lower deformation estimates, despite the twice-longer execution time. The proposed model differs significantly from previous models in terms of handling deformations in slab-support systems, comparing placement plans with each other rather than calculating precise reinforcement estimates, which is often unnecessary at the early stages. The proposed genetic algorithm scheme increases the number of found nondominated solutions without losing their diversity, and can be used to solve other multi-objective problems, taking into account the specified features. The developed algorithm was easily integrated into the CAD-based decision support software and can be used in practice by building designers.

**Keywords:** optimization in building design, wall placement problem, multiobjective genetic algorithm, combinatorial optimization, comparative estimation of slab deformations.

## References

1. Petprakob W. Beam-slab floor optimization using genetic and particle swarm optimization algorithms. A Thesis for degree of master of science in engineering and technology. Thailand: Thammasat University, Sirindhorn International Institute of Technology. 2014. 90 p. DOI: 10.14457/TU.the.2014.490.
2. Nimitawat A., Nanakorn P. Automated layout design of beam-slab floors using a genetic algorithm. *Computers & Structures*. 2009. vol. 87(21-22). pp. 1308–1330. DOI: 10.1016/j.compstruc.2009.06.007.
3. Sharafi P. Cost optimization of the preliminary design layout of reinforced concrete framed buildings. A Thesis for degree of Doctor of Philosophy in civil engineering. Australia: University of Wollongong, School of Civil, Mining and Environmental Engineering. 2013. 286 p.
4. Hadi M.H., Sharafi P., Teh L.H. A new formulation for the geometric layout optimisation of flat slab floor systems. *Proceeding of the Australasian Structural*



- Engineering Conference 2012: The Past, Present and Future of Structural Engineering. Perth, Western Australia: Engineers Australia, 2012. pp. 122–129. DOI: 10.3316/informit.026275045488938.
5. Sahab M.G., Ashour A.F., Toropov V.V. Cost optimisation of reinforced concrete flat slab buildings. *Engineering Structures*. 2005. vol. 27(3). pp. 313–322. DOI: 10.1016/j.engstruct.2004.10.002.
  6. Meng X., Lee T.U., Xiong Y., Huang X., Xie Y.M. Optimizing support locations in the roof–column structural system. *Applied Science*. 2021. vol. 11(6). DOI: 10.3390/app11062775.
  7. Zelickman Y., Amir O. Optimization of column layouts in buildings considering structural and architectural constraints. *Engineering Archive (engrXiv)*. 2024. 45 p. DOI: 10.31224/2723.
  8. Steiner B., Mousavian E., Mehdizadeh Saradj F., Wimmer M., Musialski P. Integrated structural-architectural design for interactive planning. *Computer Graphics Forum*. 2016. vol. 36(8). pp. 80–94. DOI: 10.1111/cgf.12996.
  9. Valiakhmetova Yu.I., Vasil'eva L.I., Zinov V.I. Algoritm reshenija zadachi opredelenija racional'nogo plana razmeshhenija nesushhih konstrukcij pri stroitel'stve mnogojetazhnyh zdanij [An algorithm to solving a rational placement of load-bearing structures problem in the design of a multi-storey building]. *Sovremennye problemy i perspektivy razvitiya estestvoznaniya: Sb. nauchn. tr. Nacional'noj nauchno-prakticheskoy konferencii [Contemporary issues and perspectives on natural science: Collected research papers of National research-to-practise conference]*. Ufa: BSPU, 2020. pp. 46–54. (In Russ.).
  10. Eleftheriadis S., Duffour P., Stephenson B., Mumovic D. Automated specification of steel reinforcement to support the optimisation of RC floors. *Automation in Construction*. 2018. vol. 96. pp. 366–377. DOI: 10.1016/j.autcon.2018.10.005.
  11. Wang J., Chen K., Yang H., Zhang L. Ensemble deep learning enabled multi-condition generative design of aerial building machine considering uncertainties. *Automation in Construction*. 2024. vol. 157. DOI: 10.1016/j.autcon.2023.105134.
  12. Liashkevich A. Ispol'zovanie iskusstvennyh nejronnyh setej dlja ocenki armirovaniya zhelezobetonyh plit perekrytija [Application of artificial neural networks for assessing the reinforcement of reinforced-concrete floor slabs Использование искусственных нейронных сетей для оценки армирования железобетонных плит перекрытия] *Problemy sovremennogo betona i zhelezobetona: Sb. nauchn. tr. [Contemporary Issues of Concrete and Reinforced Concrete: Collected Research Papers]*. no. 11. Minsk: Institute BelNIIS, 2019. pp. 51–62. DOI: 10.35579/2076-6033-2019-11-04. (In Russ.).
  13. Liao W.J., Lu X.Z., Fei Y.F., Gu Y., Huang Y.L. Generative AI design for building structures. *Automation in Construction*. 2024. vol. 157. DOI: 10.1016/j.autcon.2023.105187.
  14. Zinov V.I., Kartak V.M., Valiakhmetova Yu.I. [An algorithm for estimating the deformation of floor slabs based on the span-support topology of the building]. *Sistemy analiza i obrabotki dannyh – Analysis and data processing systems*. 2023. no. 4(92). pp. 35–54. DOI: 10.17212/2782-2001-2023-4-35-54. (In Russ.).
  15. Long Q., Wu C.Z., Wang X.Y., Jiang L., Li J. A multiobjective genetic algorithm based on a discrete selection procedure. *Mathematical Problems in Engineering*. 2015. vol. 2015. 17 p. DOI: 10.1155/2015/349781.
  16. Valiakhmetova Yu.I., Prokudina E.I., Zinov V.I. Mnogokriterial'nyj geneticheskij algoritm reshenija zadachi racional'nogo razmeshhenija opornyh konstrukcij na plane zdaniy [Multiobjective genetic algorithm for solution of support structure rational placement on the building plan problem] *Sovremennye fizika, matematika, cifrovyje i nanotehnologii v nauke i obrazovanii: Sb. nauchn. tr. II Vserossijskoj molodezhnoj*

- shkoly-konferencii [Modern physics, mathematics, digital and nanotechnologies in science and education: Collected research papers of II All-Russia youth school-conference]. Ufa: BSPU, 2023. pp. 95–100. (In Russ.).
17. Rahimi I., Gandomi A.H., Nikoo M.R., Chen F. A comparative study on evolutionary multi-objective algorithms for next release problem. *Applied Soft Computing*. 2023. vol. 144. DOI: 10.1016/j.asoc.2023.110472.
  18. Srinivas N., Deb K. Multiobjective optimization using nondominated sorting in genetic algorithms. *Evolutionary Computation*. 1994. vol. 2(3). pp. 221–248. DOI: 10.1162/evco.1994.2.3.221.
  19. Deb K., Pratap A., Agarwal S., Meyarivan T. A fast and elitist multiobjective genetic algorithm: NSGAII. *IEEE Transactions on Evolutionary Computation*. 2002. vol. 6(2). pp. 182–197. DOI: 10.1109/4235.996017.
  20. Zitzler E., Laumanns M., Thiele L. SPEA2: Improving the strength Pareto evolutionary algorithm. *TIK-Report*. 2001. vol. 103. DOI: 10.3929/ethz-a-004284029.
  21. Liu M., Zou X., Chen Y., Wu Z. Performance assessment of DMOEA-DD with CEC 2009 MOEA Competition Test Instances. *Proceeding of the IEEE Congress on Evolutionary Computation (CEC 2009)*. Norway: Trondheim, 2009. pp. 2913–2918. DOI: 10.1109/CEC.2009.4983309.
  22. He M., Wang X., Chen H., Li X. A Knee Point-Driven Many-Objective Evolutionary Algorithm with Adaptive Switching Mechanism. *Journal of Applied Mathematics*. 2024. vol. 2024(1). 43 p. DOI: 10.1155/2024/4737604.
  23. Ester M., Kriegel H.-P., Sander J., Xu X. A density-base algorithm for discovering clusters in large spatial databases with noise. *Proceeding of 2nd International Conference on Knowledge Discovery and Data Mining (KDD-96)*. Germany: Munich, 1996. pp. 226–231.

**Zinov Vladislav** — Postgraduate student, Department of technical cybernetics, Ufa University of Science and Technology. Research interests: optimization in the building design industry, heuristic and metaheuristic methods, decision support systems. The number of publications — 6. zinovvladislavufa@gmail.com; 32, Z. Validi St., 450076, Ufa, Russia; office phone: +7(917)435-5560.

**Kartak Vadim** — Ph.D., Dr.Sci., Professor, head of the department, Department of computing and information security, Ufa University of Science and Technology. Research interests: big data, optimization methods, information security. The number of publications — 80. kvmail@mail.ru; 32, Z. Validi St., 450076, Ufa, Russia; office phone: +7(917)433-3919.

**Valiakhmetova Yulia** — Ph.D., Associate professor of the department, Department of computational mathematics and engineering cybernetics, Ufa University of Science and Technology. Research interests: heuristic algorithms in discrete optimization problems, mathematical modeling, numerical analysis, software systems. The number of publications — 60. julikas@inbox.ru; 32, Z. Validi St., 450076, Ufa, Russia; office phone: +7(917)348-1947.

В.В. КУРЕЙЧИК, В.И. ДАНИЛЬЧЕНКО, Е.В. ДАНИЛЬЧЕНКО  
**МАРШРУТИЗАЦИЯ АВТОНОМНЫХ УСТРОЙСТВ  
В ТРЁХМЕРНОМ ПРОСТРАНСТВЕ**

*Курейчик В.В., Данильченко В.И., Данильченко Е.В. Маршрутизация автономных устройств в трёхмерном пространстве.*

**Аннотация.** Статья посвящена решению проблемы маршрутизации автономных устройств в трёхмерном пространстве, что является актуальной задачей в области интеллектуального управления. Трёхмерное пространство отличается высокой степенью свободы, сложной топологией и динамическими изменениями среды, что значительно усложняет задачу эффективного планирования траекторий. Разработка методов маршрутизации, обеспечивающих безопасность, энерго и вычислительную эффективность, имеет ключевое значение для повышения производительности автономных систем. В работе рассматривается комплексная система маршрутизации, основанная на гибридном подходе, объединяющем высокоуровневое моделирование рабочего пространства с метаэвристическими методами оптимизации. Для представления трёхмерной среды используются иерархические структуры данных, такие как октодеревья, что обеспечивает компактность и гибкость пространственных моделей. Эти модели преобразуются в графовые структуры, что позволяет описать маршрутизацию в виде оптимизационной задачи на графах. Предложен модифицированный метаэвристический муравьиный алгоритм, относящийся к классу роевых методов оптимизации. Алгоритм ориентирован на построение безопасных и энергоэффективных маршрутов, а также на решение задач поиска кратчайших гамельтоновых циклов и динамической перенастройки маршрута в условиях изменяющейся внешней среды. В работе представлены результаты вычислительного эксперимента, включающие тестирование алгоритма в условиях трёхмерного пространства, и сравнительный анализ с другими алгоритмами маршрутизации. Вычислительный эксперимент подтвердил эффективность разработанного алгоритма маршрутизации, включая сокращение времени вычислений и повышение энергоэффективности автономных устройств. Перспективы дальнейших исследований включают интеграцию предложенной системы в широкий спектр приложений для автономных устройств, направленных на оптимизацию процессов управления и повышение эффективности в динамически изменяющейся внешней среде. Отметим, что разработанный алгоритм может быть адаптирован для решения комплексных задач, в которых маршрутизация и размещение ветрогенераторов на плоскости взаимосвязаны. Задача размещения напрямую связана с построением маршрутов для обслуживания этих объектов, что требует комплексного подхода для эффективного решения этих задач. Это станет частью системы поддержки принятия решений, предназначенной для планирования и обслуживания ветрогенераторных комплексов, обеспечивая их эффективное функционирование и управление ресурсами.

**Ключевые слова:** метаэвристический алгоритм, муравьиная оптимизация, графовые математические модели, маршрутизация, моделирование трёхмерного пространства, энергетические системы.

**1. Введение.** Задача маршрутизации автономных устройств в трёхмерном пространстве является одной из ключевых проблем в области робототехники, автономных транспортных систем

и беспилотных летательных аппаратов. С развитием технологий автономных систем, увеличением вычислительных мощностей и повышением требований к безопасности и энергоэффективности, проблема эффективного планирования траекторий приобрела особую актуальность. В последние десятилетия для решения этой задачи применяются как классические алгоритмы, так и современные методы, такие как гибридные метаэвристические подходы, основанные на комбинировании различных стратегий поиска оптимальных решений [1 – 3]. Таким образом, задача маршрутизации автономных устройств в трехмерном пространстве является важным и актуальным направлением исследований, которое требует модифицированных подходов и методов для повышения качества и эффективности навигации в сложных и динамичных условиях.

**2. Анализ существующих решений.** Среди известных и широко применяемых методов маршрутизации выделяется алгоритм Дейкстры, предложенный Эдсгером Дейкстрой в 1956 году [1]. Он используется для поиска кратчайших путей в статичных графах и гарантирует нахождение оптимального маршрута. Отметим, что его вычислительная сложность возрастает с увеличением размера пространства, что ограничивает его эффективность для больших или динамично изменяющихся сред [4 – 6]. Алгоритм «A star» или «A\*», предложенный в 1968 году Патриком Хартом, является улучшенной версией алгоритма Дейкстры, так как использует эвристическую функцию для оценки проектного решения, что позволяет ускорить процесс поиска [1]. Но вместе с тем данный алгоритм имеет ограничения в работе с трехмерными пространствами, так как для его реализации требуется большие вычислительные затраты, то есть необходима квадратичная вычислительная сложность. В литературе представлено множество его модификаций, таких как Theta\*, Lifelong Planning A (LPA\*), D\* и его расширенная версия D\* Lite [1 – 8].

Среди множества модификаций алгоритма A\* особый интерес представляют методы, направленные на его адаптацию к динамическим средам. Одним из примеров таких алгоритмов является Theta\* [7], представляющий собой модификацию традиционного алгоритма A\*. Он позволяет строить квазиоптимальные маршруты, учитывая прямую видимость между узлами. Однако, несмотря на это преимущество, Theta\* требует значительных вычислительных ресурсов, особенно в сложных трехмерных средах [7, 8]. В отличие от него, алгоритм Lifelong Planning A\* (LPA\*) [8] предназначен для эффективного перерасчета маршрутов при изменении внешних условий. Тем не менее, LPA\* сохраняет структуру классического A\*,

что ограничивает его применение в динамических пространствах с большим количеством препятствий.

Для решения задачи маршрутизации разработаны алгоритмы семейства  $D^*$ . Оригинальный алгоритм  $D^*$  [7] обеспечивает адаптацию маршрута в реальном времени без необходимости полного пересчета маршрута, что повышает его эффективность для автономных систем. Улучшенная версия,  $D^*$  Lite [8], упрощает вычисления и ускоряет процесс обновления маршрута, сохраняя высокую точность. Однако оба метода ориентированы преимущественно на двухмерные пространства и имеют значительные вычислительные затраты при расширении на трехмерные среды, поскольку количество возможных направлений движения возрастает экспоненциально, что значительно увеличивает сложность поиска.

Альтернативные подходы к маршрутизации включают графовые методы, такие как дорожные карты и диаграммы Вороного, которые широко применяются для навигации в сложных пространствах. Одним из известных и эффективных методов является Probabilistic Roadmap (PRM) [9], строящий граф возможных маршрутов, при этом используя случайную дискретизацию пространства свободного движения. Такой метод эффективен в работе с предварительно известной средой (предварительно отсортированными данными), так как позволяет заранее сформировать карту маршрутов, что позволяет ускорить процесс поиска. Однако его эффективность значительно снижается в динамически изменяющихся условиях, где требуется частое перестроение графа, что увеличивает вычислительные затраты.

Другим популярным и эффективным методом является Rapidly-exploring Random Trees (RRT) [9], который выполняет исследование пространства с помощью случайных деревьев. Данный алгоритм применяется для построения маршрута в трехмерных пространствах с высокой сложностью конфигурации, поскольку он эффективно охватывает большие области и находит маршруты в условиях высокой неопределенности. Важное преимущество RRT заключается в его способности быстро адаптироваться к изменениям в среде. Однако базовая версия алгоритма не гарантирует нахождение оптимального маршрута, а в условиях динамически изменяющейся среды приводит к нестабильным решениям, требующим дополнительной оптимизации.

В таблице 1 приведены вычислительные сложности рассматриваемых алгоритмов маршрутизации, включая как классические методы, так и метаэвристические подходы, которые используются для навигации в динамических пространствах.

Таблица 1. Вычислительная сложность алгоритмов

Алгоритм	Вычислительная сложность	Примечание
A*	$O(n * \log h(n))$	$h(n)$ – эвристическая функция
Theta*	$O(n * \log n)$	Преимущество – учет прямой видимости между узлами
LPA*	$O(n * \log n)$	Эффективен для динамических изменений в графе
D*	$O(n * \log n)$	Подходит для динамических сред, пересчет маршрута в реальном времени
D* Lite	$O(n * \log n)$	Упрощенная версия D*, подходит для автономных систем
RRT	$O(n)$	Зависит от количества итераций, эффективен в высокоразмерных пространствах
PRM	$O(n * k)$	Может потребоваться повторный запуск алгоритма при неудаче в нахождении пути
ACO (Муравьиный алгоритм)	$O(n * m)$	Адаптивность и самоорганизация, подходит для динамических сред, эффективно находит квазиоптимальные маршруты
PSO (Метод роя частиц)	$O(n * m)$	Эффективен для нахождения квазиоптимальных решений за полиномиальное время, имитирует коллективное поведение агентов
GA (Генетические алгоритмы)	$O(n * m)$	Требует значительных вычислительных ресурсов для многокритериальных задач, основан на принципах естественного отбора

Несмотря на рассмотренные преимущества, перечисленные методы имеют ограничения, не позволяющие эффективно применять их для маршрутизации автономных устройств в трехмерном пространстве. Высокие вычислительные затраты, сложность обработки объемных данных и ограниченная адаптивность к быстро меняющимся условиям снижают их применимость в сложных сценариях. В связи с этим возникает необходимость разработки модифицированных эвристических подходов и методов, способных эффективно решать задачу маршрутизации в трехмерной динамической среде.

Одним из таких подходов для решения задачи маршрутизации является использование метаэвристических методов, которые реализуют различные стратегии поиска, тем самым повышая эффективность маршрутизации. Среди них метод на основе роя частиц (PSO) имитирующих коллективное поведение агентов, генетические алгоритмы (GA), основанные на принципах естественного отбора. Как известно данные методы позволяют эффективно находить квазиоптимальные решения за полиномиальное время в оптимизационных задачах. Однако, при решении многокритериальных задач, эти методы требуют значительных вычислительных ресурсов, что ограничивает их применимость в реальных условиях с динамически изменяющейся средой [1 – 4].

Одним из наиболее перспективных и широко применяемых методов в данной области является метаэвристический метод муравьиной оптимизации (ACO), который, как известно из литературных источников, позволяет находить кратчайшие маршруты используя коллективное поведение биологических агентов, обладающими такими свойствами, как автономность, коммуникабельность, адаптация и самоорганизация, присущими живым системам [1, 3]. Дополнительным преимуществом таких подходов является их способность динамически реагировать на изменение условий внешней среды. В частности, использование таких метаэвристических методов в задачах маршрутизации автономных устройств в трёхмерном пространстве позволяет значительно сократить время, затраченное на вычисления, и оптимизировать процесс планирования траекторий с учетом множества критериев, таких как скорость, потребление энергии, обход препятствий, безопасность и предотвращение столкновений.

В исследовании рассматривается модифицированный муравьиный алгоритм оптимизации, который, на основе проведенного анализа, продемонстрировал высокий потенциал для решения задачи маршрутизации в трехмерном пространстве. Этот метод обеспечивает эффективный баланс между вычислительной сложностью, адаптивностью к изменениям окружающей среды и качеством маршрутов, являясь особенно перспективным для навигации автономных систем в условиях динамичных и сложных сред.

**2. Постановка задачи.** Задача маршрутизации автономных устройств в трёхмерном пространстве заключается в поиске оптимального маршрута для обхода заданных контрольных точек. Это представляет собой адаптированную вариацию классической задачи коммивояжёра, но с учётом специфики трёхмерного пространства.

Главное отличие заключается в необходимости предварительной формализации задачи, что требует построения полной взвешенной графовой модели на основе трёхмерных данных. Формализация достигается путём вычисления кратчайших путей между контрольными точками и созданием матрицы весов, которая отражает стоимость перемещения между вершинами графа [1, 5 – 7].

Задачу маршрутизации предлагается разделить на два ключевых этапа. Первый этап включает построение математической модели в виде графа, основанной на трёхмерном пространстве с возможными препятствиями. Для этого предлагается упрощённая пространственная модель, которая сохраняет размерность исходного пространства, но с адаптированной сложностью геометрии для снижения вычислительной нагрузки. Разработка упрощённой модели трёхмерного пространства, должна учитывать геометрические особенности и возможные преграды. Модель пространства представим в виде графа  $G(X, U)$  (рисунк 1), где  $X$  – множество вершин (контрольных точек), а  $U$  – множество рёбер, соединяющих вершины.

Контрольные точки – это заранее определённые позиции в пространстве, через которые должен пройти маршрут автономного устройства. Они могут быть связаны с физическими объектами или быть абстрактными позициями, такими как места доставки или важные ориентиры на пути. Формирование контрольных точек происходит на основе анализа пространства с учётом геометрии, наличия препятствий и требований задачи. Контрольные точки выбираются так, чтобы обеспечивать оптимальный маршрут, минимизируя расстояние, время или затраты энергии, а также избегать столкновений с препятствиями.

В работе под динамической средой понимается система, в которой объекты и условия изменяются во времени. Эти изменения включают как перемещения объектов, так и изменения их характеристик или состояния, а также структуры среды, например, появления новых препятствий или исключение существующих. В динамической среде устройства или системы способны адаптироваться к этим изменениям в реальном времени, что требует высокой гибкости в планировании маршрутов и принятии решений. Ограничения на динамические объекты включают неопределённость в их поведении, трудности в прогнозировании будущих изменений и своевременной реакции на эти события [7 – 9].





$$W(i, j) = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2 + (z_i - z_j)^2} + P(i, j). \quad (2)$$

Функция штрафа  $P(i, j)$  для рёбер между вершинами  $i$  и  $j$  выражается как сумма двух критериев:

$$P(i, j) = P_{dist}(i, j) + P_o(i, j), \quad (3)$$

где  $P_{dist}(i, j)$  – штраф, зависящий от минимального расстояния до препятствия, который корректирует вес рёбер в графе в зависимости от коэффициента сложности  $K_{тип}$ :

$$P_{dist}(i, j) = \frac{K_{тип}}{dist_{min} + \epsilon}, \quad (4)$$

где  $dist_{min}$  – минимальное расстояние до препятствия,  $\epsilon$  – коэффициент для предотвращения деления на ноль. Если на траектории движения встречается препятствие, то требуется его обход, представляющий собой дополнительное увеличение веса рёбер, пропорциональное длине пути, необходимого для обхода препятствия, что определяется следующим образом:

$$P_o(i, j) = K_{тип} \cdot L_{обход}, \quad (5)$$

где  $L_{обход}$  – длина дополнительного пути.

Введение функции штрафов позволяет учесть наличие препятствий в маршруте и корректировать веса рёбер в графе в зависимости от их воздействия на движение. При наличии крупных объектов, таких как горы, здания или зоны запрета, алгоритм не только учитывает расстояние до препятствия, но и оценивает сложность его обхода. Отметим, что для динамической адаптации движения требуется проведение дополнительного анализа и применение методов, позволяющих своевременно корректировать маршрут в ответ на изменения внешней среды. Такой подход позволит обеспечить универсальность модели и её соответствие реальным условиям, в которых динамические препятствия могут существенно влиять на эффективность маршрута.

Далее производится построение гамильтонова цикла, то есть нахождения последовательности вершин, которая проходит через все

контрольные точки с возвратом в исходную точку, что выражается следующим образом:

$$C = \min_{\sigma \in S_n} \left( \sum_{i=1}^{n-1} W(v_{\sigma(i)}, v_{\sigma(i+1)}) + W(v_{\sigma(n)}, v_{\sigma(1)}) \right), \quad (6)$$

где  $\sigma$  – перестановка вершин графа,  $S_n$  – множество всех перестановок длины  $n$ , а  $v_{\sigma(1)}$  – вершина маршрута.

Построение общей длины маршрута  $C$  вычисляется в соответствии со следующим выражением [8]:

$$C = \sum_{i=1}^{n-1} W(v_i, v_{i+1}) + W(v_n, v_1), \quad (7)$$

где  $v_i$  – вершины маршрута входящие в маршрут, а  $n$  – общее количество вершин.

Для обеспечения адаптивности к динамическим изменениям условий внешней среды важным аспектом является уменьшение времени перерасчёта построенного маршрута. Это достигается с использованием муравьиной эвристики, которая позволяет эффективно обновлять маршрут в ответ на динамические изменения в среде. Для повышения вычислительной эффективности алгоритма и минимизации затрат времени перерасчёта маршрутов предлагается ввести ограничение на объём перерасчётов. Это ограничение заключается в обработке только тех участков графа, которые были изменены, что значительно снижает вычислительную нагрузку. Отметим, что это ограничение используется, когда в пространстве появляются новые препятствия или изменения в характеристиках существующих, что влияет только на определённые участки маршрута. Механизм перерасчёта маршрута с учётом изменений в графе определяется следующим образом:

$$C_{new} = \sum_{i=1}^k W(v_i, v_{i+1}) + \sum_{i=k+1}^n W(v_i, v_1), \quad (8)$$

где  $C_{new}$  – новый маршрут после перерасчёта,  $k$  – количество изменённых вершин графа.

Для повышения качества решений в модель также введен критерий энергоэффективности маршрута. Этот критерий позволяет учитывать не только кратчайшую длину маршрута, но и количество энергии, затраченное на его выполнение. Как известно из литературы [4 – 7] критерий энергоэффективности имеет одно из ключевых значений при планировании маршрутов для автономных устройств, особенно в условиях ограниченных энергетических ресурсов. Энергоэффективности маршрута вычисляется на основе следующего выражения:

$$E_{eff} = \frac{d_{opt}}{E_{spent}}, \quad (9)$$

где  $d_{opt}$  – длина кратчайшего маршрута,  $E_{spent}$  – энергия, затраченная на его выполнение.

Применение такого подхода позволяет сравнивать различные маршруты с учётом не только их длины, но и энергетических затрат, что помогает выбрать наиболее приемлемый маршрут с точки зрения баланса между длиной маршрута и затрачиваемой энергией на его прохождение. Отметим, что энергия, затраченная на выполнение маршрута, зависит от нескольких факторов, таких как сопротивление воздуха, скорости направления ветра, углы подъёма/спуска, скорость перемещения устройства и др. Для автономных устройств, уменьшение энергии, требуемой для выполнения маршрута, позволяет значительно увеличить время их работы. Тогда расчет энергии, требуемой на преодоление каждого сегмента маршрута, вычисляется на основе следующего выражения:

$$E_{segment} = \alpha \cdot d_{segment}^2 + \beta \cdot d_{segment} + \gamma, \quad (10)$$

где  $\alpha$ ,  $\beta$ , и  $\gamma$  – коэффициенты, учитывающие физические особенности движения. При этом энергия, затраченная на весь маршрут, вычисляется как сумма энергий, затраченных на прохождения каждого сегмента:

$$E_{spent} = \sum_{i=1}^{n-1} E_{segment}(v_i, v_{i+1}), \quad (11)$$

где  $E_{segment}$  – энергия, необходимая для преодоления пути между вершинами  $v_i$  и  $v_{i+1}$ . Таким образом, учет энергоэффективности в задаче маршрутизации позволяет повысить эксплуатационные характеристики автономных устройств, обеспечивая баланс между длиной маршрута и продолжительностью работы устройства. Отметим, что при применении данного подхода к кусочно-непрерывной траектории возникают разрывы в скорости движения автономного устройства и нарушается физическая реализуемость маршрута. Для улучшения маршрутизации автономных устройств в сложных средах предлагается модификация алгоритма маршрутизации, включающая процедуру сглаживания траектории, которая минимизирует острые углы и избыточные точки, повышая точность и энергоэффективность при сохранении баланса между оптимизацией пути и его точностью. Что позволяет устранить разрывы в скорости и обеспечить более реалистичное моделирование движения устройства.

Данная модель позволяет учесть не только длину маршрута, но и его сложность, обусловленную особенностями среды и динамическими характеристиками устройства. Важно отметить, что при моделировании маршрутов необходимо учитывать дополнительные энергозатраты, связанные с маневрами устройства, такими как торможение и разгон, что особенно важно при резких поворотах и изменениях направления движения. Для учёта этих факторов можно дополнительно ввести компоненту, отражающую энергозатраты, возникающие при маневрах:

$$E_{segment} = \alpha \cdot d_{segment}^2 + \beta \cdot d_{segment} + \gamma + \delta f(\theta), \quad (12)$$

где  $f(\theta)$  – функция, описывающая дополнительные энергозатраты, вызванные маневрами, а  $\delta$  – коэффициент, учитывающий влияние этих затрат [8 – 10].

Полиномиальная модель энергозатрат учитывает различные аспекты движения устройства. Квадратичная компонента  $\alpha \cdot d_{segment}^2$  отражает увеличение энергозатрат с ростом расстояния, включая аэродинамическое сопротивление и затраты на разгон или торможение. Линейная зависимость  $\beta \cdot d_{segment}$  моделирует постоянные потери энергии, такие как трение и механические потери. Константная составляющая  $\gamma$  учитывает фиксированные энергозатраты на работу сенсоров или систем управления устройства.

При этом автономное устройство в данной задаче рассматривается как объект с тремя степенями свободы, обладающий определёнными формой, размерами и весом, а также точками приложения силы. Эти характеристики учитываются при построении маршрута, что позволяет обеспечить точное соответствие модели реальным условиям эксплуатации и учитывать факторы, влияющие на маневренность, энергоэффективность и безопасность движения устройства в трёхмерном пространстве.

### **3. Математическая модель трёхмерного пространства.**

Проблема эффективной обработки больших трёхмерных пространств с множеством объектов является одной из ключевых в области робототехники, компьютерных наук и геоинформационных систем. В рамках этой задачи большое внимание уделяется разработке методов для упрощения представления пространства и ускорения процессов обработки. Одним из наиболее распространённых решений является использование древовидных структур данных, таких как октодеревья и  $k$ -мерные деревья, которые широко применяются в таких областях, как компьютерная графика, геодезия и других [7 – 9].

Октодеревья представляют собой одну из наиболее эффективных структур для представления трёхмерного пространства [3, 9]. В отличие от  $k$ -мерных деревьев, которые являются бинарными и делят пространство на два подмножества, октодерево делит каждую ячейку на восемь равных частей, что позволяет более эффективно обрабатывать данные в трёхмерной среде. Это свойство делает октодеревья особенно подходящими для задач, связанных с анализом соседства, пространственной индексацией, а также для поиска коллизий между объектами. В своей основе октодерево является структурой, способной динамически адаптировать уровень детализации пространства, что позволяет эффективно управлять объёмами данных и минимизировать вычислительные затраты. Отметим, что октодеревья демонстрируют высокую эффективность при выполнении задач поиска ближайших соседей в трёхмерной среде. Иерархическая структура, которая последовательно делит пространство, позволяет быстро исключать области, не содержащие искомым объектов, существенно сокращая объём вычислений. Такое преимущество особенно важно при решении задач в масштабах реального времени, таких как маршрутизация автономных устройств. Также, универсальность октодеревьев заключается в их способности работать как со статичными, так и с динамически изменяемыми сценариями маршрутизации.

При этом модель октодерва в трёхмерном пространстве имеет следующие ограничения.

1. Зависимость производительности от уровня детализации (глубины дерева). С увеличением глубины дерева растёт количество узлов, что повышает вычислительные затраты. Количество ячеек на каждом уровне дерева пропорционально  $8^d$ , где  $d$  – глубина дерева. Следовательно, производительность системы будет зависеть от точности, требуемой для представления объектов. Для поиска ближайших соседей или коллизий, например, число вычислений будет возрастать экспоненциально с увеличением глубины поиска. Тогда количество ячеек в дереве при глубине  $d$  определяется следующим образом:

$$N_{cells} = 8^d, \quad (13)$$

где  $N_{cells}$  – количество ячеек дерева. Заметим, если глубина октодерва превышает заданный порог, это приводит к значительным затратам памяти и времени, так как необходимо хранить и обрабатывать большое количество данных. При этом глубина октодерва определяет степень детализации модели: с увеличением числа уровней повышается точность представления структуры пространства, однако это ведёт к увеличению вычислительных нагрузок [3, 10]. При выборе оптимальной глубины октодерва необходимо учитывать несколько факторов. Для объектов, занимающих большую часть рабочего пространства, достаточно использовать октодерво минимальной глубины, что позволяет эффективно представлять их положение и взаимодействие без избыточного деления ячеек, тем самым сокращая объём вычислений. Для учета мелких объектов необходимо повышать детализацию, так как размеры ячеек на верхних уровнях октодерва могут быть слишком большими для их точного представления. При решении задач в динамически изменяющихся средах, оптимальная глубина дерева позволяет минимизировать объём перерасчётов при сохранении точности модели. Приемлемая глубина октодерва рассчитывается на основе следующего выражения:

$$d = \left\lceil \log_2 \frac{L}{L_{min}} \right\rceil, \quad (14)$$

где  $L$  – размер пространства, а  $L_{min}$  – минимальный размер объекта.

2. Управление памятью при работе с большими пространствами. При работе с большими пространствами и динамически изменяющимися сценариями, управление памятью становится актуальной задачей. В случае изменения положения объектов в рабочем пространстве, необходимо обновлять только те ветви дерева, которые затронуты, что существенно сокращает вычислительные затраты и увеличивает производительность системы.

Опишем процесс динамического обновления октодерева, где каждый узел может быть «свободным» или «занятым». Пусть  $T$  – это октодерево,  $T(i)$  – это  $i$  узел этого дерева,  $O_i$  – объект, который может занимать или не занимать пространство в ячейке  $T(i)$ . Тогда состояние узла  $T(i)$  в определённый момент времени будет определяться следующим образом [8 – 10]:

$$T(i) = \begin{cases} \text{занят,} & \text{если } O_i \text{ пересекает или находится в } T(i) \\ \text{свободен,} & \text{если } O_i \text{ не пересекает } T(i). \end{cases} \quad (15)$$

Алгоритм определения множества узлов, которые изменили своё состояние, включает следующие этапы:

1. для каждого объекта  $O_i$ , который изменил своё положение, вычисляется его новое пространственное расположение;
2. выполняется пересечение нового объёма объекта с существующей структурой октодерева;
3. формируется множество  $a_n$ , состоящее из всех узлов  $T(i)$ , которые пересекаются с новым положением объекта  $O_i$ , но не пересекались с его предыдущим положением;
4. аналогично, в  $a_n$  добавляются узлы, которые перестали пересекаться с объектом после его перемещения;
5. только узлы из множества  $a_n$  подвергаются обновлению состояния, что значительно снижает вычислительные затраты.

Такой подход позволяет минимизировать обновления структуры, оптимизируя вычислительные ресурсы при динамическом изменении сцены [10, 11].

При перемещении объект  $O_i$  необходимо производить пересчет только тех узлов, которые изменили свое состояние, в противном случае обновление состояния будет происходить для всех узлов. Данный процесс обновления состояния узлов дерева определяется следующим образом:

$$\forall i \in a_n \quad T(i) = \text{update}(O_i), \quad (16)$$



где  $a_n$  – это множество узлов, которые пересекаются с перемещаемым объектом. Это значительно снижает вычислительные затраты, особенно в случае динамических сцен с частыми изменениями положения объектов в пространстве за счет обновления только части используемого дерева.

Такой механизм обновления узлов деревьев в динамических сценариях помогает поддерживать высокую производительность системы маршрутизации и минимизировать ресурсоёмкость вычислений, что критически важно для автономных устройств, работающих в сложных и изменяющихся внешних условиях.

**4. Построение модели маршрутизации.** Для решения задачи маршрутизации автономных устройств в трёхмерном пространстве, моделируемом в виде графа, применяется эвристический алгоритм  $A^*$  (A-star), который является модификацией алгоритма Дейкстры [1, 8–11]. Этот подход позволяет быстро находить оптимальные пути в графах, с учетом только прямолинейных маршрутов, однако его эффективность снижается, когда необходимо учитывать сложные траектории, в том числе те, которые имеют форму зигзагообразных путей, характерных для навигации в реальных или ограниченных пространствах. В условиях, когда движущийся объект должен преодолевать пространство с препятствиями или работать в ограниченных и сильно структурированных средах, прямолинейные пути могут быть невозможными, и возникает необходимость в более гибком планировании маршрута [11]. Алгоритм  $A^*$  строит зигзагообразную траекторию, поскольку он стремится найти наиболее короткие и прямые отрезки, что определяется следующим выражением:

$$f(n) = g(n) + h(n), \quad (17)$$

где  $g(n)$  – длина пути до текущей точки  $n$ , а  $h(n)$  – оценка расстояния до цели. Тогда траектория состоит из множества коротких отрезков  $d(P_i, P_{i+1})$ , сумма которых задаёт общую длину и вычисляется следующим образом:

$$L_t = \sum_{i=1}^{N-1} d(P_i, P_{i+1}), \quad (18)$$

Дискретизация пространства приводит к формированию зигзагообразных маршрутов, так как фиксированные узлы

ограничивают выбор точек, соединяя их по принципу ближайшего соседства, образуя острые углы между последовательными отрезками. Они вычисляются на основе следующего выражения:

$$\cos(\theta_i) = \frac{(P_i - P_{i-1}) \cdot (P_{i+1} - P_i)}{\|P_i - P_{i-1}\| \cdot \|P_{i+1} - P_i\|}, \quad (19)$$

Такой маршрут, построенный с помощью алгоритма А\* приведен на рисунке 2.

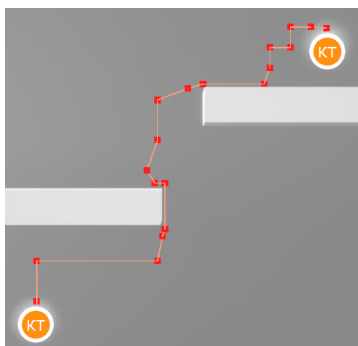


Рис. 2. Построение маршрута на основе алгоритма А\*

Эти «изгибы» увеличивают суммарную длину маршрута, снижают точность, а также повышают вычислительные затраты и время движения автономных устройств, за счет выполнения дополнительных маневров, что отрицательно влияет на общую эффективность, особенно в контексте ограниченных ресурсов.

Для уменьшения числа острых углов применяется процедура сглаживания траектории, что достигается путём оптимизации отрезков пути, исключая избыточные точки. Итоговое решение (длина маршрута) включает сглаживание поворотов, что определяется следующим выражением:

$$L_s = \sum_{i=1}^{N-1} \|P_i - P_{i-1}\| + \beta \sum_{i=2}^{N-1} (1 - \cos(\theta_i)), \quad (20)$$

где первый компонент минимизирует длину, а второй сглаживает углы между отрезками.

С учетом этого в работе предлагается модифицированный алгоритм маршрутизации, реализующий процедуру сглаживания маршрута, за счет уменьшения количество избыточных точек на пути. При этом устраняются ненужные колебания и ошибки, что способствует оптимизации маршрута, повышению его точности, сокращению вычислительных затрат, а также снижается общая длина пути, уменьшается время обхода контрольных точек, что эффективно сказывается при обработке в сложных и динамически изменяющихся средах.

Процесс сглаживания траектории в алгоритмах маршрутизации, несмотря на повышение энергоэффективности автономного устройства, обладает рядом недостатков, которые необходимо учитывать при его реализации. Одним из ключевых недостатков является снижение точности маршрута: сглаживание упрощает траекторию, устраняя острые углы и зигзаги, что приводит к отклонению маршрута от заданных контрольных точек. Это критично в ситуациях, когда необходимо точное соблюдение траектории, например, для обхода препятствий или при движении в ограниченных пространствах. Для устранения данных недостатков в работе предлагается использовать адаптивные методы сглаживания.

Предложенный алгоритм отличается от существующих решений тем, что использует адаптивное сглаживание маршрута, динамически изменяя степень удаления промежуточных точек в зависимости от структуры окружающей среды и требований к точности движения. В отличие от стандартных методов, где точки удаляются по фиксированному критерию (например, углу между отрезками или расстоянию до соседних узлов), в предложенном алгоритме вводится весовая функция, учитывающая плотность контрольных точек, наличие препятствий и необходимую точность навигации. Что позволяет сохранять критически важные точки, обеспечивая надежность маршрута, и одновременно исключать избыточные узлы, сокращая вычислительные затраты.

Ключевой модификацией является введение механизма приоритетного сохранения контрольных точек, основанный на оценке значимости узлов. Для этого используются параметры: минимальная дистанция между сохранёнными точками, степень изменения угла маршрута и наличие внешних ограничений (стены, узкие коридоры, динамические препятствия). В сравнении с традиционными методами сглаживания, которые с одной стороны упрощают маршрут, приводя к его отклонению от заданных контрольных точек, с другой не обеспечивают достаточную эффективность в устранении избыточных

узлов. Данный метод актуален для автономных устройств, функционирующих в условиях сложных динамически изменяющихся сред, где требуется учитывать вариативность топологии и пространственные ограничения.

Реализация модифицированного алгоритма построения матрицы на основе контрольных точек и поиска оптимальных путей между ними с процедурой сглаживания реализован на Псевдокоде 1.

Алгоритм построения матрицы маршрутов

Входные данные: P - множество контрольных точек  $\{P_1, P_2, \dots, P_N\}$ , где каждая точка имеет координаты в 3D-пространстве.

Выходные данные: M - матрица маршрутов, где  $M[i][j]$  содержит путь и его длину между  $P_i$  и  $P_j$ .

Шаг 1. Инициализация:

Создать пустую матрицу маршрутов M размером  $N \times N$ .

Для всех  $i, j$ : установить  $M[i][j]$  как пустой маршрут.

Шаг 2. Формирование маршрутов:

Для каждого  $i$  от 1 до N:

Для каждого  $j$  от  $i+1$  до N:

маршрут = Найти маршрут ( $P_i, P_j$ )

длина = Рассчитать длину (маршрут)

$M[i][j]$  = (маршрут, длина)

$M[j][i]$  = Инвертировать (маршрут, длина)

Шаг 3. Поиск маршрута ( $P_{start}, P_{end}$ ):

Если существует прямой путь между  $P_{start}$  и  $P_{end}$ :

вернуть  $\{P_{start}, P_{end}\}$

Иначе:

Использовать алгоритм поиска пути (например, A\*) для нахождения оптимального маршрута.

вернуть список точек маршрута.

Шаг 4. Вычисление длины маршрута (маршрут):

длина = 0

Для каждого  $i$  от 1 до длины (маршрут) - 1:

длина += Расстояние(маршрут[i], маршрут[i+1])

вернуть длина

Шаг 5. Инверсия маршрута (маршрут, длина):

инвертированный\_маршрут = Обратить последовательность (маршрут)

вернуть (инвертированный\_маршрут, длина)

Шаг 6. Оптимизация маршрутов:

Для каждого маршрута в M:

удалить избыточные точки, не влияющие на траекторию;

минимизировать количество резких поворотов;

обновить длину маршрута.

Шаг 7. Выход:

Вернуть матрицу маршрутов M.

Псевдокод 1. Реализация модуля построение матрицы маршрутов

Рассмотрим работу этого алгоритма по шагам.

- Инициализация. Создается пустая матрица для хранения информации о маршрутах между всеми возможными парами точек. Каждая ячейка матрицы содержит данные о маршруте, включая последовательность точек и его длину. Что описывается на основе формулы 1.

- Получение входных данных. Система получает набор контрольных точек, которые необходимо связать. Эти точки представляют собой координаты в трёхмерном пространстве.

- Поиск маршрутов. Для каждой пары точек выполняется соединение. Это может быть прямой путь или более сложная траектория, в зависимости от препятствий или ограничений.

- Сохранение маршрутов. После нахождения соединения между двумя точками, данные о маршруте (последовательность точек на пути и длина маршрута) сохраняются в соответствующей ячейке матрицы.

- Обратные маршруты. Сохраняются инвертированные маршруты (обратные пути), что позволяет использовать симметричные данные, и при необходимости вернуться из конечной точки в начальную.

- Формирование матрицы решений. Заполняется итоговая матрица маршрутов, которая содержит всю необходимую информацию о соединениях между точками. Эта матрица используется для дальнейшего анализа, оптимизации или планирования траекторий в контексте навигации автономных устройств.

На рисунке 3 представлен маршрут после реализации механизма сглаживания, который уменьшает количество острых углов и зигзагов оптимизируя траекторию. Это позволяет сформировать эффективный маршрут, сокращая время движения и повышая точность при обходе препятствий.

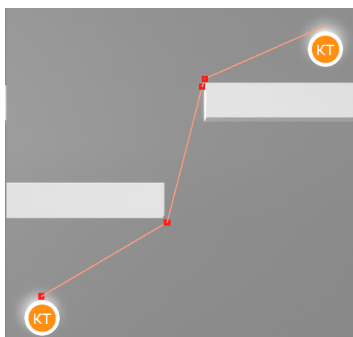


Рис. 3. Маршрут после применения механизма сглаживания

Таким образом, применение механизма сглаживания позволяет улучшить характеристики маршрута, обеспечивая более плавное и эффективное движение автономных устройств, что особенно важно в условиях сложной топологии и динамически изменяющейся сред.

**5. Модифицированный муравьиный алгоритм.** Для реализации всей маршрутизации в работе перелagается модифицированный метаэвристический муравьиный алгоритм. Алгоритм маршрутизации основывается на принципах поведения муравьиной колонии, которая при поиске пищи оставляет феромоны на рёбрах графа, тем самым формируя различные возможные маршруты. По мере того, как феромоны испаряются, более длинные и менее эффективные пути становятся менее выраженными, в то время как более короткие маршруты, наоборот, усиливаются, так как на них остаётся больше феромона [12 – 15]. Этот процесс позволяет моделировать поиск оптимального маршрута, аналогично поведению муравьёв в естественной среде [14].

Как известно одной из проблем при реализации муравьиных алгоритмов является большое количество параметров с широкими диапазонами значений, что усложняет настройку и адаптацию алгоритма для конкретных задач [15 – 17].

Для решения этой проблемы предлагается следующая модификация, заключающаяся в реализации двух механизмов. Это динамическая регулировка количества феромона, оставляемого муравьями, зависящая от качества пройденного маршрута, что позволяет избежать чрезмерного накопления феромона на менее удачных путях и способствует более равномерному исследованию решений. При этом процесс обновления количества феромона на ребре  $e_{ij}$  вычисляется на основе следующего выражения [1, 17]:

$$\tau_{ij}(t + 1) = (1 - \rho) \cdot \tau_{ij}(t) + \Delta \tau_{ij}(t), \quad (21)$$

где  $\tau_{ij}(t)$  – количество феромона на ребре  $e_{ij}$  в момент времени  $t$ ,  $\rho$  – коэффициент испарения феромона, а  $\Delta \tau_{ij}(t)$  – изменение феромона, которое зависит от пути, выбранного муравья в текущем цикле.

В отличие от классического подхода, в данном алгоритме учитывается адаптивное изменение феромона в зависимости от качества маршрута, что позволяет избежать ранней фиксации на локальных решениях.

Дополнительно используется механизм введения «элитных муравьёв», которые усиливают феромонные следы на наиболее перспективных маршрутах, ускоряя процесс сходимости и направляя

агентов к оптимальным траекториям. Эти изменения устраняют недостатки классического алгоритма, такие как медленная сходимость и склонность к застреванию в локальных минимумах [1].

Элитные муравьи оставляют больше феромона на наиболее перспективных маршрутах. Обновление феромона для элитных муравьёв описывается следующим образом:

$$\tau_{ij}^{ilite}(t+1) = (1 - \rho) \cdot \tau_{ij}(t) + \Delta \tau_{ij}^{ilite}(t), \quad (22)$$

где  $\Delta \tau_{ij}^{ilite}(t)$  – это дополнительное количество феромона, оставляемое элитными муравьями. Этот параметр задается как более высокое значение по отношению к  $\Delta \tau_{ij}(t)$  для обычных муравьёв:

$$\tau_{ij}^{ilite}(t+1) = \varepsilon \cdot \Delta \tau_{ij}(t), \quad (23)$$

где  $\varepsilon > 1$  – коэффициент усиления, который задает уровень феромона элитных муравьёв.

Предложенный механизм элитных муравьёв отличается от стандартного тем, что использует динамическую настройку коэффициента усиления  $\varepsilon$  в зависимости от текущего состояния поиска. Что позволяет достичь баланса между усилением лучших маршрутов и сохранением разнообразия решений. В классическом алгоритме муравьиного поиска обновление феромона осуществляется равномерно для всех муравьёв, что способствует преждевременной фиксации на неоптимальных маршрутах. В предложенной модификации данный процесс регулируется динамически с учетом эффективности пройденного пути, что повышает устойчивость поиска и ускоряет его сходимость. Кроме того, в отличие от стандартного подхода, предложенный алгоритм адаптивно регулирует влияние элитных муравьёв посредством коэффициента  $\varepsilon$ , предотвращая избыточную концентрацию феромона на отдельных маршрутах и обеспечивая сбалансированное исследование пространства решений.

Данный механизм способствует ускорению сходимости к оптимальным траекториям, снижая вероятность попадания в локальные «ямы». Он фокусирует внимание на наиболее эффективных путях, увеличивая интенсивность феромонного следа, что повышает вероятность выбора этих маршрутов на последующих итерациях. Таким образом, введение элитных муравьёв позволяет уменьшить влияние случайности и ускоряет процесс нахождения эффективного маршрута, устраняя недостатки классического муравьиного алгоритма.

Также элитные муравьи не только увеличивают интенсивность выделения феромонов, но и активно влияют на баланс глобального и локального поиска, адаптируя свою стратегию в зависимости от предыдущих итераций, что делает процесс нахождения оптимального маршрута более устойчивым и эффективным.

Модифицированный муравьиный алгоритм, в котором реализованы механизмы динамического обновления феромонов и использования элитных муравьёв представлен в виде Псевдокода 2.

Шаг 1. Инициализация параметров: задаём основные параметры алгоритма, включая количество муравьёв, число итераций, коэффициент испарения феромонов и влияния эвристики.

```
num_ants = 50
num_iterations = 100
evaporation_rate = 0.5
alpha = 1.0
beta = 2.0
elite_factor = 2.0
pheromones = [[0.1 for _ in range(len(M))] for _ in range(len(M))]
```

Шаг 2. Построение маршрутов: каждый муравей строит маршрут, выбирая следующий узел с учётом феромонов и эвристики.

```
def build_route():
    route, visited = [], set()
    current = random.randint(0, len(M) - 1)
    route.append(current)
    visited.add(current)
    while len(visited) < len(M):
        next_node = select_next_node(current, visited)
        route.append(next_node)
        visited.add(next_node)
        current = next_node
    return route
def select_next_node(current, visited):
    probabilities = [pheromones[current][neighbor]**alpha * (1/M[current][neighbor])**beta
                    for neighbor in range(len(M)) if neighbor not in visited]
    nodes = [neighbor for neighbor in range(len(M)) if neighbor not in visited]
    return random.choices(nodes, weights=probabilities)[0]
```

Шаг 3. Обновление феромонов: феромоны обновляются с учётом всех маршрутов и более сильным воздействием элитных муравьёв.

```
def update_pheromones(routes, best_route):
    for i in range(len(M)):
        for j in range(len(M)):
            pheromones[i][j] *= (1 - evaporation_rate)
    for route in routes:
        length = calculate_route_length(route)
        pheromone_value = 1.0 / length
        for i in range(len(route) - 1):
```



```

a, b = route[i], route[i + 1]
pheromones[a][b] += pheromone_value
pheromones[b][a] += pheromone_value
if best_route:
    elite_pheromone = elite_factor / calculate_route_length(best_route)
    for i in range(len(best_route) - 1):
        a, b = best_route[i], best_route[i + 1]
        pheromones[a][b] += elite_pheromone
        pheromones[b][a] += elite_pheromone
Шаг 4. Обновление феромонов: феромоны обновляются с учётом всех маршрутов и
более сильным воздействием элитных муравьёв.
best_route, best_length = None, float("inf")
for iteration in range(num_iterations):
    routes = [build_route() for _ in range(num_ants)]
    for route in routes:
        length = calculate_route_length(route)
        if length < best_length:
            best_length = length
            best_route = route
update_pheromones(routes, best_route)
Шаг 5. Результат: выводим лучший маршрут и его длину.
print("Лучший маршрут:", best_route)
print("Длина маршрута:", best_length)
Псевдокод 2. Реализация модифицированного муравьиного алгоритма

```

Рассмотрим работу модифицированного муравьиного алгоритма по шагам.

- Инициализация параметров. На первом этапе задаются начальные параметры алгоритма, включая количество муравьёв, коэффициенты испарения феромона и максимальное число итераций. Также инициализируются данные об уровне феромонах на рёбрах графа.

- Построение и обновление уровня феромонов. Количество феромона на рёбрах зависит от качества маршрута. Больше феромона остаётся на коротких, эффективных путях, а менее эффективных путях происходит испарение феромонов с течением времени.

- Использование «элитных муравьёв». Для ускорения сходимости алгоритма и предотвращения попадания в локальные «ямы» вводится механизм элитных муравьёв. Эти муравьи оставляют большее количество феромона на наиболее перспективных маршрутах, что позволяет быстрее направить поиск к искомому решению. Это улучшает скорость сходимости и направляет муравьёв к более эффективным траекториям.

- Основной цикл и итерации. Выполняется основной цикл, в котором муравьи начинают своё движение от стартовой точки,

выбирая пути на основе вероятностей, пропорциональных количеству феромона на рёбрах, что вычисляется на основе следящей формулы [1, 11]:

$$P_{ij} = \frac{\tau_{ij}^{\alpha} \cdot \mu_{ij}^{\beta}}{\sum_{k \in N_i} \tau_{ij}^{\alpha} \cdot \mu_{ij}^{\beta}}, \quad (24)$$

где  $\tau_{ij}^{\alpha}$  – количество феромона на ребре  $e_{ij}$ ,  $\mu_{ij}^{\beta}$  – эвристическая информация (например, обратная длина пути),  $N_i$  – множество соседей текущего узла, а  $\alpha$  и  $\beta$  – параметры, определяющие количество феромона и эвристической информации.

– Постобработка и обновление количества феромонов. После завершения всех итераций, выполняется обновление феромонов на рёбрах, по которым прошли муравьи, с учётом элитных муравьёв. Обновление количества феромона позволяет зафиксировать наилучшие пути [1, 11].

– Оценка и завершение работы. Продолжается процесс до достижения заданного числа итераций или критерия останова. В конце работы выбирается оптимальный маршрут.

Модифицированный муравьиный алгоритм реализует механизмы динамического обновления феромонов и использования элитных муравьёв для поиска оптимальных маршрутов в динамических и сложных средах. Данный алгоритм повышает эффективность поиска квазиоптимальных решений за счет быстрой и устойчивой сходимости.

Таким образом, модифицированный муравьиный алгоритм не только использует элитных муравьёв и динамическое обновление феромонов, но и обеспечивает баланс между локальным и глобальным поиском, предотвращая преждевременную сходимость алгоритма.

**6. Компьютерное моделирование.** Данный подход к маршрутизации автономных устройств в трёхмерном пространстве был реализован в среде компьютерного моделирования Unreal Engine 5 (UE 5) [17–19]. На рисунке 4 приведено октодереве рабочего пространства, представленного в виде подразделяющейся кубической структуры.

Здесь в качестве модели используется ветрогенераторный парк, состояние которого анализируется автономным устройством, управляемым предложенным методом. Этот подход иллюстрирует одну из потенциальных областей применения разработанной системы.



Рис. 4. Октодерево рабочего пространства

На рисунке 5 изображен полный граф всех возможных маршрутов.

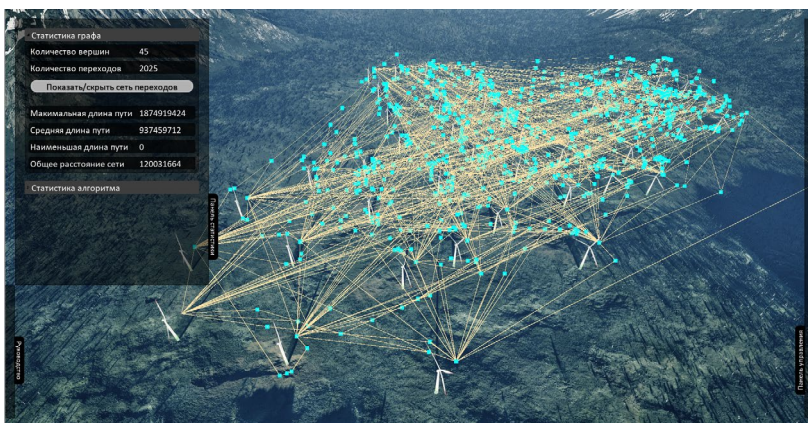


Рис. 5. Полный граф всех возможных маршрутов

Отметим, что вблизи объектов, структура дерева делится на максимальное количество уровней, что позволяет более точно моделировать сложные геометрические особенности и препятствия моделируемого ландшафта.

На рисунке 6 показан результат реализации модифицированного муравьиного алгоритма при построении кратчайшего маршрута, интерпретированного в виде точек в пространстве.

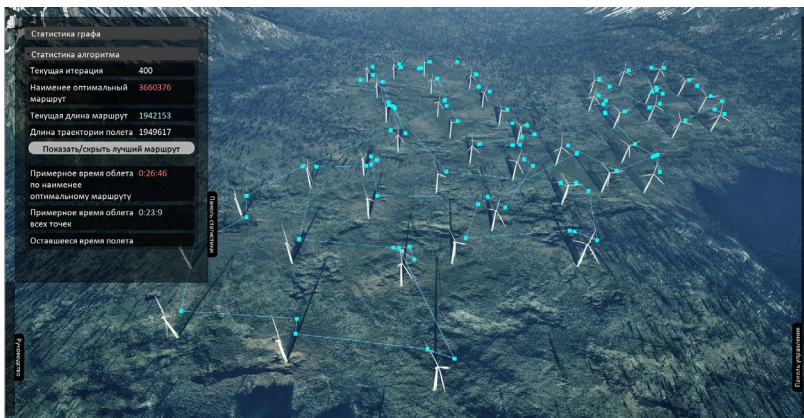


Рис. 6. Кратчайший маршрут (квазиоптимальное решение)

В результате проведенного компьютерного моделирования и предложенного подхода маршрутизации автономных устройств в 3D-пространстве с использованием Unreal Engine 5 продемонстрировано высокое качество реализации и производительности. Моделирование в виртуальной среде, включая использование октодеревя и реализации модифицированных алгоритмов, показало точность в обработке сложных геометрических объектов при построении квазиоптимальных траекторий, что подтверждает эффективность предложенного подхода при решении задачи маршрутизации.

**7. Вычислительный эксперимент.** Для подтверждения эффективности и качества предложенного подхода, проведена серия экспериментальных исследований. В качестве аналога был рассмотрен классический муравьиный алгоритм.

По результатам предварительного анализа исходного графа маршрутизации было найдено эталонное значение целевой функции, которое составило 1,500,000 (в единицах Unreal Engine, где одна UE единица соответствует одному сантиметру) [17 – 20].

Для экспериментальной проверки предложенного подхода использовались сгенерированные случайные графы. Проведена серия тестов с целью оценки изменения параметров сходимости и качества маршрутов.

Оценка качества маршрутов  $Q$  играет ключевую роль в оценке эффективности алгоритма маршрутизации, поскольку она позволяет определить, насколько найденный маршрут приближен к оптимальному решению. Что рассчитывается по следующей формуле [1, 19]:

$$Q = \frac{D_{opt} - D_{current}}{D_{opt}} \cdot 100\%, \quad (25)$$

где  $D_{opt}$  – эталонное значение целевой функции, а  $D_{current}$  – текущий результат, полученный с использованием алгоритма.

В качестве базового использовался классический муравьиный алгоритм с настройками, включающими следующие параметры: количество муравьёв (50), число итераций (100), коэффициент испарения феромонов (0.5), влияние феромонов ( $\alpha = 1.0$ ) и влияние эвристики ( $\beta = 2.0$ ) [1]. Эти параметры были выбраны на основе предварительных экспериментов, направленных на балансировку качества решения и вычислительных затрат.

Результаты экспериментов, представленные на графике (рисунок 7), демонстрируют зависимость качества решения от времени выполнения различных алгоритмов. Что позволяет проанализировать эффективность предложенного подхода и сравнить его с классическими методами маршрутизации.

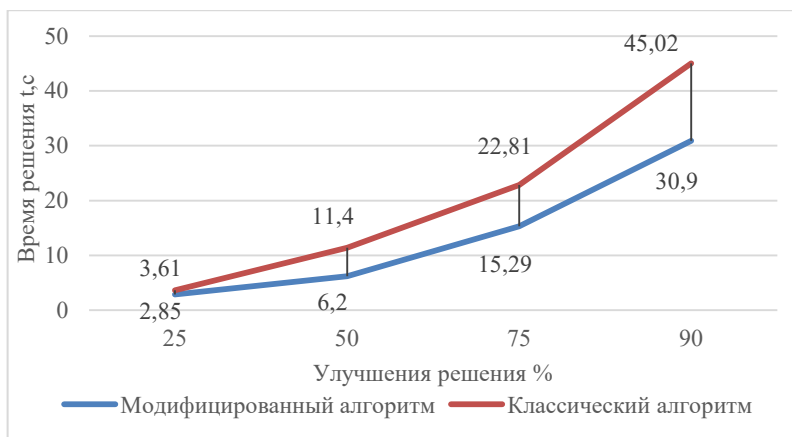


Рис. 7. График зависимости качества решения от времени выполнения алгоритмов

На рисунке 8 представлена зависимость качества решения от количества вершин графа (ветрогенераторов). График демонстрирует, как увеличивается эффективность алгоритмов с ростом числа объектов, что подчеркивает улучшение качества решения при расширении задачи.

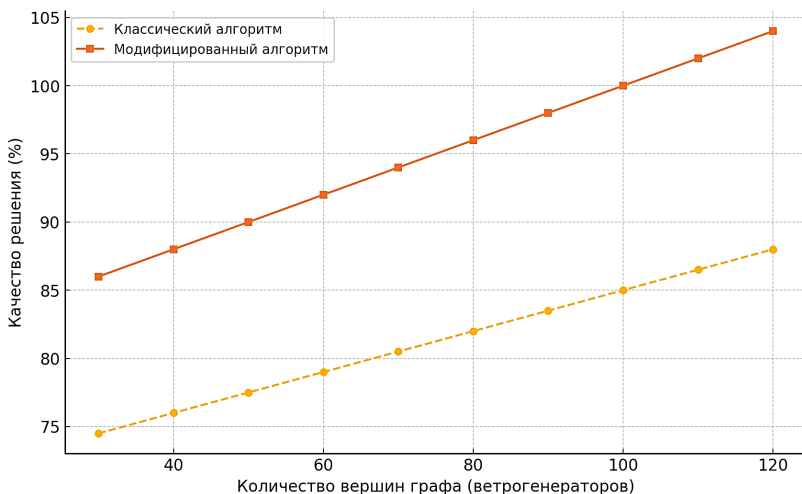


Рис. 8. График зависимости качества решения от количества вершин графа (ветрогенераторов)

При увеличении числа вершин в графе растёт и вычислительная сложность задачи. Однако применение модификаций муравьиного алгоритма, таких как динамическая корректировка уровня феромонов и использование элитных муравьёв, позволяет сохранять высокую эффективность и адаптироваться к динамическим изменениям в пространстве поиска, что позволяет улучшить результаты маршрутизации в условиях динамических и сложных средах.

Модифицированный муравьиный алгоритм продемонстрировал лучшие результаты, достигая 95% качества решений при 120 вершинах, в то время как классический алгоритм ограничивается на 90%. Предложенная модификация алгоритма, позволяет улучшить эффективность поиска, сокращая время для достижения 90% качества с 45,02 до 30,9 секунды. Несмотря на увеличение времени одной итерации из-за механизма «элитных муравьёв» и дополнительных операций с матрицей феромонного следа, ускорение сходимости за счет усиления феромонного влияния на перспективные маршруты компенсирует этот эффект. Таким образом, предложенный подход позволил улучшить качество и эффективность маршрутизации в среднем на 10-15%.

**8. Заключение.** В исследовании рассмотрена проблема маршрутизации автономных устройств в трёхмерном пространстве, которая представляет собой одну из ключевых задач в области интеллектуального управления. Актуальность данной проблемы

обусловлена необходимостью эффективного решения задач оптимизации траекторий в условиях высокой степени свободы, сложной топологии и динамично изменяющейся среды. Особенности трёхмерной среды, включающие высокую степень свободы, сложную топологическую организацию и динамически изменяющиеся параметры, существенно усложняют процесс построения эффективных и безопасных маршрутов. В связи с этим, разработанный подход к решению задачи маршрутизации интегрирует высокоуровневое моделирование рабочего пространства с использованием иерархических структур данных, таких как октодеревья. Он основан на применении метаэвристических алгоритмов оптимизации, которые обеспечивают высокую адаптивность к динамическим изменениям и вычислительную эффективность, необходимую для решения задачи маршрутизации в условиях сложной топологии. Данный подход к маршрутизации автономных устройств в трёхмерном пространстве был реализован в среде компьютерного моделирования Unreal Engine 5.

Для подтверждения эффективности предложенного подхода была проведена серия вычислительных экспериментов, которая подтвердила эффективность разработанного алгоритма маршрутизации, включая сокращение времени вычислений и повышение энергоэффективности автономных устройств в среднем на 10-15%.

Для задачи маршрутизации в трёхмерном пространстве оптимальным решением является использование модифицированного муравьиного алгоритма, который продемонстрировал высокую производительность и энергоэффективность.

Перспективы дальнейших исследований включают расширение возможностей системы, включая интеграцию пользовательских алгоритмов маршрутизации без необходимости перекомпиляции исходного кода, внедрение асинхронной обработки данных и оптимизацию времени выполнения итераций. Предполагается, что такие улучшения позволят существенно повысить адаптивность и вычислительную эффективность системы, сохранив её работоспособность в заданных режимах.

### **Литература**

1. Курейчик В.В., Гладков Л.А., Кравченко Ю.А., Родзин С.И. Интеллектуальные системы: модели и методы метаэвристической оптимизации: Монография // Чебоксары: ООО «Издательский дом «Среда», 2024. 228 с. DOI: 10.31483/a-10639.
2. Даринцев О.В., Мигранов А.Б. Аналитический обзор подходов к распределению задач в группах мобильных роботов на основе технологий мягких вычислений //

- Информатика и автоматизация. 2022. Т. 21. № 4. С. 729–757. DOI: 10.15622/ia.21.4.4.
3. Курейчик В.В., Родзин С.И. Вычислительные модели эволюционных и роевых биоэвристик (обзор) // Информационные технологии. 2021. Т. 27. № 10. С. 507–520. DOI: 10.17587/it.27.507-520.
  4. Перепелкин Д.А., Иванчикова М.А., Нгуен В.Т. Нейросетевая многопутевая маршрутизация в программно-конфигурируемых сетях на основе генетического алгоритма // Информационные технологии. 2023. Т. 29. № 12. С. 622–629. DOI: 10.17587/it.29.622-629.
  5. Казакова Е.М. Применение метода роя частиц в задачах оптимизации // Известия Кабардино-Балкарского научного центра РАН. 2022. № 5(109). С. 48–57. DOI: 10.35330/1991-6639-2022-5-109-48-57.
  6. Казаков К.А., Семенов В.А. Обзор современных методов планирования пути // Тр. ИСП РАН. 2016. Т. 28(4). С. 241–294. DOI: 10.15514/ISPRAS-2016-28(4)-14.
  7. Гладков Л.А., Гладкова Н.В. Эволюционирующие многоагентные системы и эволюционное проектирование // Известия ЮФУ. Технические науки. 2020. № 4(214). С. 48–59. DOI: 10.18522/2311-3103-2020-4-48-59.
  8. Рачков Т.И., Кузьмина И.А. Мета-эвристический алгоритм децентрализованного управления группой // Математические методы в технологиях и технике. 2021. № 3. С. 96–99. DOI: 10.52348/2712-8873\_MMTT\_2021\_3\_96.
  9. Костин А.С., Майоров Н.Н. Исследование моделей и методов маршрутизации и практического выполнения автономного движения беспилотными транспортными системами для доставки грузов // Вестник государственного университета морского и речного флота им. адмирала С.О. Макарова. 2023. Т. 15. № 3. С. 524–536. DOI: 10.21821/2309-5180-2023-15-3-524-536.
  10. Акопов А.С., Бекларян Л.А., Бекларян А.Л. Оптимизация характеристик интеллектуальной транспортной системы с использованием генетического алгоритма вещественного кодирования на основе адаптивной мутации // Информационные технологии. 2023. Т. 29. № 3. С. 115–125. DOI: 10.17587/it.29.115-125.
  11. Карпенко А.П. Эволюционные операторы популяционных алгоритмов глобальной оптимизации. Опыт систематизации // Математика и математическое моделирование. 2018. № 1. С. 59–89. DOI: 10.24108/mathm.0118.0000103.
  12. Медведев М.Ю., Костюков В.А., Пшихопов В.Х. Метод оптимизации траектории мобильного робота в поле источников-репеллеров // Информатика и автоматизация. 2021. Т. 20. № 3. С. 690–726. DOI: 10.15622/ia.2021.3.7.
  13. Литвиненко А.М., Кудрявцев Г.В., Ибрагимов М.У.У. Исследование адаптивной системы управления электроэнергетическим комплексом с ветроэлектрогенератором // Вести высших учебных заведений Черноземья. 2021. № 3(65). С. 10–17. DOI: 10.53015/18159958\_2021\_3\_10.
  14. Курочкин А.Г., Тигенко Е.А. Модифицированный алгоритм сглаживания точек маршрута // Известия Юго-Западного государственного университета. 2016. № 5(68). С. 43–51.
  15. Шмалько Е.Ю., Румянцев Ю.А., Байназаров Р.Р., Ямшанов К.Л. Идентификация нейросетевой модели робота для решения задачи оптимального управления // Информатика и автоматизация. 2021. Т. 20. № 6. С. 1254–1278. DOI: 10.15622/ia.20.6.3.
  16. Ургатов В.А. Компьютерное моделирование основных типов движения мультикоптера в трёхмерном пространстве // Вестник Совета молодых ученых Рязанского государственного аграрно-технологического университета имени П.А. Костычева. 2016. № 1(2). С. 216–220.



17. Жарков С.Н. Стохастическое формирование проактивного множества при кластеризации в мобильных беспроводных сенсорных сетях // Т-Comm: Телекоммуникации и транспорт. 2013. Т. 7. № 5. С. 29–34.
18. Шилов Н.Г., Пономарев А.В., Смирнов А.В. Анализ методов онтолого-ориентированного нейро-символического интеллекта при коллаборативной поддержке принятия решений // Информатика и автоматизация. 2023. Т. 22. № 3. С. 576–615. DOI: 10.15622/ia.22.3.4.
19. Danilchenko V.I., Danilchenko E.V., Kureichik V.M. Application of Genetic Algorithms in Solving the Problem of Placing Elements on a Crystal Taking into Account the Criterion of the Maximum Number of Linear Segments // Proceedings of the Fifth International Scientific Conference “Intelligent Information Technologies for Industry” (IITI’21). Lecture Notes in Networks and Systems. 2022. vol. 330. pp. 276–284. DOI: 10.1007/978-3-030-87178-9\_28.
20. Валькман Ю.Р., Тарасов В.Б. От онтологий проектирования к когнитивной семиотике // Онтология проектирования. 2018. Т. 8. № 1(27). С. 8–34. DOI: 10.18287/2223-9537-2018-8-1-8-34.

**Курейчик Владимир Викторович** — д-р техн. наук, профессор, заведующий кафедрой, САПР им. В.М. Курейчика, ИКТИБ, Южный Федеральный университет. Область научных интересов: компьютерное моделирование, автоматизация проектирования. Число научных публикаций — 411. vkur@sfedu.ru; улица Энгельса, 1, к. Г-435, 347928, Таганрог, Россия; р.т.: +7(863)438-3451.

**Данильченко Владислав Иванович** — канд. техн. наук, доцент кафедры, САПР им. В.М. Курейчика, ИКТИБ, Южный Федеральный университет. Область научных интересов: компьютерное моделирование, автоматизация проектирования. Число научных публикаций — 58. vdanielchenko@sfedu.ru; улица Энгельса, 1, к. Г-435, 347928, Таганрог, Россия; р.т.: +7(863)437-1651.

**Данильченко Евгения Владимировна** — ассистент кафедры, САПР им. В.М. Курейчика, ИКТИБ, Южный Федеральный университет. Область научных интересов: автоматика, вычислительная техника. Число научных публикаций — 20. lipkina@sfedu.ru; улица Энгельса, 1, к. Г-435, 347928, Таганрог, Россия; р.т.: +7(863)437-1651.

**Поддержка исследований.** Исследование выполнено за счет гранта Российского научного фонда № 24-71-00035, <https://rscf.ru/project/24-71-00035/> в Южном федеральном университете.

V. KUREYCHIK, V. DANILCHENKO, E. DANILCHENKO  
**ROUTING OF AUTONOMOUS DEVICES IN THREE-  
DIMENSIONAL SPACE**

*Kureychik V., Danilchenko V., Danilchenko E. Routing of Autonomous Devices in Three-Dimensional Space.*

**Abstract.** The article addresses the problem of routing autonomous devices in three-dimensional space, which is a relevant task for intelligent control. The three-dimensional space is characterized by a high degree of freedom, complex topology, and dynamic environmental changes, which significantly complicate the task of effective trajectory planning. The development of routing methods that ensure safety, energy efficiency, and computational efficiency is crucial for improving the performance of autonomous systems. The paper presents a comprehensive routing system based on a hybrid approach that combines high-level modeling of the working space with metaheuristic optimization methods. Hierarchical data structures, such as octrees, are used to represent the three-dimensional environment, providing compactness and flexibility for spatial models. These models are transformed into graph structures, allowing the routing problem to be described as an optimization problem on graphs. A modified metaheuristic ant colony optimization algorithm, belonging to the class of swarm optimization methods, is proposed. The algorithm is designed to build safe and energy-efficient routes, as well as to solve problems related to finding the shortest Hamiltonian cycles and dynamically reconfiguring routes in a changing external environment. The paper presents the results of computational experiments, including algorithm testing in three-dimensional space and a comparative analysis with other routing algorithms. The computational experiment confirmed the effectiveness of the developed routing algorithm, including reduced computation time and improved energy efficiency of autonomous devices. The prospects for further research include integrating the proposed system into a wide range of applications for autonomous devices aimed at optimizing control processes and enhancing performance in a dynamically changing external environment. It is worth noting that the developed algorithm can be adapted to solve complex tasks where routing and wind generator placement on a plane are interrelated. The placement problem is directly connected to route construction for servicing these objects, which requires a comprehensive approach for an efficient solution. This will be part of a decision support system designed for the planning and servicing of wind power complexes, ensuring their effective operation and resource management.

**Keywords:** metaheuristic algorithm, ant colony optimization, graph-based mathematical models, routing, three-dimensional space modeling, energy systems.

## References

1. Kureichik V.V., Gladkov L.A., Kravchenko Yu.A., Rodzin S.I. *Intellektual'nye sistemy: modeli i metody metatejvristicheskoj optimizacii: Monografija [Intelligent Systems: Models and Methods of Metaheuristic Optimization. Monograph].* Cheboksary: Publishing House «Sreda», 2024. 228 p. DOI: 10.31483/a-10639. (In Russ.).
2. Darintsev O.V., Migranov A.B. [Analytical Review of Approaches to Task Allocation in Groups of Mobile Robots Based on Soft Computing Technologies]. *Informatika i avtomatizacija – Informatics and Automation.* 2022. vol. 21. no. 4. pp. 729–757. DOI: 10.15622/ia.21.4.4. (In Russ.).

3. Kureichik V.V., Rodzin S.I. [Computational Models of Evolutionary and Swarm Bioheuristics: A Review]. *Informacionnye tehnologii – Information Technologies*. 2021. vol. 27. no. 10. pp. 507–520. DOI: 10.17587/it.27.507-520. (In Russ.).
4. Perepelkin D.A., Ivanchikova M.A., Nguyen V.T. [Neural Network Multi-Path Routing in Software-Defined Networks Based on Genetic Algorithm]. *Informacionnye tehnologii – Information Technologies*. 2023. vol. 29. no. 12. pp. 622–629. DOI: 10.17587/it.29.622-629. (In Russ.).
5. Kazakova E.M. [Application of Particle Swarm Optimization in Optimization Problems]. *Izvestija Kabardino-Balkarskogo nauchnogo centra RAN – Proceedings of the Kabardino-Balkar Scientific Center of the Russian Academy of Sciences*. 2022. no. 5(109). pp. 48–57. DOI: 10.35330/1991-6639-2022-5-109-48-57. (In Russ.).
6. Kazakov K.A., Semenov V.A. [An overview of modern methods for motion planning]. *Trudy ISP RAN – Труды ИСП РАН*. 2016. vol. 28(4). pp. 241–294. DOI: 10.15514/ISPRAS-2016-28(4)-14. (In Russ.).
7. Gladkov L.A., Gladkova N.V. [Evolving Multi-Agent Systems and Evolutionary Design]. *Izvestija JuFU. Tehnicheskie nauki – Proceedings of Southern Federal University. Technical Sciences*. 2020. no. 4(214). pp. 48–59. DOI: 10.18522/2311-3103-2020-4-48-59. (In Russ.).
8. Rachkov T.I., Kuzmina I.A. [Metaheuristic Algorithm for Decentralized Group Control]. *Matematicheskie metody v tehnologiijah i tehnike – Mathematical Methods in Technologies and Engineering*. 2021. no. 3. pp. 96–99. DOI: 10.52348/2712-8873\_MMTT\_2021\_3\_96. (In Russ.).
9. Kostin A.S., Maiorov N.N. [Study of Models and Routing Methods for Autonomous Movement of Unmanned Transport Systems for Freight Delivery]. *Vestnik gosudarstvennogo universiteta morskogo i rechnogo flota im. admirala S.O. Makarova – Bulletin of the Admiral S.O. Makarov State University of Maritime and Inland Shipping*. 2023. vol. 15. no. 3. pp. 524–536. DOI: 10.21821/2309-5180-2023-15-3-524-536. (In Russ.).
10. Akopov A.S., Beklaryan L.A., Beklaryan A.L. [Optimization of Intelligent Transportation System Characteristics Using Genetic Algorithm with Real-Valued Encoding Based on Adaptive Mutation]. *Informacionnye tehnologii – Information Technologies*. 2023. vol. 29. no. 3. pp. 115–125. DOI: 10.17587/it.29.115-125. (In Russ.).
11. Karpenko A.P. [Evolutionary Operators of Population Algorithms for Global Optimization: A Systematic Review]. *Matematika i matematicheskoe modelirovanie – Mathematics and Mathematical Modeling*. 2018. no. 1. pp. 59–89. DOI: 10.24108/mathm.0118.0000103. (In Russ.).
12. Kostyukov V.A., Medvedev M.Yu., Pshikhov V.Kh. [Trajectory Optimization Method for Mobile Robots in the Field of Repeller Sources]. *Informatika i avtomatizacija – Informatics and Automation*. 2021. vol. 20. no. 3. pp. 690–726. DOI: 10.15622/ia.2021.3.7. (In Russ.).
13. Litvinenko A.M., Kudryavtsev G.V., Ibrahimov M.U. [Study of an Adaptive Control System for an Electric Power Complex with Wind Turbine Generators]. *Vesti vysshih uchebnyh zavedenij Chernozem'ja – Herald of Higher Educational Institutions of the Black Earth Region*. 2021. no. 3(65). pp. 10–17. DOI: 10.53015/18159958\_2021\_3\_10. (In Russ.).
14. Kurochkin A.G., Titenko E.A. [Modified Algorithm for Smoothing Route Points]. *Izvestija Jugo-Zapadnogo gosudarstvennogo universiteta – Proceedings of the Southwest State University*. 2016. no. 5(68). pp. 43–51. (In Russ.).
15. Shmalko E.Yu., Rumyantsev Yu.A., Baynazarov R.R., Yamshanov K.L. [Identification of a Neural Network Model for Optimal Control of a Robot].

- Informatika i avtomatizacija – Informatics and Automation. 2021. vol. 20. no. 6. pp. 1254–1278. DOI: 10.15622/ia.20.6.3. (In Russ.).
16. Urgapov V.A. [Computer Modeling of Basic Multicopter Motion Types in Three-Dimensional Space]. Vestnik Soveta molodyh uchenyh Rjazanskogo gosudarstvennogo agrotehnologicheskogo universiteta imeni P.A. Kostycheva – Herald of the Young Scientists Council of the Ryazan State Agrotechnological University named after P.A. Kostychev. 2016. no. 1(2). pp. 216–220. (In Russ.).
  17. Zharkov S.N. [Stochastic Formation of a Proactive Set in Clustering of Mobile Wireless Sensor Networks]. T-Comm: Telekommunikacii i transport – T-Comm: Telecommunications and Transport. 2013. vol. 7. no. 5. pp. 29–34. (In Russ.).
  18. Shilov N.G., Ponomaryov A.V., Smirnov A.V. [Analysis of Ontology-Oriented Neuro-Symbolic Intelligence Methods in Collaborative Decision-Making Support]. Informatika i avtomatizacija – Informatics and Automation. 2023. vol. 22. no. 3. pp. 576–615. DOI: 10.15622/ia.22.3.4. (In Russ.).
  19. Danilchenko V.I., Danilchenko E.V., Kurechik V.M. Application of Genetic Algorithms in Solving the Problem of Placing Elements on a Crystal Taking into Account the Criterion of the Maximum Number of Linear Segments. Proceedings of the Fifth International Scientific Conference «Intelligent Information Technologies for Industry» (IITI'21). Lecture Notes in Networks and Systems. 2022. vol. 330. pp. 276–284. DOI: 10.1007/978-3-030-87178-9\_28.
  20. Valkman Yu.R., Tarasov V.B. [From Design Ontologies to Cognitive Semiotics]. Ontologija proektirovanija – Ontology of Design. 2018. vol. 8. no. 1(27). pp. 8–34. DOI: 10.18287/2223-9537-2018-8-1-8-34. (In Russ.).

**Kureychik Vladimir** — Ph.D., Dr.Sci., Professor, Head of the department, V.M. Kureychik CAD department, ICTIS, Southern Federal University. Research interests: computer modeling, design automation. The number of publications — 411. vkur@sfedu.ru; 1, cor. G-435, Engels St., 347928, Taganrog, Russia; office phone: +7(863)438-3451.

**Danilchenko Vladislav** — Ph.D., Associate professor of the department, V.M. Kureychik CAD department, ICTIS, Southern Federal University. Research interests: computer modeling, design automation. The number of publications — 58. vdanilchenko@sfedu.ru; 1, cor. G-435, Engels St., 347928, Taganrog, Russia; office phone: +7(863)437-1651.

**Danilchenko Evgeniya** — Assistant of the department, V.M. Kureychik CAD department, ICTIS, Southern Federal University. Research interests: automation, computer engineering. The number of publications — 20. lipkina@sfedu.ru; 1, cor. G-435, Engels St., 347928, Taganrog, Russia; office phone: +7(863)437-1651.

**Acknowledgements.** The research was funded by the Russian Science Foundation project No. 24-71-00035, <https://rscf.ru/project/24-71-00035/> implemented by the Southern Federal University.

Д.В. НАГИБИН, А.С. ПЕТРЕНКО, В.С. ДАВЫДЕНКО, И.В. КОТЕНКО,  
Е.В. ФЕДОРЧЕНКО

## **ИНВАЗИВНЫЙ ПОДХОД К ВЕРИФИКАЦИИ ФУНКЦИОНАЛЬНО-СТРУКТУРНЫХ СПЕЦИФИКАЦИЙ, РЕАЛИЗОВАННЫХ В ЗАКАЗНЫХ ИНТЕГРАЛЬНЫХ СХЕМАХ**

*Нагибин Д.В., Петренко А.С., Давыденко В.С., Котенко И.В., Федорченко Е.В.*  
**Инвазивный подход к верификации функционально-структурных спецификаций, реализованных в заказных интегральных схемах.**

**Аннотация.** Представлен подход к верификации функционально-структурных спецификаций, реализованных в заказных интегральных схемах, основанный на инвазивных методах исследования. Актуальность проведённого исследования обусловлена необходимостью проведения верификации функционально-структурных спецификаций, поставляемых сторонними исполнителями аппаратных реализаций алгоритмов обеспечения информационной безопасности, сложностью выявления на аппаратном уровне модификаций этих алгоритмов и внедрённых в них недокументированных возможностей и отсутствием единых универсальных или стандартизированных методов решения этой задачи. Сформулирована математическая постановка задачи исследования, суть которой состоит в проверке равенства значений параметров заявленной спецификации с их значениями, восстановленными методом обратного проектирования. Представлены результаты применения предложенного подхода к верификации функционально-структурных спецификаций на примерах аппаратно-реализованных алгоритмов шифрования DES и AES. Восстановленные функционально-структурные блоки алгоритмов (в частности – блок подстановок) были успешно верифицированы.

**Ключевые слова:** заказная интегральная схема, идентификация, верификация, функционально-структурная спецификация, алгоритм шифрования.

**1. Введение.** Стойкость криптографических алгоритмов определяется используемым ключом шифрования, который для нарушителя изначально неизвестен. Сами алгоритмы шифрования и их параметры открыты для ознакомления, изучения, проверок и поисков уязвимостей. При этом внесение любого изменения в криптографические алгоритмы может привести к появлению уязвимостей и, как следствие, к снижению стойкости алгоритма. Поэтому проведение проверок в поставляемом оборудовании или программном обеспечении на предмет соответствия реализованных в них криптографических алгоритмов заявленным спецификациям является актуальной задачей для информационной безопасности [1, 2].

Особенно остро эта задача стоит при верификации алгоритмов, реализованных аппаратно, так как, по сравнению с программными реализациями, аппаратные реализации предоставляют гораздо больше вариантов для скрытного внесения модификаций и недокументированных возможностей в технические устройства.

Указанная задача декомпозируется на 1) идентификацию алгоритмов в исследуемом устройстве и 2) их верификацию.

Решение первой из указанных задач было предложено в статье [1]. В данной работе рассматривается верификация алгоритмов на примерах исследования аппаратно-реализованных алгоритмов шифрования DES (Data Encryption Standard) [3 – 5] и AES (Advanced Encryption Standard) [6 – 8], который де-факто является международным стандартом шифрования.

Хотя попытки автоматизировать решение задачи верификации программных и/или аппаратных систем в мировом сообществе ведутся давно, на сегодняшний день отсутствуют единые универсальные или стандартизированные методы решения этой задачи. Связано это с разнородностью реализаций, осуществляемых различными производителями. Кроме того, трудность этой задачи обусловлена и архитектурой построения современных технических устройств, целевая функция которых распределяется между программной и аппаратной реализациями. Само это распределение не стандартизировано. Таким образом, для решения задачи верификации технических устройств или систем необходимо проводить исследования как программных, так и аппаратных частей их реализаций.

Настоящая работа направлена на решение задачи верификации функционально-структурных спецификаций аппаратных реализаций технических устройств. Данная работа отличается применением инвазивных методов исследования и нацеленностью на анализ аппаратных реализаций алгоритмов шифрования, что определило некоторые её особенности. Исследование вносит вклад в практику решения проблемы верификации функционально-структурных спецификаций заказных интегральных схем (ЗИС) недоверенных производителей, например, иностранного производства.

При этом термин «спецификация» в разных отраслях определён по-разному. В рамках настоящего исследования будем руководствоваться следующей формулировкой: спецификация – это документ, содержащий подробное перечисление узлов и деталей какого-либо изделия, конструкции, установки, и т.п., входящих в состав сборочного или монтажного чертежа, а также документ с перечислением условий, которым должен удовлетворять производственный заказ.

Работа организована следующим образом. В разделе 2 дается общая характеристика релевантных исследований. В разделе 3 формулируется задача исследования. В разделе 4 предлагается подход

к верификации функционально-структурных спецификаций, реализованных в заказных интегральных схемах. В разделе 5 описывается применение данного подхода на примере алгоритма шифрования DES. В разделе 6 рассматривается применение подхода к верификации на примере алгоритма шифрования AES. В разделе 7 оценивается результативность применения этого подхода. В разделе 8 подводятся итоги проделанной работы и определяются направления дальнейших исследований.

Алгоритмы DES и AES выбраны, потому что мировым сообществом они признаны в качестве стандартов шифрования. В дальнейшем планируется применение предложенных решений к другим аппаратно реализованным алгоритмам.

**2. Релевантные исследования.** Можно выделить следующие исследования, посвященные автоматизации решения задачи верификации программных и/или аппаратных систем.

В [9 – 11] представлены общие принципы проектирования защищенных систем, включая рассмотрение вопросов, связанных с верификацией различных классов систем: систем на основе микроконтроллеров [9], защищенных встроенных устройств на примере системы охраны периметра [10] и безопасных встроенных систем [11].

Ряд работ направлен на статическую верификацию исполняемых программ и исходного кода программ. В.В. Ковалев и др. [12] предложили подход к статической верификации исполняемых программ, основанный на сопоставлении семантических аспектов вычислений. В [13] рассматривается статический анализ исходного кода программ на основе абстрактной интерпретации. В [14] авторы разработали подход к верификации программ с «нулевыми» знаниями, т.е. без раскрытия исходного кода.

В [15] авторы рассмотрели подход к верификации протоколов безопасности на основе комбинированного использования существующих методов и средств. Tran D.D. и Ogata K. разработали инвариантный генератор оценки доказательства IPSG [16] на основе интерпретатора SafeOBJ [17, 18]. В [19] описано его применение для верификации протокола TLS. Saeed и др. предложили подход к формальной верификации криптографических протоколов с использованием методов проверки на модели и частичного порядка [20]. В [21] авторы разработали инструмент CryptoFormalEval автоматического выявления уязвимостей в криптографических протоколах на основе больших языковых моделей и формальной верификации.

В [22] авторы предложили метод верификации функционально-структурных спецификаций технических устройств на основе интерпретатора SafeOBJ с поддержкой тестовых оценок и проверкой теорем, а также генератором тестовых сценариев на основе тестовых оценок с возможностью исправления ошибок.

В [23] авторы разработали метод установления эквивалентности скомпилированных двоичных файлов для конкретного программного обеспечения на основе проверки того, являются ли исходный и двоичный графы потока управления изоморфными после преобразований графа, сохраняющих семантику.

В [24] авторы представили метод адаптивного извлечения структурных признаков аппаратных реализаций троянских программ на уровне шлюза GateDet, который включает в себя усовершенствованный метод моделирования графов схем и разработанный метод обнаружения, основанный на двунаправленных графовых нейронных сетях.

В [25] авторы предложили метод идентификации аппаратно-реализованных троянов на основе анализа данных из побочных каналов с применением сетевых архитектур ResNeXt и ARA, позволяющий эффективно различать типы аппаратных троянских программ, в том числе и в условиях отсутствия эталонного образца.

В [26] авторы разработали модель процесса обратного проектирования сложных производственных систем, которая представляет собой сложную систему динамических, итеративных, параллельных, рекурсивных и зависящих от времени процессов.

В [27] авторы разработали многопараметрическую систему обнаружения троянских программ на основе машинного обучения с поддержкой анализа логики передачи регистров.

В [28] авторы предложили неинвазивный метод обнаружения аппаратно-реализованных троянских программ, основанный на использовании аппаратных средств с нулевой задержкой (с использованием пространственных корреляций для подавления методов скрытия троянских программ).

В [29] авторы рассмотрели методику преобразования программного обеспечения, направленную на внедрение обфускации при выполнении программ. Целью исследования являлась защита микропроцессорных систем от недокументированных возможностей аппаратных реализаций на основе создания модифицированной версии ассемблерного кода.



В [30] авторы описали методику идентификации аппаратно-реализованных троянов в конвейерных микропроцессорах на основе использования семейства классификаторов.

В [31] авторы представили методику идентификации аппаратно-реализуемых троянов, внедряемых в аппаратное обеспечение средствами систем автоматизированного проектирования, на основе методов машинного обучения.

Недостатком проанализированных исследований является применение неинвазивных методов, которые нацелены, в основном, именно на поиск недокументированных возможностей. К таким методам относится анализ побочных каналов, пространственная корреляция, машинное обучение, анализ управляющих программ и др. Использование только процедуры поиска не позволяет говорить однозначно о гарантированном отсутствии недокументированных возможностей.

Настоящая работа применяет инвазивные (разрушающие) методы исследования и имеет более общую цель – верификацию функционально-структурных спецификаций технических устройств, которая включает также выявление недокументированных возможностей. Специфической особенностью предлагаемого подхода является нацеленность на анализ аппаратных реализаций алгоритмов шифрования.

**3. Постановка задачи исследования.** Нелинейность криптографическим преобразованиям в блочных шифрах придают блоки подстановок (так называемые S-блоки, от английского «substitution»), что определило высокий интерес специалистов в области информационной безопасности к анализу этих блоков.

*Определение 1.* S-блок (S-box, блок замены, блок подстановки) – представляет собой отображение из множества двоичных векторов длины  $n$  в множество двоичных векторов длины  $m$  [32].

При этом числа  $n$  и  $m$  относительно малы, например 4, 6, 8, 16, 32. Часто рассматриваются S-блоки, являющиеся взаимно однозначными преобразованиями, в этом случае  $n = m$ .

Обычно они хранятся в виде таблиц как массивы данных. На рисунке 1 представлен фрагмент схемы криптографического преобразования информации с помощью S-блоков в алгоритме DES [33], который можно считать типовым для подобных преобразований. Структура криптографического преобразования с помощью S-блоков алгоритма DES насчитывает 8 таких блоков. Кроме того, на ней наглядно проиллюстрирована разноразрядность

входного и выходного блока. Однако эти параметры в других алгоритмах могут отличаться.

Обычно входные данные для S-блоков рассматриваются как адрес ячейки S-блока, значение которой считывается и подаётся в качестве выходной информации из S-блока.

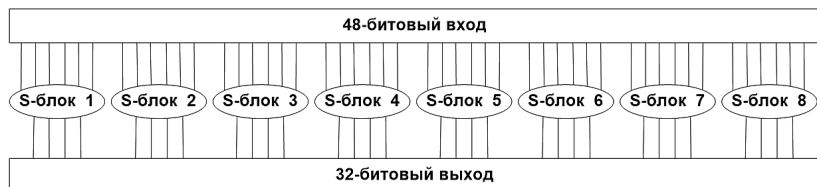


Рис. 1. Схема криптографического преобразования информации с помощью S-блоков в алгоритме DES

Математически S-блок является векторной булевой функцией.

*Определение 2.* Векторная булева функция – функция вида

$$F : \mathbb{Z}_2^n \rightarrow \mathbb{Z}_2^m .$$

Другое название векторной булевой функции:  $(n,m)$ -функция [32].

На рисунке 2 представлены шестнадцатеричные значения блока подстановок алгоритма AES [6].

	0	1	2	3	4	5	6	7	8	9	a	b	c	d	e	f
0	63	7c	77	7b	f2	6b	6f	c5	30	01	67	2b	fe	d7	ab	76
1	ca	82	c9	7d	fa	59	47	f0	ad	d4	a2	af	9c	a4	72	c0
2	b7	fd	93	26	36	3f	f7	cc	34	a5	e5	f1	71	d8	31	15
3	04	c7	23	c3	18	96	05	9a	07	12	80	e2	eb	27	b2	75
4	09	83	2c	1a	1b	6e	5a	a0	52	3b	d6	b3	29	e3	2f	84
5	53	d1	00	ed	20	fc	b1	5b	6a	cb	be	39	4a	4c	58	cf
6	d0	ef	aa	fb	43	4d	33	85	45	f9	02	7f	50	3c	9f	a8
7	51	a3	40	8f	92	9d	38	f5	bc	b6	da	21	10	ff	f3	d2
8	cd	0c	13	ec	5f	97	44	17	c4	a7	7e	3d	64	5d	19	73
9	60	81	4f	dc	22	2a	90	88	46	ee	b8	14	de	5e	0b	db
a	e0	32	3a	0a	49	06	24	5c	c2	d3	ac	62	91	95	e4	79
b	e7	c8	37	6d	8d	d5	4e	a9	6c	56	f4	ea	65	7a	ae	08
c	ba	78	25	2e	1c	a6	b4	c6	e8	dd	74	1f	4b	bd	8b	8a
d	70	3e	b5	66	48	03	f6	0e	61	35	57	b9	86	c1	1d	9e
e	e1	f8	98	11	69	d9	8e	94	9b	1e	87	e9	ce	55	28	df
f	8c	a1	89	0d	bf	e6	42	68	41	99	2d	0f	b0	54	bb	16

Рис. 2. Значения блока подстановок алгоритма AES

В рассматриваемом примере разрядность входных данных равна 8 битам (4 бита определяют адрес строки, 4 бита определяют адрес столбца в таблице) и размерность выходных данных равна 8 битам (определяется количеством бит, необходимых для представления шестнадцатеричных чисел от 0x00 до 0xFF).

Самый распространённый вариант хранения этой таблицы – хранение в памяти. Однако блоки памяти можно считывать и перезаписывать. В этом случае довольно проблематично скрытно внести в них модификации. То есть, такой вариант хранения значений блоков подстановок прост для реализации, но и прост для верификации этих значений.

Второй вариант представления подобных таблиц подстановок – в виде булевых функций каждого выходного бита от входных битов.

*Определение 3.* Произвольная функция  $f$  из множества  $\mathbb{Z}_2^n$  в множество  $\mathbb{Z}_2$  называется булевой функцией от  $n$  переменных [33].

Приведём пример булевой функции от шести переменных:  
$$f(x_5, x_4, x_3, x_2, x_1, x_0) = (x_5 \cdot x_3 \cdot x_1) \oplus (x_4 + x_2 + x_0),$$
 где  $x_0, \dots, x_5 \in \mathbb{Z}_2$ .

Если этот вариант реализован аппаратно, то для его проверки (верификации) необходимо использовать средства аппаратного обратного проектирования. Внедрённые модификации на аппаратном уровне могут долгое время оставаться не выявленными для конечных пользователей технических устройств, что (как ранее было сказано во введении) является уязвимостью, через которую могут быть реализованы различные угрозы информационной безопасности (получение доступа к закрытой переписке, искажение информации и прочее) [34, 35].

Следует отметить, что таблицы подстановок для каждого алгоритма шифрования подвергаются тщательному изучению и проверке на уязвимости мировым криптографическим сообществом [36]. Потому в основных криптографических алгоритмах и стандартах значения таблиц S-блоков не рекомендуется изменять, так как их исходные значения уже прошли указанные проверки и общепризнаны стойкими к различным методам криптоанализа. Считается, что любое изменение значений таблиц подстановок (как и перестановок) может существенно снизить стойкость алгоритма шифрования к атакам на него [37].

Таким образом, задачу верификации функционально-структурных спецификаций, реализованных в заказных интегральных схемах, можно сформулировать следующим образом.

**Дано:**

1.  $X_{Pr}$  – исходный файл с описанием элементной базы исследуемой ЗИС, представленной в виде помеченного связного графа [1];

2.  $Z_M$  – база знаний интеллектуальной системы поддержки принятия решений (ИСППР), содержащая знания о спецификациях известных и восстановленных ЗИС [1];

3.  $U^{SpO} = \{u_1^{SpO}, u_2^{SpO}, \dots, u_e^{SpO}\}$  – множество спецификаций, потенциально реализуемых на знаниях  $Z_M$  [1];

4.  $U^{SpC} = \{u_1^{SpC}, u_2^{SpC}, \dots, u_c^{SpC}\}$  – множество спецификаций, потенциально реализуемых на основе восстановленной структуры ЗИС [1].

На основе анализа исходного файла  $X_{Pr}$  и ИСППР  $Z_M$ , содержащей знания о спецификациях известных и восстановленных ЗИС, сформируем универсальный для каждого криптографического алгоритма, содержащего блоки подстановок, набор параметров:

$$P = \{n, r, m; c, d\}, \quad (1)$$

где:

$n$  – разрядность входного блока;

$r$  – количество S-блоков;

$m$  – разрядность выходного блока;

$c$  – разрядность входных данных на каждом S-блоке;

$d$  – разрядность выходных данных на каждом S-блоке.

**Найти (проверить):** рассматриваемая задача верификации сводится к проверке равенства:

$$u_i^{SpO} = u_j^{SpC}, \quad (2)$$

где:

$u_i^{SpO}$  – спецификация исследуемого устройства согласно поставляемой производителем технической документации;

$u_j^{SpC}$  – спецификация исследуемого устройства, которая формируется в ходе проведения исследований на основе значений выражения (1) и значений, полученных при выявлении полного

множества значений исследуемых блоков подстановок методом полного перебора всего диапазона их адресов.

Выражение (2) будет выполняться при выполнении следующего равенства:

$$P_{Спец} = P_{Иссл}, \quad (3)$$

где:

$P_{Спец}$  – значения параметров блоков подстановок согласно заявленной спецификации исследуемого устройства;

$P_{Иссл}$  – выявленные значения параметров блоков подстановок исследуемого устройства.

**4. Подход к верификации.** Структурная схема алгоритма верификации функционально-структурных спецификаций, реализованных в заказных интегральных схемах, представлена на рисунке 3.

Блокам 2 и 3 рисунка 3 соответствуют пункты 1 – 4 и формирование выражения (1) математической постановки задачи.

Так как входные данные рассматриваются как адреса ячеек в S-блоках (значения которых необходимо подать в качестве выходных данных), то для идентификации каждого S-блока предлагается анализ считываемых значений этих блоков по следующим адресам:

- адрес нулевой ячейки (на вход каждого блока подстановок подаются все нули);

- старший адрес в таблице (на вход каждого блока подстановок подаются все единицы);

- так называемое «протягивание» в адресе считываемой ячейки единицы (количество считанных таким образом ячеек равно размерности входных данных блока подстановок  $c$ );

- так называемое «протягивание» в адресе считываемой ячейки нуля (количество считанных таким образом ячеек равно размерности входных данных блока подстановок  $c$ ).

Количество  $a$  считанных для анализа ячеек в каждом блоке подстановок равно

$$a(S_i) = 2 + 2 \cdot c, \quad (4)$$

где:

$S_i$  – тестируемый блок подстановок и  $i \in [1; r]$ ;

$r$  – количество S-блоков;

$c$  – разрядность входных данных на каждом S-блоке.

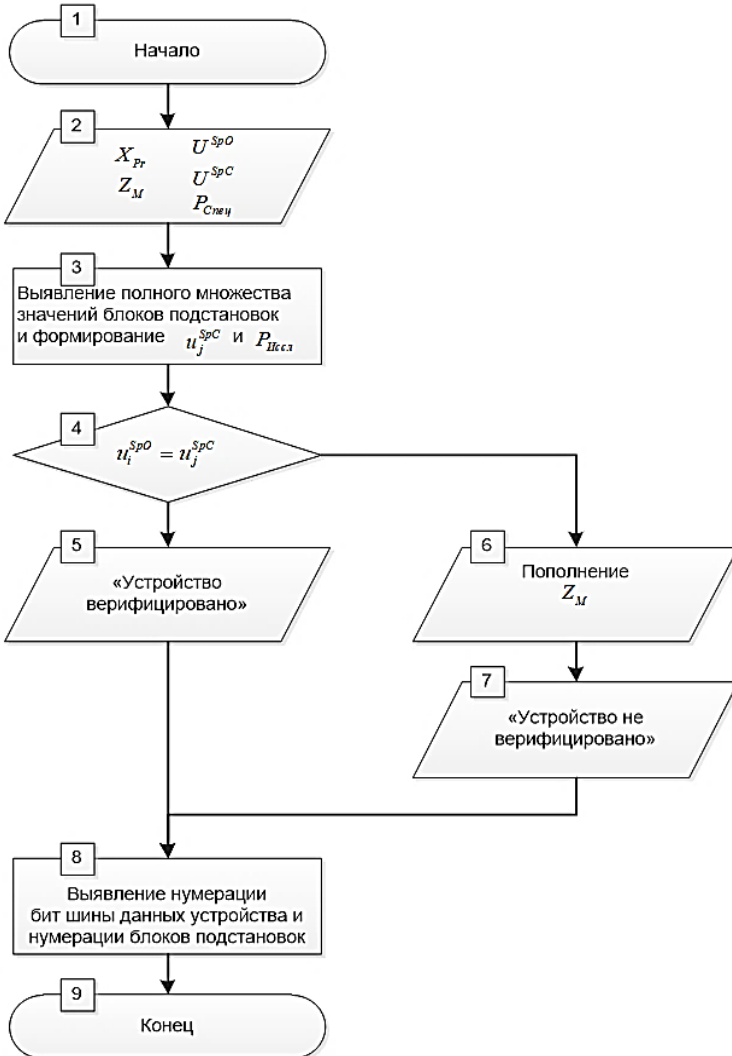


Рис. 3. Структурная схема алгоритма верификации функционально-структурных спецификаций, реализованных в заказных интегральных схемах

Общее количество считанных для анализа ячеек из всех блоков подстановок представлено выражением:

$$\sum_{i=1}^r a(S_i) = 2 \cdot r \cdot (1 + c). \quad (5)$$

В общем случае количество блоков подстановок  $r$  в алгоритмах шифрования больше единицы. Поэтому сначала проводится «первичная» идентификация исследуемого блока – это осуществляется указанным «протягиванием» единицы и нуля. Анализ исследованных алгоритмов показал, что во всех случаях применения такого подхода идентификация проходила успешно.

Такой результат был получен и при анализе указанного алгоритма DES (результаты представлены в разделе 5).

В случае алгоритма AES можно сразу переходить к процессу верификации, так как в AES имеется только один блок подстановок (раздел 6).

Тестируемый алгоритм может оказаться известным и содержаться в базе знаний ИСППР  $Z_M$ . В этом случае переходим к процессу верификации тестируемых блоков подстановок с их каноническими значениями из имеющейся базы знаний – проверке равенств (2) и (3).

В случае подтверждения равенства (2), делается вывод о *верификации* исследуемого устройства (или части устройства) его заявленным производителем спецификациям (блок 5 рисунка 3), а также вывод о доверии исследуемому устройству с точки зрения информационной безопасности.

В противном случае переходим к блоку 6 рисунка 3 и пополняем базу знаний ИСППР  $Z_M$  новым алгоритмом или выявленными модификациями уже известных алгоритмов. Далее делается вывод о *несоответствии* исследуемого устройства (или части устройства) его заявленным производителем спецификациям (блок 7 рисунка 3). Следовательно – исследуемое устройство не может быть признано доверенным для использования с точки зрения информационной безопасности.

В блоке 8 рисунка 3 происходит выявление нумерации S-блоков и выявление нумерации бит шины данных исследуемого алгоритма, что является важным при исследовании аппаратно реализованных алгоритмов шифрования. В начале исследования такая информация

отсутствует, как и значения параметров блоков подстановок, указанных в выражении (1).

На рисунке 4 представлена общая схема использования блоков подстановок в алгоритмах блочного шифрования, на которой проиллюстрировано отсутствие информации о параметрах и характеристиках этих алгоритмов.

Стоит заметить, что в алгоритме DES параметр  $r$  (количество S-блоков) имеет значение 8, а в алгоритме AES этот параметр равен 1 – для этого случая, конечно, выявление нумерации блоков подстановок не требуется.

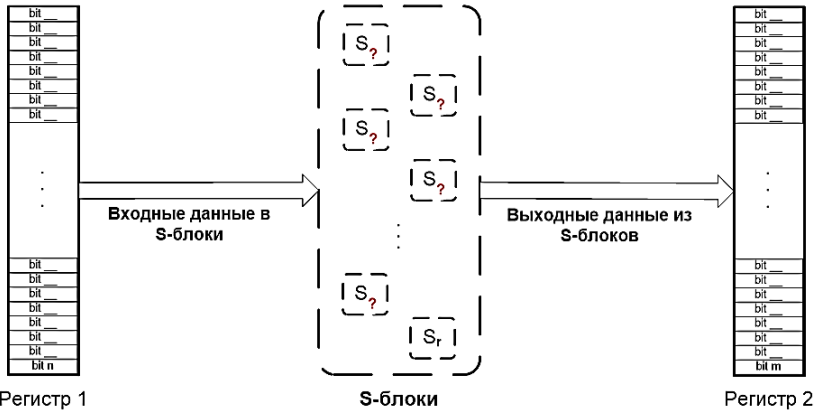


Рис. 4. Общая схема использования блоков подстановок в блочных шифрах при отсутствии информации об их параметрах и характеристиках

Предполагается, что исследование проводится в условиях отсутствия информации о структуре анализируемых технических устройств.

**5. Применение предложенного подхода к блокам подстановок алгоритма DES.** Результат применения предложенного подхода представлен на примере блока подстановок  $S_2$  алгоритма шифрования DES.

Согласно представленной в [1] методике идентификации реализованных в заказных интегральных схемах функционально-структурных спецификаций проводится обратное проектирование исследуемого функционально-структурного блока с укрупнениями (объединениями) элементов исследуемой интегральной схемы по предложенным уровням укрупнения при восстановлении функционально-



структурных спецификаций ЗИС. Указанная методика состоит из следующих шагов:

1. сбор и анализ первичной информации об исследуемой технике и ЗИС;
2. обратное проектирование заказной интегральной схемы;
3. выявление множества возможных режимов функционирования исследуемой заказной интегральной схемы.

Таким образом, настоящая работа является логическим продолжением работы [1] и заключается в переходе от процесса идентификации функционально-структурных блоков ЗИС к процессу их верификации.

На рисунке 5 схематично представлен пример восстановления одного логического блока. Из исходных данных в виде простейших логических элементов (блок 1 рисунка 5) происходит выделение задействованных элементов для создания нового более крупного элемента (блок 2 рисунка 5).

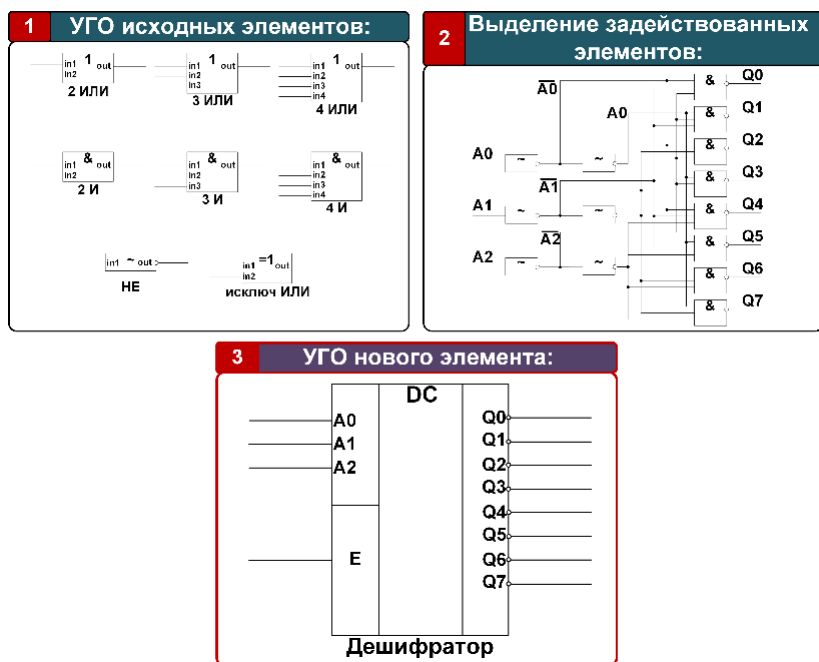


Рис. 5. Пример восстановления логического блока «дешифратор»

Далее проводится анализ логики функционирования выделенной группы элементов, на основе чего формируется условно-графическое обозначение (УГО) нового элемента – дешифратора трёх входных сигналов (блок 3 рисунка 5), который управляет обращением к одному из восьми блоков подстановок алгоритма DES.

Вновь образованный элемент (дешифратор) соответствует уровню 4 укрупнения функционально-структурных элементов заказных интегральных схем [1], что продемонстрировано на рисунке 6.

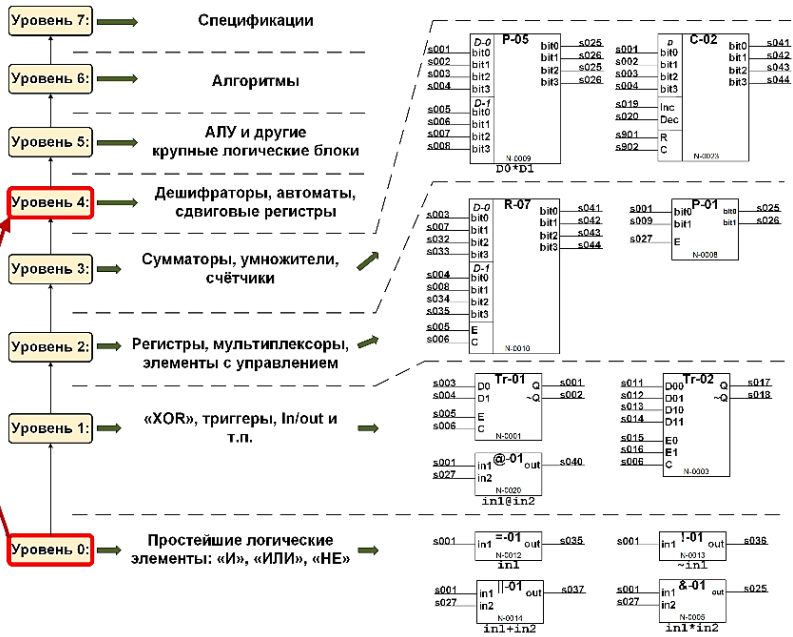


Рис. 6. Укрупнение логических элементов в функционально-структурные блоки

Пример применения предложенного подхода для блока подстановок  $S_2$  алгоритма шифрования DES показан на рисунке 7.

Слева на рисунке 7 в таблице приведены значения тестируемого блока подстановок  $S_2$  с представлением в двоичном виде адресов ячеек. Адреса строк в диапазоне  $[0;3]$ , адреса столбцов в диапазоне  $[0;15]$ .

Справа на рисунке 7 приведена таблица адресов считываемых ячеек блока подстановок  $S_2$ .

Результаты исследований на основе анализа блоков подстановок алгоритма DES показали, что считывание значений ячеек с указанных адресов позволяют однозначно идентифицировать тестируемые блоки подстановок с каноническими блоками подстановок алгоритма DES [3].

Предложенный подход к идентификации основан на анализе, согласно выражению (4), 14-ти из 64-х значений каждого блока подстановок алгоритма шифрования DES.

Как видно из рисунка 7, для анализа считываются по 5 значений ячеек из старшей и младшей строки, и по 2 значения ячеек (младший и старший столбец) из строки с номером  $1_{10}=01_2$  и предпоследней строки (в рассмотренном примере – строка с номером  $2_{10}=10_2$ ).

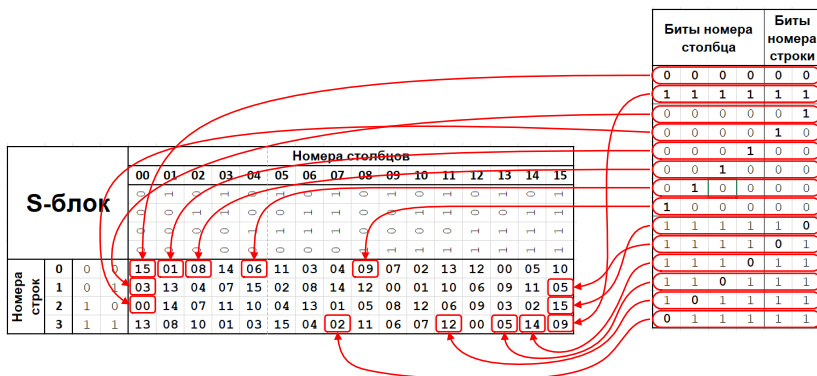


Рис. 7. Пример применения предлагаемого подхода к блоку подстановок  $S_2$  алгоритма шифрования DES

Предложенный авторами подход не зависит от наличия у исследователя информации о нумерации бит на входе в блоки подстановок, так как «протягивание» сначала единицы и потом нуля по входным данным происходит по всем их разрядам. Эта нумерация, как и нумерация битов выходных значений блоков подстановок, была восстановлена в ходе проведения исследований.

В ходе исследования были проанализированы (согласно выражению (5)) 112 значений блока подстановок алгоритма шифрования DES. Как видно из рисунка 7, для анализа считываются по 5 значений ячеек из старшей и младшей строки, а также по 2 значения ячеек из младшего и старшего столбцов.

Таким образом, для рассмотренного примера выражения (1) и (5) имеют вид:

$$P = (n, r, m; c, d) = (48, 8, 32; 6, 4), \quad \sum_{i=1}^r a(S_i) = 2 \cdot r \cdot (1 + c) = 112.$$

По считанным указанным способом значениям происходит первичная идентификация блока подстановок.

Далее происходит последовательное считывание каждой ячейки блоков подстановок с верификацией исследуемых блоков подстановок или с пополнением базы данных алгоритмов шифрования (в случае, когда исследуемый алгоритм не идентифицирован как известный).

**6. Применение разработанного подхода к S-блоку алгоритма AES.** Результат применения предложенного подхода для S-блока алгоритма шифрования AES показан на рисунках 8 и 9.

Согласно [1] проводится обратное проектирование исследуемого функционально-структурного блока с укрупнениями (объединениями) элементов исследуемой ЗИС при восстановлении функционально-структурных спецификаций ЗИС по уровням укрупнения, представленным на рисунке 6. От «Уровня 0» (который соответствует простейшим логическим элементам «И», «Или», «Не») до «Уровня 4» (дешифраторы, автоматы, сдвиговые регистры).

Результат работы отображен на рисунке 8.

В верхней части рисунка 8 представлена логически связанная группа простейших логических элементов исследуемой ЗИС. Согласно структуре уровней детализации элементов (рисунок 6), образуемых в процессе восстановления функционально-структурных спецификаций заказных интегральных схем, на рисунке 8 (часть 1) представлены элементы, которые соответствуют нулевому уровню этой структуры. Анализ взаимосвязей локализованной группы элементов позволил создать новый функционально-логический элемент более высокого уровня укрупнения (структурно и логически представлен внизу на рисунке 8). Вновь образованный элемент (S-блок алгоритма AES) соответствует уровню 4 детализации.

В свою очередь, восстановленный блок подстановок является структурным элементом функции SubBytes (пятый уровень на рисунке 6) алгоритма AES (шестой уровень на рисунке 6).

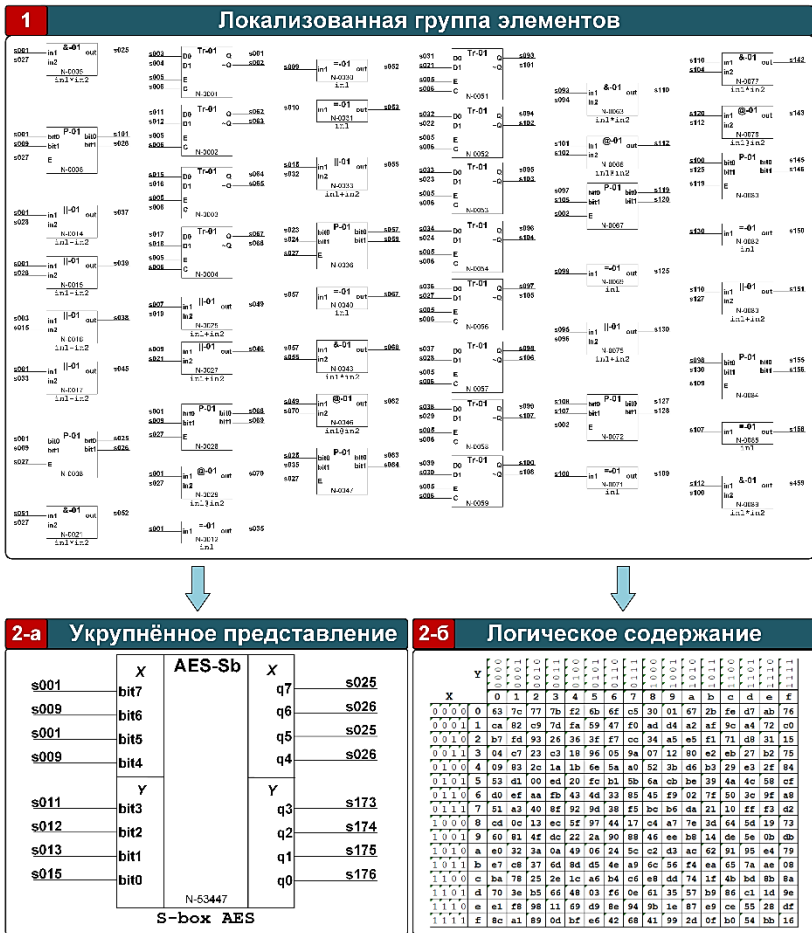


Рис. 8. Пример проведения укрупнения S-блока алгоритма шифрования AES

В верхней части рисунка 9 приведены значения проверяемого блока подстановок с представлением в двоичном виде адресов ячеек в таблице. Адреса строк в диапазоне [0;15], адреса столбцов в диапазоне [0;15].

В нижней части рисунка 9 приведена таблица с адресами и значениями считываемых ячеек исследуемого блока подстановок, которые также выделены в таблице верхней части рисунка.

X	Y	0000	0001	0010	0011	0100	0101	0110	0111	1000	1001	1010	1011	1100	1101	1110	1111
		0	1	2	3	4	5	6	7	8	9	a	b	c	d	e	f
00000	0	63	7c	77	7b	f2	6b	6f	c5	30	01	67	2b	fe	d7	ab	76
00001	1	ca	82	c9	7d	fa	59	47	f0	ad	d4	a2	af	9c	a4	72	c0
00100	2	b7	fd	93	26	36	3f	f7	cc	34	a5	e5	f1	71	d8	31	15
00101	3	04	c7	23	c3	18	96	05	9a	07	12	80	e2	eb	27	b2	75
01000	4	09	83	2c	1a	1b	6e	5a	a0	52	3b	d6	b3	29	e3	2f	84
01001	5	53	d1	00	ed	20	fc	b1	5b	6a	cb	be	39	4a	4c	58	cf
01100	6	d0	ef	aa	fb	43	4d	33	85	45	f9	02	7f	50	3c	9f	a8
01101	7	51	a3	40	8f	92	9d	38	f5	bc	b6	da	21	10	ff	f3	d2
10000	8	cd	0c	13	ec	5f	97	44	17	c4	a7	7e	3d	64	5d	19	73
10001	9	60	81	4f	dc	22	2a	90	88	46	ee	b8	14	de	5e	0b	db
10100	a	e0	32	3a	0a	49	06	24	5c	c2	d3	ac	62	91	95	e4	79
10101	b	e7	c8	37	6d	8d	d5	4e	a9	6c	56	f4	ea	65	7a	ae	08
11000	c	ba	78	25	2e	1c	a6	b4	c6	e8	dd	74	1f	4b	bd	8b	8a
11001	d	70	3e	b5	66	48	03	f6	0e	61	35	57	b9	86	c1	1d	9e
11100	e	e1	f8	98	11	69	d9	8e	94	9b	1e	87	e9	ce	55	28	df
11101	f	8c	a1	89	0d	bf	e6	42	68	41	99	2d	0f	b0	54	bb	16

№ п.п.	Представление входных данных			Представление выходных данных	
	X	Y		X	Y
1	0000000000	00	63	011000011	
2	0000000001	01	7c	011111100	
3	0000000010	02	77	011110111	
4	0000001000	05	f2	111110010	
5	0000010000	08	30	001100000	
6	0000100000	10	ca	110010100	
7	0001000000	20	b7	101110111	
8	0100000000	40	09	000001001	
9	1000000000	80	cd	110001101	
10	1111111111	ff	16	000010110	
11	1111111110	fe	bb	101111011	
12	1111111011	fd	54	010101000	
13	1111110111	fb	0f	000001111	
14	1111101111	f7	68	011001000	
15	1110111111	ef	df	110111111	
16	1101111111	df	9e	100111110	
17	1011111111	bf	08	000001000	
18	0111111111	7f	d2	110100010	

Рис. 9. Пример применения предлагаемого подхода к S-блоку алгоритма шифрования AES

В ходе исследования были проанализированы (согласно выражению (5)) 18 значений блока подстановок алгоритма шифрования AES. Как видно из рисунка 9, для анализа считываются

по 5 значений ячеек из старшей и младшей строки, а также по 4 значения ячеек из младшего и старшего столбцов.

Таким образом, для рассмотренного примера выражения (1) и (5) имеют вид:

$$P = (n, r, m; c, d) = (8, 1, 8; 8, 8), \quad \sum_{i=1}^r a(S_i) = 2 \cdot r \cdot (1 + c) = 18.$$

По считанным указанным способом значениям происходит первичная идентификация блока подстановок.

Далее происходит последовательное считывание каждой ячейки блока подстановок с верификацией исследуемого блока подстановок или с пополнением базы данных алгоритмов шифрования (в случае, когда исследуемый алгоритм не идентифицирован как известный).

Проведённое авторами исследование не зависит от наличия информации о нумерации бит на входе в блок подстановок, так как «протягивание» сначала единицы и потом нуля по входным данным происходит по всем их разрядам. Эта нумерация, как и нумерация битов выходных значений блока подстановок, восстанавливается в ходе проведения исследований.

Результаты исследований на примере анализа блока подстановок алгоритма AES показали, что применение разработанного подхода позволяет однозначно верифицировать тестируемый блок подстановок с каноническим блоком подстановок алгоритма AES [6], что подтвердило работоспособность предложенных авторами решений.

### **7. Оценивание результативности применения подхода.**

Результаты применения предлагаемого подхода к исследованию аппаратных реализаций функционально-структурных элементов блочных алгоритмов шифрования для общего случая проиллюстрированы рисунком 10.

На рисунке 10 (в отличие от рисунка 4) показано, что в результате исследователю становятся известны параметры из выражения (1): размер входного блока  $n$ , количество S-блоков  $r$ , размер выходного блока  $m$ , размер входных данных на каждом S-блоке  $c$ , размер выходных данных на каждом S-блоке  $d$ . После чего проводится проверка равенства (4).

Кроме этого, в результате становится известна нумерация бит в шине данных, через которую происходит информационное взаимодействие других функционально-структурных элементов исследуемого алгоритма с блоками подстановок.

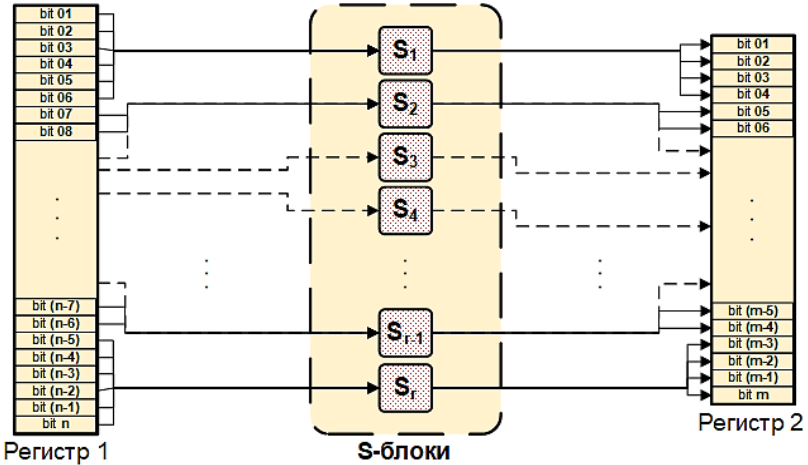


Рис. 10. Схема общего случая использования блоков перестановок в блочных шифрах с их выявленными параметрами и характеристиками

В таблице 1 представлены результаты сравнительного анализа предложенного авторами подхода и подходов, рассмотренных в разделе 2 настоящей статьи [9 – 15, 22 – 31].

Основной отличительной особенностью настоящего подхода является применение инвазивных методов исследования. Недостатком такого подхода является высокая стоимость, а достоинством – наличие полного доступа к анализу всех аппаратно-реализованных функционально-структурных блоков исследуемого устройства (в зависимости от назначения управляющей программы (так называемой «прошивки») – она может активировать, задействовать не полный перечень аппаратно-реализованного функционала устройства).

Строка 10 таблицы 1 показывает: разработанный авторами подход основан на применении методов онтологического представления знаний и разработки и применения правил для машины логического вывода, что позволяет выявлять блоки эквивалентных преобразований, конструировать новые правила для выявления подобных и других функционально-структурных блоков, использовать базу знаний исследованных ранее функционально-структурных спецификаций ЗИС. Проанализированные подходы методы онтологического представления знаний не применяют.



Таблица 1. Результаты сравнительного анализа разработанного подхода и релевантных работ

	Исследованные труды (нумерация согласно списку литературы)	[9]	[10]	[11]	[12]	[13]	[22]	[23]	[24]	[25]	[26]	[27]	[28]	[29]	[30]	[31]	Предложенный подход
		Применяемые методы															
1.	Общие принципы проектирования защищённых систем	+	+	+													
2.	Методы машинного обучения									+		+				+	+
3.	Методы тестовых оценок и генератора на их основе тестовых сценариев						+										
4.	Методы анализа обфусцированного программного кода														+		
5.	Методы сопоставления семантических аспектов вычислений					+											
6.	Методы анализа побочных каналов данных									+		+	+				
7.	Методы моделирования процесса обратного проектирования										+						
8.	Методы анализа САПР, которыми разрабатываются микросхемы												+				+
9.	Методы моделирования и анализа графов исследуемых объектов					+			+	+							+
10.	Методы онтологического представления знаний и разработки и применения правил для машины логического вывода																+
11.	Учёт особенностей построения криптографических алгоритмов ЗИ												+				+
12.	Неинвазивные методы					+	+	+	+	+	+	+	+	+	+	+	+
13.	Инвазивные методы																+

В статье [27] авторы рассмотрели в качестве примера применения своего подхода, основанного на методах анализа

побочных каналов данных (то есть – это тоже неинвазивные методы) и методах машинного обучения, один из режимов алгоритма AES. В отличие от такого подхода, настоящее исследование направлено на выявление особенностей аппаратных реализаций именно алгоритмов шифрования, иными словами – на создание модели функционально-структурных спецификаций аппаратных реализаций алгоритмов шифрования.

Проанализированные подходы представлены в таблице 1 в соответствии с применяемыми в них методами исследования. Стоит отметить, что в качестве примеров применения этих подходов были рассмотрены различные промышленные отрасли: авиастроение, автомобилестроение и другие; среди которых защиты информации криптографическими методами рассмотрено не было.

Интерес представляют исследования [27] и [31], которые рассматривают вредоносные аппаратные модификации, которые могут внедряться средствами систем автоматизированного проектирования (САПР) на этапе проектирования.

Таким образом, представленный подход может быть применён при проведении верификации функционально-структурных спецификаций или при проведении обратного проектирования интегральных микросхем. Особенностью настоящего подхода является его нацеленность на исследование аппаратных реализаций алгоритмов шифрования.

Применение данного подхода (как инвазивного метода исследования) в совокупности с неинвазивными методами даст полное представление о функционально-структурном наполнении исследуемого устройства, так как позволит выявить неиспользуемый управляющими программами аппаратно реализованный функционал и подтвердить или опровергнуть выдвинутые некоторыми статистическими неинвазивными методами гипотезы или выводы.

**8. Заключение.** Представлен подход к верификации функционально-структурных спецификаций, реализованных в заказных интегральных схемах.

Подтверждена работоспособность настоящего подхода на примерах исследования аппаратных реализаций алгоритмов шифрования DES и AES: начиная с процесса аппаратного обратного проектирования указанных алгоритмов (идентификации реализованных в заказных интегральных схемах функционально-структурных блоков) до верификации их функционально-логического наполнения.

Подтверждена работоспособность разработанного подхода для выявления нумерации исследуемых блоков подстановок и нумерации бит внутри каждого блока подстановки в отдельности.

Направлением будущей работы является проведение исследований по выявлению и локализации на стадии предварительных исследований функционально-структурных блоков, реализующих стандартные процессоры, восстановление и верификация которых может не представлять интереса.

### Литература

1. Нагибин Д.В., Платонов А.А., Сабиров Т.Р. Методика идентификации реализованных в заказных интегральных схемах функционально-структурных спецификаций // Труды Военно-космической академии имени А.Ф. Можайского. 2024. № 690. С. 112–120.
2. Mustafa Dhiaa Al-Hassan, Qusay Zuhair Abdulla. Robust Password Encryption Technique with an Extra Security Layer // Iraqi Journal of Science. 2023. vol. 64. no. 3. pp. 1477–1486. DOI: 10.24996/ij.s.2023.64.3.36.
3. Alsuwaiedi H.K.A., Rahma A.M.S. A new modified DES algorithm based on the development of binary encryption functions // Journal of King Saud University – Computer and Information Sciences. 2023. vol. 35(8). DOI: 10.1016/j.jksuci.2023.101716.
4. Agate V., Concone F., de Paola A., Ferraro P., Lo Re G., Morana M. Bayesian Modeling for Differential Cryptanalysis of Block Ciphers: A DES Instance // IEEE Access. 2023. vol. 11. pp. 4809–4820. DOI: 10.1109/ACCESS.2023.3236240.
5. Wu Y., Dai X. Encryption of accounting data using DES algorithm in Computing Environment // Journal of Intelligent & Fuzzy Systems. 2020. vol. 39. pp. 5085–5095. DOI: 10.3233/JIFS-179994.
6. Nitaj A., Susilo W., Tonien J. Enhanced S-boxes for the Advanced Encryption Standard with maximal periodicity and better avalanche property // Computer Standards & Interfaces. 2024. vol. 87. DOI: 10.1016/j.csi.2023.103769.
7. Zahid A.H., Arshad M.J. An innovative design of substitution-boxes using cubic polynomial mapping // Symmetry. 2019. vol. 11(3). DOI: 10.3390/sym11030437.
8. Zahid A.H., Arshad M.J., Ahmad M. Substitution-boxes using cubic fractional transformation // Entropy. 2019. vol. 21(3). DOI: 10.3390/e21030245.
9. Котенко И.В., Левшун Д.С., Чечулин А.А., Ушаков И.А., Красов А.В. Комплексный подход к обеспечению безопасности киберфизических систем на системе микроконтроллеров // Вопросы кибербезопасности. 2018. № 3(27). С. 29–38. DOI: 10.21681/2311-3456-2018-3-29-38.
10. Десницкий В.А., Чечулин А.А., Котенко И.В., Левшун Д.С., Коломеец М.В. Комбинированная методика проектирования защищенных встроенных устройств на примере системы охраны периметра // Труды СПИИРАН. 2016. № 5(48). С. 5–31. DOI: 10.15622/sp.48.1.
11. Desnitsky V., Levshun D., Chechulin A., Kotenko I. Design technique for secure embedded devices: application for creation of integrated cyber-physical security system // Journal of Wireless Mobile Networks, Ubiquitous Computing, and Dependable Applications (JoWUA). 2016. vol. 7. pp. 60–80.
12. Ковалев В.В., Компаниец Р.И., Новиков В.А. Верификация программ на основе соотношений подобия // Труды СПИИРАН. 2015. vol. 1(38). С. 233–245. DOI: 10.15622/sp.38.13.

13. Cousot P., Cousot R. A Galois connection calculus for abstract interpretation // Proceedings of the 41st Annual ACM SIGPLAN-SIGACT Symposium on Principles of Programming Languages (POPL'14). ACM, 2014. pp. 3–4. DOI: 10.1145/2535838.2537850.
14. Fang Z., Darais D., Near J.P., Zhang Y. Zero Knowledge Static Program Analysis // Proceedings of the 2021 ACM SIGSAC Conference on Computer and Communications Security (CCS'21). New York, USA: Association for Computing Machinery, 2021. pp. 2951–2967. DOI: 10.1145/3460120.3484795.
15. Котенко И.В., Резник С.А., Шоров А.В. Верификация протоколов безопасности на основе комбинированного использования существующих методов и средств // Труды СПИИРАН. 2009. № 8. С. 292–310. DOI: 10.15622/sp.8.14.
16. Tran D.D., Ogata K. IPSG: Invariant Proof Score Generator // 2022 IEEE 46th Annual Computers, Software, and Applications Conference (COMPSAC). 2022. pp. 1050–1055. DOI: 10.1109/COMPSAC54236.2022.00164.
17. Riesco A., Ogata K., Futatsugi K. A Maude environment for cafeOBJ // Formal Aspects of Computing. 2017. vol. 29. pp. 309–334. DOI: 10.1007/s00165-016-0398-7.
18. Clavel M., Durán F., Eker S., Lincoln P., Marti-Oliet N., Meseguer J., Talcott C. All about Maude – a High-Performance Logical Framework. How to Specify, Program and Verify Systems in Rewriting Logic // Lecture Notes in Computer Science. 2007. 802 p. DOI: 10.1007/978-3-540-71999-1.
19. Tran D.D., Ogata K. Formal verification of TLS 1.2 by automatically generating proof scores // Computers & Security. 2022. vol. 123. DOI: 10.1016/j.cose.2022.102909.
20. Saeed Mahlaqa, Ibrar Muhammad, Mahmood Dua, Delshadi Amirmohammad, Saleemi Aqdas. Enhancing Computer Security through Formal Verification of Cryptographic Protocols Using Model Checking and Partial Order Techniques // The Asian Bulletin of Big Data Management. 2024. vol. 4(2). pp. 225–238. DOI: 10.62019/abbdm.v4i02.176.
21. Curaba C., D'Ambrosi D., Minisini A., Antol'in N.P.-C. CryptoFormalEval: Integrating LLMs and Formal Verification for Automated Cryptographic Protocol Vulnerability Detection. 2024. arXiv preprint arXiv:2411.13627.
22. Riesco A., Ogata K. An integrated tool set for verifying CafeOBJ specifications // Journal of Systems and Software. 2022. vol. 189. DOI: 10.1016/j.jss.2022.111302.
23. Awadhutkar P., Tamrawi A., Goluch R., Kothari S. Control flow equivalence method for establishing sanctity of compiling // Computers & Security. 2022. vol. 115. DOI: 10.1016/j.cose.2022.102608.
24. Cheng D., Dong Ch., He W., Chen Zh., Liu X., Zhang H. A fine-grained detection method for gate-level hardware Trojan based on bidirectional Graph Neural Networks // Journal of King Saud University – Computer and Information Sciences. 2023. vol. 35(10). DOI: 10.1016/j.jksuci.2023.101822.
25. Chen Sh., Wang T., Huang Zh., Hou X. Detection method of Golden Chip-Free Hardware Trojan based on the combination of ResNeXt structure and attention mechanism // Computers & Security. 2023. vol. 134. DOI: 10.1016/j.cose.2023.103428.
26. Rozesara M., Ghazinoori S., Manteghi M., Tabatabaeian S.H. A reverse engineering-based model for innovation process in complex product systems: Multiple case studies in the aviation industry // Journal of Engineering and Technology Management. 2023. vol. 69. DOI: 10.1016/j.jengtecman.2023.101765.
27. Lavanya T., Rajalakshmi K. Heterogenous ensemble learning driven multi-parametric assessment model for hardware Trojan detection // Integration. 2023. vol. 89. pp. 217–228. DOI: 10.1016/j.vlsi.2022.12.011.

28. Esirci F.N., Bayrakci A.A. Delay based hardware Trojan detection exploiting spatial correlations to suppress variations // *Integration*. 2023. vol. 91. pp. 107–118. DOI: 10.1016/j.vlsi.2023.03.006.
29. Cassano L., Iamundo M., Lopez T.A., Nazzari A., Di Natale G. DETON: DEfeating hardware Trojan horses in microprocessors through software Obfuscation // *Journal of Systems Architecture*. 2022. vol. 129. DOI: 10.1016/j.sysarc.2022.102592.
30. Damljanovic A., Ruospo A., Sanchez E., Squillero G. Machine learning for hardware security: Classifier-based identification of Trojans in pipelined microprocessors // *Applied Soft Computing*. 2022. vol. 116. DOI: 10.1016/j.asoc.2021.108068.
31. Palumbo A., Cassano L., Luzzi B., Hernández J.A., Reviriego P., Bianchi G., Ottavi M. Is your FPGA bitstream Hardware Trojan-free? Machine learning can provide an answer // *Journal of Systems Architecture*. 2022. vol. 128. DOI: 10.1016/j.sysarc.2022.102543.
32. Токарева Н.Н. Симметричная криптография. Краткий курс: учебное пособие / Под ред. А.Л. Перегожина // Новосибирск: Новосиб. гос. Ун-т., 2012. 234 с.
33. Шнайер Б. Прикладная криптография. Протоколы, алгоритмы и исходный код на С. М.: Диалектика. 2017. 1040 с.
34. Макаренко С.И. Информационное противоборство и радиоэлектронная борьба в сетевых войнах начала XXI века. Санкт-Петербург: Научное издательство «Лань», 2017. 549 с. URL: [http://psyfactor.org/t/Makarenko-InfPro\\_2017.pdf](http://psyfactor.org/t/Makarenko-InfPro_2017.pdf).
35. Elçi A., Pieprzyk J., Chefranov A.G., Orgun M.A., Wang H., Shankaran R. *Theory and Practice of Cryptography Solutions for Secure Information Systems*. Hershey, PA: IGI Global Scientific Publishing, 2013. 351 p. DOI: 10.4018/978-1-4666-4030-6.
36. Smart N.P. *Cryptography Made Simple. Information Security and Cryptography series*. Springer, 2016. 481 p. DOI: 10.1007/978-3-319-21936-3.
37. Бабенко Л.К., Ищукова Е.А. Современные алгоритмы блочного шифрования и методы их анализа. М.: Гелиос АРВ. 2006. 376 с.

**Нагибин Дмитрий Владимирович** — преподаватель, кафедра систем сбора и обработки информации, Федеральное государственное бюджетное военное образовательное учреждение высшего образования «Военно-космическая академия имени А.Ф. Можайского» Министерства обороны Российской Федерации. Область научных интересов: обратное проектирование аппаратных и программно-аппаратных комплексов, методы верификации сложных технических устройств, криптографические методы защиты информации. Число научных публикаций — 25. [nagibin.86@internet.ru](mailto:nagibin.86@internet.ru); улица Ждановская, 13, 197198, Санкт-Петербург, Россия; р.т.: +7(950)043-0162.

**Петренко Алексей Сергеевич** — инженер-исследователь, Федеральное государственное бюджетное военное образовательное учреждение высшего образования «Военно-космическая академия имени А.Ф. Можайского» Министерства обороны Российской Федерации. Область научных интересов: информационная безопасность, криптография, постквантовая криптография, квантовые вычисления, блокчейн, искусственный интеллект. Число научных публикаций — 280. [a.petrenko1999@rambler.ru](mailto:a.petrenko1999@rambler.ru); улица Ждановская, 13, 197198, Санкт-Петербург, Россия; р.т.: +7(999)239-3901.

**Давыденко Владислав Сергеевич** — курсант, кафедра систем сбора и обработки информации, Федеральное государственное бюджетное военное образовательное учреждение высшего образования «Военно-космическая академия имени А.Ф. Можайского» Министерства обороны Российской Федерации. Область научных интересов: криптографические методы защиты информации, математические методы в криптографии. Число научных публикаций — 5. [nagibin.86@internet.ru](mailto:nagibin.86@internet.ru); улица Ждановская, 13, 197198, Санкт-Петербург, Россия; р.т.: +7(950)043-0162.

**Котенко Игорь Витальевич** — д-р техн. наук, профессор, заслуженный деятель науки Российской Федерации, главный научный сотрудник, руководитель лаборатории, лаборатория проблем компьютерной безопасности, Федеральное государственное бюджетное учреждение науки «Санкт-Петербургский Федеральный исследовательский центр Российской академии наук» (СПб ФИЦ РАН). Область научных интересов: безопасность компьютерных сетей, в том числе управление политиками безопасности, разграничение прав доступа, аутентификация, анализ защищенности, обнаружение компьютерных атак, межсетевые экраны, защита от вирусов и сетевых червей, анализ и верификация протоколов безопасности и систем защиты информации, защита программного обеспечения от взлома и управление цифровыми правами, технологии моделирования и визуализации для противодействия кибер-терроризму. Число научных публикаций — 850. ivkote@comsec.spb.ru; 14-я линия В.О., 39, 199178, Санкт-Петербург, Россия; р.т.: +7(812)328-7181; факс: +7(812)328-4450.

**Федорченко Елена Владимировна** — канд. техн. наук, доцент, старший научный сотрудник, лаборатория проблем компьютерной безопасности, Федеральное государственное бюджетное учреждение науки «Санкт-Петербургский Федеральный исследовательский центр Российской академии наук» (СПб ФИЦ РАН). Область научных интересов: безопасность информационных систем, методы анализа рисков компьютерных сетей, управление информационными рисками, анализ данных, поддержка принятия решений по повышению защищенности. Число научных публикаций — 137. doynikova@comsec.spb.ru; 14-я линия В.О., 39, 199178, Санкт-Петербург, Россия; р.т.: +7(812)328-7181.

**Поддержка исследований.** Работа выполнена при частичной финансовой поддержке бюджетной темы FFZF-2025-0016.

D. NAGIBIN, A. PETRENKO, V. DAVYDENKO, I. KOTENKO, E. FEDORCHENKO  
**INVASIVE APPROACH TO VERIFICATION OF FUNCTIONAL  
AND STRUCTURAL SPECIFICATIONS IMPLEMENTED  
IN CUSTOM INTEGRATED CIRCUITS**

*Nagibin D., Petrenko A., Davydenko V., Kotenko I., Fedorchenko E.* **Invasive Approach to Verification of Functional and Structural Specifications Implemented in Custom Integrated Circuits.**

**Abstract.** An approach to verification of functional and structural specifications implemented in custom integrated circuits based on invasive research methods is presented. The relevance of this research is determined by the necessity of verification of functional-structural specifications supplied by third-party implementers of hardware implementations of information security algorithms, the difficulty of detecting modifications of these algorithms and undocumented capabilities implemented at the hardware level, and the lack of uniform, universal or standardized methods for solving this problem. The mathematical formulation of the research problem is specified; its essence is to verify the equality of the values of the declared specification parameters and their values restored by the reverse engineering method. The results of the application of the verification technique of functional and structural specifications are presented using examples of its adaptation to the study of hardware-implemented DES and AES encryption algorithms. The restored functional and structural blocks of the algorithms (in particular, the substitution block) were successfully verified.

**Keywords:** custom integrated circuit, identification, verification, functional and structural specifications, cryptographic algorithms.

## References

1. Nagibin D.V., Platonov A.A., Sabirov T.R. [Technique for identification of functional-structural specifications realized in custom integrated circuits]. Trudy Voenno-kosmicheskoy akademii imeni A.F. Mozhajskogo – Proceedings of the Mozhaisky Military Space Academy. 2024. no. 690, pp. 112–120. (In Russ.).
2. Mustafa Dhiaa Al-Hassan, Qusay Zuhair Abdulla. Robust Password Encryption Technique with an Extra Security Layer. Iraqi Journal of Science. 2023. vol. 64. no. 3. pp. 1477–1486. DOI: 10.24996/ij.s.2023.64.3.36.
3. Alsuaiedi H.K.A., Rahma A.M.S. A new modified DES algorithm based on the development of binary encryption functions. Journal of King Saud University – Computer and Information Sciences. 2023. vol. 35(8). DOI: 10.1016/j.jksuci.2023.101716.
4. Agate V., Concone F., de Paola A., Ferraro P., Lo Re G., Morana M. Bayesian Modeling for Differential Cryptanalysis of Block Ciphers: A DES Instance. IEEE Access. 2023. vol. 11. pp. 4809–4820. DOI: 10.1109/ACCESS.2023.3236240.
5. Wu Y., Dai X. Encryption of accounting data using DES algorithm in Computing Environment. Journal of Intelligent & Fuzzy Systems. 2020. vol. 39. pp. 5085–5095. DOI: 10.3233/JIFS-179994.
6. Nitaj A., Susilo W., Tonien J. Enhanced S-boxes for the Advanced Encryption Standard with maximal periodicity and better avalanche property. Computer Standards & Interfaces. 2024. vol. 87. DOI: 10.1016/j.csi.2023.103769.
7. Zahid A.H., Arshad M.J. An innovative design of substitution-boxes using cubic polynomial mapping. Symmetry. 2019. vol. 11(3). DOI: 10.3390/sym11030437.

8. Zahid A.H., Arshad M.J., Ahmad M. Substitution-boxes using cubic fractional transformation. *Entropy*. 2019. vol. 21(3). DOI: 10.3390/e21030245.
9. Kotenko I.V., Levshun D.S., Chechulin A.A., Ushakov I.A., Krasov A.V. [Integrated approach to provide security of cyber-physical systems based on microcontrollers]. *Voprosy kiberbezopasnosti – Cybersecurity issues*. 2018. no. 3(27). pp. 29–38. DOI: 10.21681/2311-3456-2018-3-29-38. (In Russ.).
10. Desnitsky V., Chechulin A., Kotenko I., Levshun D., Kolomeec M. [Combined Design Technique for Secure Embedded Devices Exemplified by a Perimeter Protection System]. *SPIIRAS Proceedings*. 2016. no. 5(48). pp. 5–31. DOI: 10.15622/sp.48.1. (In Russ.).
11. Desnitsky V., Levshun D., Chechulin A., Kotenko I. Design technique for secure embedded devices: application for creation of integrated cyber-physical security system. *Journal of Wireless Mobile Networks, Ubiquitous Computing, and Dependable Applications (JoWUA)*. 2016. vol. 7. pp. 60–80.
12. Kovalyov V., Kompaniets R., Novikov V [Verification of Programs Based on Similarity Relations]. *SPIIRAS Proceedings*. 2015. vol. 1(38). pp. 233–245. DOI: 10.15622/sp.38.13. (In Russ.).
13. Cousot P., Cousot R. A Galois connection calculus for abstract interpretation. *Proceedings of the 41st Annual ACM SIGPLAN-SIGACT Symposium on Principles of Programming Languages (POPL'14)*. ACM, 2014. pp. 3–4. DOI: 10.1145/2535838.2537850.
14. Fang Z., Darais D., Near J.P., Zhang Y. Zero Knowledge Static Program Analysis. *Proceedings of the 2021 ACM SIGSAC Conference on Computer and Communications Security (CCS'21)*. New York, USA: Association for Computing Machinery, 2021. pp. 2951–2967. DOI: 10.1145/3460120.3484795.
15. Kotenko I., Reznik S., Shorov A. [Security protocols verification combining existing approaches and tools]. *SPIIRAS Proceedings*. 2009. no. 8. pp. 292–310. DOI: 10.15622/sp.8.14. (In Russ.).
16. Tran D.D., Ogata K. IPSG: Invariant Proof Score Generator. 2022 IEEE 46th Annual Computers, Software, and Applications Conference (COMPSAC). 2022. pp. 1050–1055. DOI: 10.1109/COMPSAC54236.2022.00164.
17. Riesco A., Ogata K., Futatsugi K. A Maude environment for cafeOBJ. *Formal Aspects of Computing*. 2017. vol. 29. pp. 309–334. DOI: 10.1007/s00165-016-0398-7.
18. Clavel M., Durán F., Eker S., Lincoln P., Marti-Oliet N., Meseguer J., Talcott C. All about Maude – a High-Performance Logical Framework. How to Specify, Program and Verify Systems in Rewriting Logic. *Lecture Notes in Computer Science*. 2007. 802 p. DOI: 10.1007/978-3-540-71999-1.
19. Tran D.D., Ogata K. Formal verification of TLS 1.2 by automatically generating proof scores. *Computers & Security*. 2022. vol. 123. DOI: 10.1016/j.cose.2022.102909.
20. Saeed Mahlaqa, Ibrar Muhammad, Mahmood Dua, Delshadi Amirmohammad, Saleemi Aqdas. Enhancing Computer Security through Formal Verification of Cryptographic Protocols Using Model Checking and Partial Order Techniques. *The Asian Bulletin of Big Data Management*. 2024. vol. 4(2). pp. 225–238. DOI: 10.62019/abbdm.v4i02.176.
21. Curaba C., D'Ambrosi D., Minisini A., Antol'in N.P.-C. CryptoFormalEval: Integrating LLMs for Automated Verification of Automated Cryptographic Protocol Vulnerability Detection. 2024. arXiv preprint arXiv:2411.13627.
22. Riesco A., Ogata K. An integrated tool set for verifying CafeOBJ specifications. *Journal of Systems and Software*. 2022. vol. 189. DOI: 10.1016/j.jss.2022.111302.
23. Awadhutkar P., Tamrawi A., Goluch R., Kothari S. Control flow equivalence method for establishing sanctity of compiling. *Computers & Security*. 2022. vol. 115. DOI: 10.1016/j.cose.2022.102608.



24. Cheng D., Dong Ch., He W., Chen Zh., Liu X., Zhang H. A fine-grained detection method for gate-level hardware Trojan based on bidirectional Graph Neural Networks. *Journal of King Saud University – Computer and Information Sciences*. 2023. vol. 35(10). DOI: 10.1016/j.jksuci.2023.101822.
25. Chen Sh., Wang T., Huang Zh., Hou X. Detection method of Golden Chip-Free Hardware Trojan based on the combination of ResNeXt structure and attention mechanism. *Computers & Security*. 2023. vol. 134. DOI: 10.1016/j.cose.2023.103428.
26. Rozesara M., Ghazinoori S., Manteghi M., Tabatabaeian S.H. A reverse engineering-based model for innovation process in complex product systems: Multiple case studies in the aviation industry. *Journal of Engineering and Technology Management*. 2023. vol. 69. DOI: 10.1016/j.jengetecman.2023.101765.
27. Lavanya T., Rajalakshmi K. Heterogenous ensemble learning driven multi-parametric assessment model for hardware Trojan detection. *Integration*. 2023. vol. 89. pp. 217–228. DOI: 10.1016/j.vlsi.2022.12.011.
28. Esirci F.N., Bayrakci A.A. Delay based hardware Trojan detection exploiting spatial correlations to suppress variations. *Integration*. 2023. vol. 91. pp. 107–118. DOI: 10.1016/j.vlsi.2023.03.006.
29. Cassano L., Iamundo M., Lopez T.A., Nazzari A., Di Natale G. DETON: DDefeating hardware Trojan horses in microprocessors through software Obfuscation. *Journal of Systems Architecture*. 2022. vol. 129. DOI: 10.1016/j.sysarc.2022.102592.
30. Damljanovic A., Ruospo A., Sanchez E., Squillero G. Machine learning for hardware security: Classifier-based identification of Trojans in pipelined microprocessors. *Applied Soft Computing*. 2022. vol. 116. DOI: 10.1016/j.asoc.2021.108068.
31. Palumbo A., Cassano L., Luzzi B., Hernández J.A., Reviriego P., Bianchi G., Ottavi M. Is your FPGA bitstream Hardware Trojan-free? Machine learning can provide an answer. *Journal of Systems Architecture*. 2022. vol. 128. DOI: 10.1016/j.sysarc.2022.102543.
32. Tokareva N.N. Simmetrichnaya kriptografija. Kratkij kurs: uchebnoe posobie [Symmetric cryptography (Symmetric cryptography). Short Course: Study Guide]. Novosibirsk: Novosib. State. Un-ty., 2012. 234 p. (In Russ.).
33. Schneier B. Protokoly, algoritmy i ishodnyj kod na S [Applied cryptography. Protocols, algorithms and source code in C Прикладная криптография]. М.: Дialektika. 2017. 1040 p. (In Russ.).
34. Makarenko S.I. Informacionnoe protivoborstvo i radioelektronnaja bor'ba v setecentricheskix vojnax nachala XXI veka [Information warfare and electronic warfare in network-centric warfare of the early 21st century]. Sankt-Peterburg: Naukoemkie tehnologii, 2017. 549 p. (In Russ.).
35. Elçi A., Pieprzyk J., Chefranov A.G., Orgun M.A., Wang H., Shankaran R. *Theory and Practice of Cryptography Solutions for Secure Information Systems*. Hershey, PA: IGI Global Scientific Publishing, 2013. 351 p. DOI: 10.4018/978-1-4666-4030-6.
36. Smart N.P. *Cryptography Made Simple. Information Security and Cryptography series*. Springer, 2016. 481 p. DOI: 10.1007/978-3-319-21936-3.
37. Babenko L.K., Ishhukova E.A. *Sovremennye algoritmy blochnogo shifrovaniya i metody ih analiza [Modern block encryption algorithms and methods of their analysis]*. М.: Gelios ARV, 2006. 376 p. (In Russ.).

**Nagibin Dmitry** — Lecturer, Department of information collection and processing systems, Federal State Budgetary Military Educational Institution of Higher Education «Military Space Academy named after A.F. Mozhaisky» of the Ministry of Defense of the Russian Federation. Research interests: reverse engineering of hardware and hardware-software complexes, verification methods of complex technical devices, cryptographic methods of information

protection. The number of publications — 25. nagibin.86@internet.ru; 13, Zhdanovskaya St., 197198, St. Petersburg, Russia; office phone: +7(950)043-0162.

**Petrenko Alexey** — Research engineer, Federal State Budgetary Military Educational Institution of Higher Education «Military Space Academy named after A.F. Mozhaisky» of the Ministry of Defense of the Russian Federation. Research interests: information security, cryptography, post-quantum cryptography, quantum computing, blockchain, artificial intelligence. The number of publications — 280. a.petrenko1999@rambler.ru; 13, Zhdanovskaya St., 197198, St. Petersburg, Russia; office phone: +7(999)239-3901.

**Davydenko Vladislav** — Cadet, Department of information collection and processing systems, Federal State Budgetary Military Educational Institution of Higher Education «Military Space Academy named after A.F. Mozhaisky» of the Ministry of Defense of the Russian Federation. Research interests: cryptographic methods of information security, mathematical methods in cryptography. The number of publications — 5. nagibin.86@internet.ru; 13, Zhdanovskaya St., 197198, St. Petersburg, Russia; office phone: +7(950)043-0162.

**Kotenko Igor** — Ph.D., Dr.Sci., Professor, Honored scientist of the Russian Federation, Chief researcher, head of the laboratory, Laboratory of computer security problems, St. Petersburg Federal Research Center of the Russian Academy of Sciences (SPC RAS). Research interests: computer network security, including security policy management, access control, authentication, network security analysis, intrusion detection, firewalls, deception systems, malware protection, verification of security systems, digital right management, modeling, simulation and visualization technologies for counteraction to cyber terrorism. The number of publications — 850. ivkote@comsec.spb.ru; 39, 14-th Line V.O., 199178, St. Petersburg, Russia; office phone: +7(812)328-7181; fax: +7(812)328-4450.

**Fedorchenko Elena** — Ph.D., Associate Professor, Senior researcher, Laboratory of computer security problems, St. Petersburg Federal Research Center of the Russian Academy of Sciences (SPC RAS). Research interests: information systems security, risk analysis methods for computer networks, information security risk management, data analysis, security decision support. The number of publications — 137. doynikova@comsec.spb.ru; 39, 14-th Line V.O., 199178, St. Petersburg, Russia; office phone: +7(812)328-7181.

**Acknowledgements.** The reported study was partially funded by the budget project FFZF-2025-0016.

N.V. HUNG, N.A. QUAN, N. TAN, T.T. HAI, D.K. TRUNG, L.M. NAM,  
B.T. LOAN, N.T.T. NGA

## BUILDING PREDICTIVE SMELL MODELS FOR VIRTUAL REALITY ENVIRONMENTS

---

*Nguyen Viet Hung, Nguyen Anh Quan, Nguyen Tan, Tran Trieu Hai, Dang Khanh Trung, Le Mai Nam, Bui Thanh Loan, Nguyen Thi Thuy Nga.* **Building Predictive Smell Models for Virtual Reality Environments.**

**Abstract.** In a sensory-rich environment, human experiences are shaped by the complex interplay of multiple senses. However, digital interactions predominantly engage visual and auditory modalities, leaving other sensory channels, such as olfaction, largely unutilized. Virtual Reality (VR) technology holds significant potential for addressing this limitation by incorporating a wider range of sensory inputs to create more immersive experiences. This study introduces a novel approach for integrating olfactory stimuli into VR environments through the development of predictive odor models, termed SPRF (Sensory Predictive Response Framework). The objective is to enhance the sensory dimension of VR by tailoring scent stimuli to specific content and context with the collection of information about the location of scent sources and their identification through features to serve to reproduce them in the space of the VR environment, thereby enriching user engagement and immersion. Additionally, the research investigates the influence of various scent-related factors on user perception and behavior in VR, aiming to develop predictive models optimized for olfactory integration. Empirical evaluations demonstrate that the SPRF model achieves superior performance, with an accuracy of 98.13%, significantly outperforming conventional models such as Convolutional Neural Networks (CNN, 79.46%), Long Short-Term Memory (LSTM, 80.37%), and Support Vector Machines (SVM, 85.24%). Additionally, SPRF delivers notable improvements in F1-scores (13.05%-21.38%) and accuracy (12.89%-18.67%) compared to these alternatives. These findings highlight the efficacy of SPRF in advancing olfactory integration within VR, offering actionable insights for the design of multisensory digital environments.

**Keywords:** virtual reality, odor, model selection, user experience, imagination, odor prediction.

---

**1. Introduction.** The rapid expansion of virtual worlds and advancements in 3D space technology have ushered in a new era of human interaction and perception. As research in these fields progresses, the integration of virtual environments with real-life experiences becomes increasingly significant [1]. This evolution reflects a growing interest in enhancing human vision and creating immersive experiences that bridge the gap between virtual and real worlds. Consequently, there is a concerted effort to explore how these developments can enrich human life and foster a future characterized by greater enjoyment and connectivity [2].

Moreover, today's multisensory digital experiences have enhanced human interaction with technology, with the aim of replicating real-world sensory perceptions. The objective is to integrate human senses into digital environments, creating a seamless and immersive experience. However,

delivering a comprehensive sensory digital experience poses significant challenges due to various influencing factors. This endeavor is prioritized in the future technological development [3]. Current digital techniques often lack full auditory stimulation and integration with other senses, which presents a major challenge in increasing consumer value [4]. Besides, the future expansion of VR systems has captured the keen interest of researchers, making studies on transmission and prediction increasingly essential. In particular, the ability to predict and simulate odors in virtual environments is expected to play a pivotal role, especially in enhancing the online transmission of 360-degree videos [5-7]. This area of research is not only necessary but also holds immense potential to revolutionize immersive experiences in the future.

The integration of olfactory elements into virtual reality (VR) environments marks a groundbreaking advancement in immersive technology, addressing a sensory dimension that has traditionally been overlooked in digital spaces. Olfaction, with its profound influence on human perception, memory, and emotion, offers significant potential to enhance user immersion and realism in virtual environments. Recent technological advancements have enabled the incorporation of scents into VR, thereby providing a more holistic sensory experience. While much of the recent research in VR has focused on areas such as viewport position prediction and the evaluation of 360-degree video streaming quality [8-11], studies on integrating olfactory stimuli into VR remain limited. However, predictive modeling, as demonstrated in studies such as [9, 10], plays a critical role in enhancing VR experiences by anticipating user interactions and optimizing immersion. Building on these advancements, this work seeks to enhance user perception in virtual environments by combining predictive modeling with engaging olfactory stimuli. By doing so, we aim to create a novel and immersive virtual atmosphere that deepens realism and enriches the overall user experience.

The proposed scent recognition system in virtual reality works by collecting and analyzing scents in the surrounding space through customized electronic noses integrated directly under the virtual reality glasses with gas sensors used in the e-nose such as MP503, BME680, MQ3, MQ5, MQ9 and WSP2110 it is shown in Figure 1. When the user wears the glasses, these electronic noses continuously scan the environment, detecting and collecting scent molecules in the air. The collected data is then processed by advanced algorithms to classify the scent, identify the characteristics of the scent, and accurately predict the distance from the source to the location of the VR glasses. The system also takes into account environmental factors that may affect the diffusion of scents, such as wind speed and direction, temperature, humidity, and air quality. As a result, the algorithm can calibrate parameters to ensure

that the prediction of the scent's location is simulated as accurately as possible. When a scent source is detected, the system displays or recreates its location in the virtual reality environment, providing the user with a more realistic experience. With this technology, users in virtual reality environments can not only see and hear but also intuitively perceive scents, opening up many potential applications in the fields of entertainment, education, scientific research and even environmental investigation.

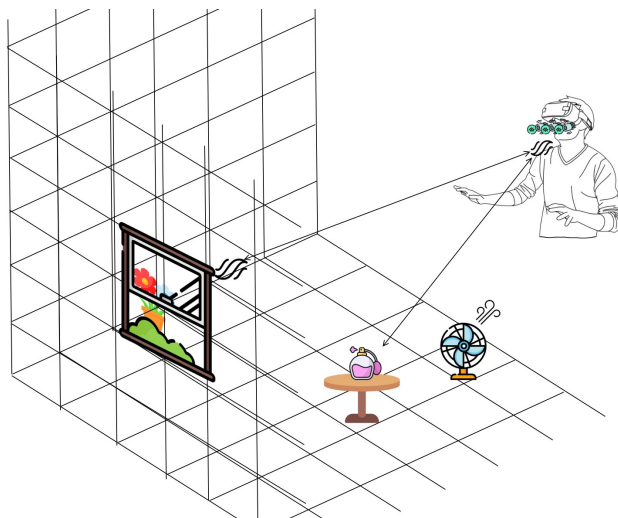


Fig. 1. System in a virtual reality environment

Predictive odor models are at the forefront of this innovation, providing the ability to simulate realistic olfactory experiences using complex algorithms and extensive scent databases. These models work by analyzing the chemical compositions of odors and predicting their perceptual attributes, thereby enhancing the realism and interactivity of VR applications. Such applications span diverse fields, including gaming, education, therapeutic interventions, and marketing strategies.

Figure 2 shows an architectural model for developing predictive odor systems in virtual reality environments. It features four key layers. The Data Collection Layer collects information through sensors and odor data repositories. Moving to the Processing Layer, chemical analysis is conducted alongside machine learning models, such as Support Vector Machines (SVM), to predict odor perceptions. The Integration Layer ensures seamless connectivity between VR software and olfactory display systems, enabling the

emission of odors within virtual environments. Finally, the User-Interaction Layer focuses on incorporating feedback mechanisms and user interfaces to refine and enhance user experiences. The flow between these layers is depicted with arrows, illustrating the sequential process from data collection to user interaction.

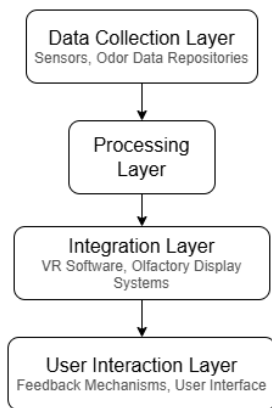


Fig. 2. Architectural model for predictive odor models in VR

The novelty of predictive odor models lies in their scientific approach to synthesizing and delivering scents with precision. Using machine learning techniques and comprehensive olfactory research, these models can accurately reproduce a broad spectrum of odors. This capability is pivotal in overcoming the challenges associated with creating consistent and repeatable olfactory experiences in VR, ensuring that users receive reliable sensory input regardless of the context.

On the one hand, predictive odor models facilitate the creation of personalized and adaptive scent experiences, representing a significant leap toward user-centric virtual environments. By tailoring scent profiles to individual preferences or specific VR scenarios, these models enable a level of customization that was previously unattainable. This personalization not only enhances user engagement but also allows for the exploration of new dimensions in digital interaction.

In addition to improving user experience, predictive odor models have significant potential to advance research on the human olfactory system and its influence on behavior and cognition. By simulating diverse scent scenarios, researchers can investigate psychological and physiological responses to

different odors within controlled virtual settings. This research can lead to a deeper understanding of the interplay between olfaction and various cognitive processes.

Ultimately, the development of predictive odor models represents a substantial leap forward in the pursuit of truly immersive virtual realities. As technology continues to evolve, these models will play an essential role in bridging the gap between the virtual and the real, providing users with a multisensory experience that closely mirrors real-world interactions. This advancement not only enriches the sensory landscape of VR but also opens new avenues for innovation across multiple disciplines.

**2. Related work.** The development of predictive odor models for virtual reality (VR) environments is a multidisciplinary endeavor that integrates insights from olfactory science, computational modeling, and immersive technology. This field has gained traction due to the increasing demand for more immersive VR experiences that engage multiple senses beyond sight and sound. Researchers have explored various approaches to simulate and predict olfactory experiences, with the aim of enhancing realism and user engagement.

Early work in olfactory science laid the groundwork by identifying the fundamental properties of odors and how they are perceived by humans. Studies such as those by the authors in [12] on olfactory receptors provided crucial insights into how humans detect and differentiate odors. This understanding is vital for creating models that can predict how different odorants will interact and be perceived in a virtual space.

In computational modeling, efforts have been made to simulate odor dispersion in virtual environments. These models often draw on fluid dynamics to predict how odor molecules move and spread. For example, the research by the authors in [13] applied computational fluid dynamics (CFD) to model odor dispersion in enclosed spaces, which can be adapted for VR scenarios. Such models help to create realistic odor propagation in virtual worlds, accounting for factors such as air flow and temperature.

Machine learning techniques have also been used to improve predictive accuracy. By training algorithms on large data sets of odorant molecules and their perceived smells, researchers aim to predict olfactory experiences more reliably. Approaches using neural networks, as discussed in [14, 15], show promise in predicting odor characteristics based on molecular structure, which is crucial for VR applications where real-time processing is needed.

In the realm of immersive technology, the integration of olfactory feedback into VR systems presents unique challenges. Devices such as scent diffusers and wearable olfactory interfaces have been developed to deliver controlled odor stimuli. The research by the authors in [16] demonstrated

how olfactory stimuli could be synchronized with visual and auditory cues to enhance the sense of presence in VR. This synchronization is key to creating a cohesive and believable virtual environment.

Furthermore, user experience studies are crucial to understanding how predictive odor models affect immersion and enjoyment in VR. Experiments often involve user testing to evaluate the effectiveness of olfactory integration. Findings from studies such as those by the authors in [17] suggest that olfactory signals can significantly enhance the perception of presence and emotional impact in virtual settings.

The challenge of standardization and calibration of olfactory devices remains a critical area of research. Differences in individual perception and the subjective nature of smell require models and devices to be highly adaptable. Collaborative efforts, such as those led by ISO working groups, aim to establish guidelines and standards for olfactory VR implementations, ensuring consistency and reliability across different systems.

Privacy and ethical considerations are emerging concerns as VR environments become more personalized. The collection and processing of olfactory data raises questions about user consent and data security. Researchers such as the authors in [18] emphasize the need for ethical frameworks to address these issues, ensuring that advances in olfactory VR respect user privacy.

Recent advances in sensor technology also play a pivotal role in the development of predictive odor models. The miniaturization and increased sensitivity of electronic noses enable more precise detection and analysis of odorants in real time. Studies by the authors in [19] highlight the potential of these sensors in VR applications, where they can provide feedback loops to dynamically adjust the virtual olfactory environment.

In summary, the development of predictive odor models for VR environments is a rapidly evolving field that bridges several scientific and technological domains. Continued research and collaboration across these areas will be essential to overcome current limitations and unlock the full potential of immersive olfactory experiences. As technology matures, it holds the promise of creating truly multisensory virtual worlds that can transform entertainment, education, and training.

In this study [20], machine learning-based classification models were developed to predict odor characteristics using the psychophysical data set created by the authors. This data set includes data on odorant properties for 480 structurally diverse compounds, each measured at two different concentrations (dilutions).



This study [21] uses data from five reservoirs in Kansas, USA, to create predictive models that relate dissolved geosmin concentrations to water quality factors. Individual reservoir-based models outperformed pooled data models in terms of performance. Events related to taste and smell occurred outside of the summer, and wintertime saw higher amounts of geosmin. The development of universal models was hampered by the strong dependence of geosmin concentrations on regional environmental conditions. Inorganic phosphorus limits have been found to play a major role in controlling the generation and release of geosmin into the water column.

This study [22] uses comprehensive data on a variety of biotic and abiotic characteristics of Taihu Lake to create predictive models for T&O (Taste and Odor) chemicals. The realistic dynamics of the T&O compounds were accurately recorded and a good match was achieved. They took into account two algal growth seasons (blooming and non-blooming) and two fractions of the T&O compounds (dissolved and particle bound), in contrast to previous odor models. The models proved to be useful for water resource managers in anticipating the possibility of T&O accidents and showed a high degree of accuracy in predicting T&O concentrations.

Due to the wider applicability and efficiency of model-based aroma design in SOR (Stimulus-Organism-Response) creation, this study [23] used it. In order to precisely characterize the olfactory characteristics and facilitate the creation of a more potent SOR, R-profile descriptors were utilized. Furthermore, the SOR was modeled using Machine Learning (ML) based on Artificial Neural Network (ANN), which demonstrates the accuracy of the model with an average  $R^2$  of 0.8807. Two case studies, aroma replacement and odor tuning, validated by tests and literature, supported the efficacy of the ML model and the computer-aided aroma design framework (CAAD) for aroma mixtures.

This study [24] presents a predictive model to predict chemical odor characteristics represented by binary values from mass spectra. The predictive model incorporates the language modeling approach Word2vec. The similarity between descriptors is minimized because, in the Sigma-Aldrich catalog data utilized in this work, descriptors representing the olfactory characteristics of molecules are used solely, even when additional descriptors express comparable odor characters.

Laboratory data were used in this study [25]. However, Internet of Things (IoT) sensor devices are used to gather information regarding odor compounds in real-world livestock scenarios. Due to the nature of data collection using sensors, missing data for a variety of causes is a frequent problem.

This study [26] uses dynamic olfactometry and analytical techniques to examine gas emissions from nine solid wastes and digestates during the active composting phase. The authors measured 22 important odorants and correlated them with odor concentration (OC) using an odor activity value (OAV) approach based on odor detection thresholds (ODT). To forecast OC trends, linear models employing OAVmax and OAVsum, as well as partial least squares (PLS) regressions, were investigated. Outperforming OAVmax and OAVsum, the PLS model explained 74-76% of the variance in OC. The important causes of odor pollution were found to be key odorants such as dimethyl sulfide, methanethiol, and hydrogen sulfide. In order to verify whether the PLS model can be applied to other processes, the validation set must be expanded.

### **3. Methodology**

**3.1. Introduction to Methodology.** Odor recognition in virtual reality (VR) environments is an interdisciplinary research field that combines sensory science and machine learning to create realistic olfactory experiences. Machine learning models such as Convolutional Neural Networks (CNN), Long Short-Term Memory (LSTM), Support Vector Machines (SVM), and Random Forests play an important role in processing and analyzing complex odor-related data. Each model has its own advantages, but not all are suitable for the current research objectives and conditions. After careful consideration, we chose Random Forest as the primary method to conduct this research and created a model named the Sensory Predictive Response Framework (SPRF).

CNN is a powerful model for processing spatial data, such as images or chemical structures. However, the main drawback of CNN is that it requires a large amount of data for training. This is a major challenge in the field of odor research, where data is often limited and difficult to collect. LSTM, with its ability to analyze temporal data and detect long-term dependence patterns, is also a potential option, but it requires high computational resources and considerable time for parameter optimization, which is not suitable for the scale of this study. Meanwhile, SVM has advantages in classifying data at high spatial scales, but when dealing with large or complex data sets, SVM is prone to becoming inefficient and resource-consuming.

Based on the current methodologies, a new model needed to be developed, and SPRF was chosen by us because it better fits the current research goals and context. The SPRF model is described in Section 4. It is a decision tree-based machine learning method distinguished by its ability to handle incomplete, noisy, and multidimensional data. It is particularly useful in early studies where olfactory data is often incomplete or non-normalized. A significant strength is its ability to perform well on small to medium-sized

datasets and provide interpretable results. This allows us to analyze and adjust the model more flexibly during the research process.

Additionally, SPRF offers a balance between accuracy and performance. Unlike deep models such as CNNs or LSTMs in [27], it does not require high hardware configurations or complex optimization techniques. With the ability to synthesize multiple decision trees, it not only reduces the risk of overfitting but also improves confidence in predictions. These characteristics make it an ideal method for performing initial analyses in the prediction of odor, laying the foundation for further research and practical applications in VR.

**3.2. Research Design.** In this section, we describe the type of research 3.2.1 and specifically outline our research design 3.2.2.

**3.2.1. Type of research.** Both qualitative and quantitative methodologies are used in this investigation. Specifically, in this study, qualitative factors are presented through the analysis of human sensory responses to odors in a virtual reality (VR) environment, in order to better understand the interaction between olfactory sensation and environmental factors. Quantitative factors, on the other hand, are applied in the analysis of data collected from sensors (e-nose), helping to determine the relationship between the chemical characteristics of odors and the recorded sensory characteristics.

### **3.2.2. Description of research design.**

– **Data Collection.** Data were collected from experiments with subjects (68 people) in a controlled environment. Important parameters such as phi (azimuth) and theta (tilt angle) were recorded from electronic sensors, and the data were stored as text files (txt) containing odor information. At the same time, participants' sensory responses to odors were recorded as color values in the CIELAB color space, allowing for the analysis of the relationship between odors and colors. In addition, using CIELAB allows for better representation of nonlinear data because gas sensors measure the concentration of compounds in the air, and the data obtained is often nonlinear. If using a color space such as RGB, the values may overlap making it difficult to distinguish between odors.

– **Experimental Design.** In a virtual reality environment, participants were exposed to different odors, such as black pepper, caramel, and cherry. The experiments were designed so that participants were not distracted and could concentrate on odor recognition under controlled environmental conditions. Environmental parameters such as light and temperature were also maintained and kept stable throughout the experiment.

– **Survey and data analysis.** After collecting sensory and chemical data, we use data analysis methods to model the relationship between odors and sensory features (color). Machine learning models are applied, including

regression, sliding windows, and data normalization, to analyze the collected data. Specifically, we use the sliding-window analysis method to generate consecutive data sequences, which enables the machine learning model to predict odors based on chemical and sensory features.

– **Model evaluation.** Machine learning models are evaluated using metrics such as the root mean square error (RMSE) to test the accuracy of predictions. The data is divided into training and testing sets to determine the generalizability of the model and test the reliability of the results.

– **Analysis of results.** Finally, the results obtained from the machine learning models will be compared with previous studies, and the validity of the data will be checked. The results from the experiments and models will provide insights into how odors are recognized in virtual reality environments, thereby helping to improve odor prediction models in future research.

**3.3. Data Collection.** In this section, we present the prediction of odor in virtual reality spaces. The environment plays a crucial role, as it encompasses the simulated settings where scents are integrated to enhance immersiveness. Participants are individuals who engage with these virtual environments, providing feedback on their sensory experiences to help refine odor prediction models. The data set comprises data collected from these interactions, including sensory responses and environmental variables, which are analyzed to improve the accuracy of odor predictions in digital simulations.

We use the dataset from [28], which is used to predict the color associated with odors using an electronic nose (e-nose). Perceptual data was collected from 68 participants who were asked to associate colors with ten different odors such as black pepper, caramel, and cherry. These participants selected colors in the CIELAB color space, which includes three channels:  $L^*$  for lightness,  $a^*$  for the red-green axis, and  $b^*$  for the yellow-blue axis. The odors were presented in a controlled environment to ensure consistent lighting conditions.

On the one hand, the chemical data was obtained using a custom-made e-nose equipped with various gas sensors, such as MP503, BME680, and MQ9. This setup allowed for the extraction of chemical features from the odors, which were then used to train the regression models. The e-nose recorded 100 samples, ten for each odor, with sensors that capture data on air quality, pollution level, and other environmental factors. The data were pre-processed by averaging sensor responses and smoothing signals before being used in the analysis.

On the other hand, the final dataset consisted of ten features per sample, excluding the time component used for packet reordering. Regression models

were trained to predict the  $L^*$ ,  $a^*$ , and  $b^*$  values for each odor based on the chemical characteristics captured by the e-nose.

The odor data in this study were collected in large datasets consisting of ten odors, each stored in ten data files. As shown in Figure 3, these files serve as examples for the remaining ten odors in the dataset and are depicted as two 3D graphs.

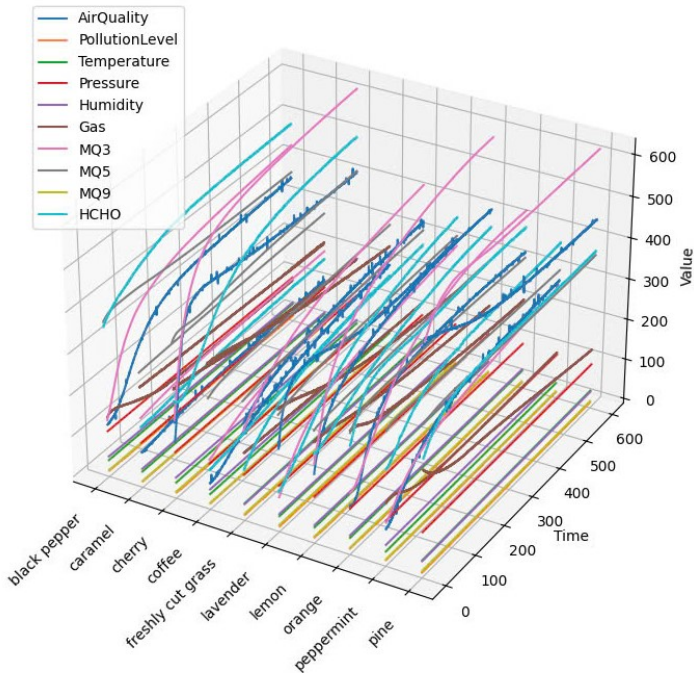


Fig. 3. Analyzed Scent Data from Sensors

As mentioned above, each collected odor dataset contains eleven columns of data describing that odor, as detailed below.

- **Time** is responsible for allocating time during the data collection process.
- **AirQuality** is responsible for measuring and evaluating air quality based on parameters such as fine dust concentration (PM2.5, PM10) and toxic gases.
- **PollutionLevel** measures the concentration of air pollution in the environment.

- **Temperature** measures the current ambient temperature, expressed in °C or °F units.
- **Pressure** measures atmospheric pressure, commonly used for weather and environmental conditions analysis, in hPa or mmHg.
- **Humidity** measures the relative humidity of the air, expressed as a percentage (%).
- **Gas** measures the concentration of flammable or toxic gases in the air, such as methane (CH<sub>4</sub>) or carbon monoxide (CO).
- **MQ3** sensor detects alcohol vapor, mainly for measuring ethanol concentration in the air.
- **MQ5** sensor is used to measure gases such as LPG, methane and butane.
- **MQ9** sensor detects carbon monoxide (CO) and other flammable gases in the air.
- **HCHO** measures the concentration of formaldehyde (HCHO), a pollutant commonly found in indoor or industrial environments.

Thus, a collected odor has ten collection files, and each file can have approximately 2125 rows of collected data with eleven columns. The recording process is done continuously, and the collected values are updated in real time. This ensures that even smallest change in the environment is accurately recorded by the "e-nose".

**3.4. Data Analysis.** In this study, data analysis was mainly performed using Python, a programming language with development tools such as PyCharm for coding and data processing. Popular libraries such as NumPy and Pandas were used for data manipulation and processing, while machine learning models were built and evaluated with the support of libraries such as Scikit-learn and Keras. Below are two sections related to data processing: the first is Data Analysis Methodology 3.4.1, which analyzes the data in this study, including various techniques for processing and understanding the data, and the next is Data Reliability and Validity Check 3.4.2. Through the above techniques, we can ensure that the collected data is accurate, reliable, and highly valid, thereby laying a solid foundation for the machine learning model in predicting odors in virtual reality environments.

#### **3.4.1. Data Analysis Methodology.**

– **Descriptive Statistics.** Before building the model, the data is explored through descriptive statistical techniques such as calculating the mean, standard deviation, and distribution of the values. This helps to identify key features of the data, such as *phi* and *theta* values, and detect outliers that may affect the model results.

– **Data transformation.** To ensure that the data falls within a suitable range for the machine learning model, normalization techniques are applied. The *transform* function is used to normalize the data, bringing the *phi* and *theta* values to a range of 0 to 1, which helps to increase the accuracy and performance of the model.

– **Regression.** To predict phi and theta values, regression models are used, in which linear regression techniques or more complex models such as Gaussian Process Regression (GPR) can be applied to analyze the relationship between chemical characteristics and sensory values.

– **Sliding window analysis.** An important technique used in this study is sliding window analysis, which aims to generate consecutive data sequences to serve as input to the machine learning model. Sliding window functions help to divide the data into consecutive sub-segments, providing training samples for the regression model. This technique helps the model to learn the temporal features in phi and theta data over time.

#### **3.4.2. Data Reliability and Validity Check.**

– **Reliability.** To check the reliability of the data, we use methods such as dividing the data into training and testing sets. This helps to evaluate the generalization ability of the model and confirm that the model does not overfit the training data. The results of the model are measured by metrics such as root mean square error (RMSE), which helps to assess the accuracy of the predictive model.

– **Validity.** The validity of the data is checked by comparing it with experimental results or previous studies in the field. The data is collected from reliable sources and under strictly controlled conditions, ensuring that outliers do not affect the results. Furthermore, the data is thoroughly processed to remove missing or unusual values that may cause bias in the analysis.

### **4. Experiments**

**4.1. Experimental Settings.** In this part, we will install and test the Python programming language on a 64-bit Windows 11 Pro computer. The system specifications include a Core i5-6300U processor, 16GB of RAM, a 512GB SSD, and a 12.5-inch HD display with a resolution of 1366x768. Additionally, we conducted data analysis and created graphs using the Python programming language, focusing on a dataset comprising 11 features collected from odor samples [28]. These samples were instrumental in developing odor prediction models, which paved the way for future research involving a broader range of odors. This expansion is crucial for improving the accuracy and applicability of our models.

To support these advancements, we plan to incorporate assessments in complex 3D spaces using virtual reality. These environments provide objective

evaluations for participants, enhancing their experience and ensuring more rigorous testing conditions. This approach will facilitate the development of more robust and versatile odor prediction models.

An overview of our research model is presented in Figure 4. This model is designed to predict scent in virtual reality (VR) environments using the Random Forest algorithm. The model-building process includes many important steps, from input data collection, processing and feature extraction, to training and evaluating prediction performance.

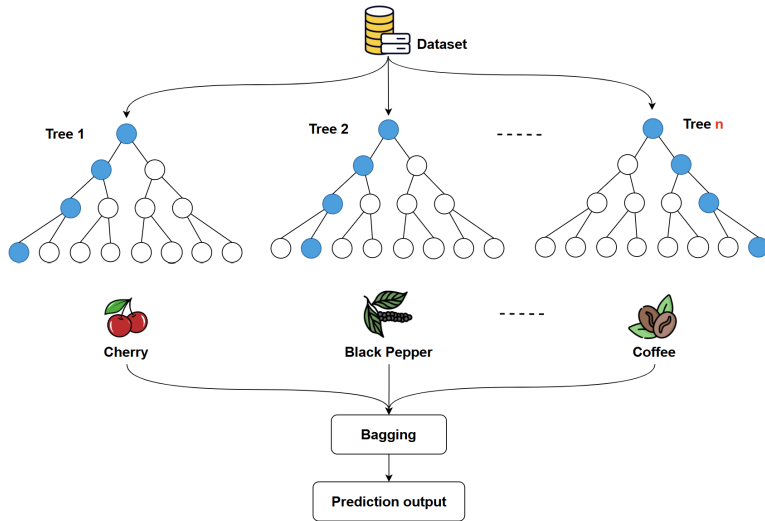


Fig. 4. Model Overview

**Bootstrap Sampling.** For each tree  $T_i$ , a subset  $\mathcal{D}_i$  is randomly sampled from the training dataset  $\mathcal{D}$ , with the size equal to that of  $\mathcal{D}$  but allowing repetition. The mathematical formulation of this sampling process is given by Equation (1). In this, the index  $j_k$  is drawn randomly from the set of available data points, as defined in Equation (2) below:

$$\mathcal{D}_i = \{x_{j_1}, x_{j_2}, \dots, x_{j_m}\}, \quad \text{where } x_{j_k} \in \mathcal{D}, \quad (1)$$

where:

- $\mathcal{D}_i$  – bootstrap sample for the  $i$ -th tree.
- $x_{j_k}$  – the  $k$ -th data point sampled from the dataset  $\mathcal{D}$ .
- $\mathcal{D}$  – the training dataset.



$$m = |\mathcal{D}| \quad \text{and} \quad j_k \sim \text{Uniform}(1, m), \quad (2)$$

**where:**

- $m$  – the total number of data points in the training set  $\mathcal{D}$ .
- $j_k$  – randomly drawn index for the  $k$ -th data point.
- $\text{Uniform}(1, m)$  – random sampling from the set  $\{1, 2, \dots, m\}$ .

**Building Each Decision Tree.** Each tree  $T_i$  is constructed based on  $\mathcal{D}_i$ , with several random features. At each split node, a random subset  $\mathcal{F}$  of features  $n$  is considered to select the best-split point. The size of  $\mathcal{F}$  is typically  $\sqrt{n}$  or  $\log_2(n)$ .

The Gini index, given in Equation (3), is used to evaluate the quality of each potential split. This index helps determine how well the split divides the data into distinct classes. The lower the Gini index, the purer the split.

$$G = 1 - \sum_{k=1}^K p_k^2, \quad (3)$$

**where:**

- $G$  – the Gini index for a given split.
- $K$  – the number of classes (labels).
- $p_k$  – the probability of a sample belonging to class  $k$  at that node.

**Aggregating the Results.** After building  $N$  decision trees, the algorithm aggregates the results from all the trees to make a prediction. For classification problems, the final prediction  $\hat{y}$  is obtained by taking the class that appears most frequently across all trees, as shown in Equation (4).

$$\hat{y} = \arg \max_k \left( \sum_{i=1}^N \mathbb{I}[T_i(x) = k] \right), \quad (4)$$

**where:**

- $\hat{y}$  – the predicted class label.
- $k$  – a class label.
- $T_i(x)$  – the prediction made by the  $i$ -th tree for input sample  $x$ .
- $\mathbb{I}[\cdot]$  – indicator function (1 if the condition is true, 0 if false).
- $N$  – the total number of trees in the forest.

For regression problems, the final prediction is the average of all the individual tree predictions, as seen in Equation (5).

$$\hat{y} = \frac{1}{N} \sum_{i=1}^N T_i(x), \quad (5)$$

**where:**

- $\hat{y}$  – the predicted value for regression.
- $T_i(x)$  – the output of the  $i$ -th tree for input sample  $x$ .
- $N$  – the total number of trees in the forest.

**Specific Parameters from Code.**

– Number of trees ( $n\_estimators$ ):  $N = 100$ . Increasing the number of trees  $N$  helps reduce the model's variance but increases training and prediction time.

– Random feature subset size at each node ( $\mathcal{F}$ ). By default in sklearn,  $\mathcal{F} = \sqrt{n}$ , where  $n = 11$ , so  $\mathcal{F} \approx 3$ .

– Maximum depth ( $max\_depth$ ):  $\infty$  (default). Allows the tree to grow until leaf nodes achieve purity ( $Gini = 0$ ).

– Minimum samples to split a node ( $min\_samples\_split$ ): 2.

– Minimum samples at each leaf node ( $min\_samples\_leaf$ ): 1.

– Bootstrap sampling. Used to ensure that each tree is built from a different dataset.

**4.2. Performance Analysis.** In the realm of machine learning, performance analysis is essential to evaluate how well different models meet the objectives of a given task. Therefore, by understanding the strengths and limitations of various algorithms, such as CNN [29], LSTM [30], and SVM [31] and the proposed (SPRF), we can make informed decisions on model selection and optimization. In this part, we delve into the comparative performance of these models, providing insights into their applicability to diverse datasets and problem domains.

In this research area, we evaluate the performance of a prediction model using accuracy as a key metric. Accuracy is defined as the ratio of correct predictions to the total number of predictions. It is calculated using the following formula:

$$Accuracy = \frac{NumberofCorrectPredictions}{TotalNumberofPredictions}. \quad (6)$$

Next, F1-Score is used as an important evaluation metric in various types of tasks to evaluate the performance of a model because it combines precision and recall scores:

$$recall = \frac{TP}{TP + FN}, \quad (7)$$

$$precision = \frac{TP}{TP + FP}, \quad (8)$$

$$F1 - Score = \frac{2 \times precision \times recall}{precision + recall}, \quad (9)$$

**where:**

- TP. The number of times the model correctly predicted a scent that actually existed.
- TN. The number of times the model correctly predicted that a scent did not exist and that scent did not exist.
- FP. The number of times the model incorrectly predicted that a scent existed but did not exist.

This metric provides a straightforward assessment of the model’s effectiveness in making accurate predictions.

Table 1 presents a comparative analysis of model accuracy metrics for four different machine learning models: CNN [29], LSTM [30], SVM [31], and SPRF. The accuracy percentages indicate the performance of each model in terms of its ability to correctly predict outcomes. The CNN model, with an accuracy of 79.46%, was the least accurate among the models presented, given that they used CNN for odor pleasantness prediction. While CNNs are typically strong in handling image data due to their convolutional layers, their lower performance here might suggest that the dataset used is not well-suited for a CNN’s architecture, or that the model was not optimally tuned. This highlights the importance of model selection and hyperparameter tuning in achieving high accuracy.

Table 1. Comparative Analysis of Model Accuracy Metrics

Models	Accuracy (%)	F1-Score (%)
CNN [29]	79.46	76.70
LSTM [30]	80.37	77.70
SVM [31]	85.24	85.03
<b>The proposed (SPRF)</b>	<b>98.13</b>	<b>98.08</b>

The LSTM model showed a slightly better accuracy of 80.37%, which represents a modest improvement over CNN, but their LSTM was built with

automatic gas source localization in an outdoor environment in mind. They set the feature transformation function in the convolutional layer to (3,3,5) with no padding. LSTMs are particularly effective in handling sequential data, such as time series or natural language processing tasks. The marginal increase in accuracy suggests that the data set may have some sequential aspect, but the improvement is not substantial. This could mean that the LSTM architecture captures some temporal dependencies better than CNN but still struggles to fully understand the underlying patterns, possibly due to insufficient data preprocessing or feature engineering.

On the one hand, the SPRF model that we developed – previously no Random Forests model has been applied to predict smells in virtual reality environment – achieved a significantly higher accuracy of 98.13%, making it the most accurate model in this comparison. Random Forests are ensemble learning methods that are robust to overfitting and can handle a wide variety of data types, which might explain their superior performance. This suggests that the dataset features are well-suited for decision tree-based models, where feature importance and interactions play a crucial role. The high accuracy of the SPRF model implies that it effectively captures the complex patterns within the data, making it a reliable choice for similar datasets.

On the other hand, the Support Vector Machine (SVM) model also performed well, with an accuracy of 85.24%. SVM has not been applied to odor recognition in virtual environments; it has been used for emotion recognition in VR scenes [31] and some other VR-related categories such as odor source localization [32] so we further developed the SVM model to apply to odor prediction to provide more model choices for odor prediction in this study and demonstrated that SPRF performs well, SVMs are known for their effectiveness in high-dimensional spaces and are especially useful in classification tasks with clear boundary distances. The performance of the SVM model shows that it can handle the data efficiently, although not as well as the SPRF model. This shows that although SVM can delineate classes to a good extent, it may not capture all complex patterns as effectively as the ensemble approach of the SPRF model.

In conclusion, the data indicate that the model selection should be tailored to the characteristics of the data set. The Random Forest model outperforms the others, suggesting its suitability for these particular data. However, the choice between models should also consider other factors such as computational efficiency, interpretability, and specific application requirements. The results emphasize the importance of understanding the strengths and limitations of each model type, as well as the necessity for a thorough data analysis to guide model selection and optimization.

**4.3. Discussing Smell in VR Environment.** To integrate odor sensor models into the VR system, first, it is necessary to collect data from odor and gas sensors, then use data analysis algorithms to convert the measured values into signals that can be interacted with in the VR environment. Sensors such as MQ3, MQ5, and HCHO [33] will provide data on specific gases, while the data processing system will analyze and reproduce odors in the virtual space.

System installation requires connecting VR software to sensors via data transmission protocols, such as MQTT or WebSocket [34, 35], to ensure that sensor data are continuously updated and responded to promptly. VR software will need to integrate biological models and odor analysis to create realistic odor responses in the virtual environment. The odor generation system will be installed to reproduce odors based on sensor data while ensuring stability and high performance when operating in the VR environment.

This process requires both software and hardware to work in sync, with the VR software controlling and simulating the environment, and the hardware performing the odor reproduction.

Potential issues that may arise during this process:

– *Accuracy of sensor data.* Odor and gas sensors need to be highly accurate to ensure that the data collected is accurate and reliable in simulations. If the collected data is skewed or inaccurate, this can lead to incorrect odor reproduction or even negative effects in the simulation.

– *Interaction between odors and the VR environment.* One of the major challenges in integrating odor sensors into VR is how to make the odors perceived realistically and in line with other elements in the virtual environment. This requires a sophisticated system that can combine odors with other elements such as images, sounds, and sensations in the VR space.

– *Odor technology in virtual environments.* To create a realistic experience, the odor system must be able to reproduce odors with high accuracy and be adjustable. However, this technology still faces some limitations in the ability to reproduce various odors, the consistency of the odors over time, and the precision of the odor distribution in virtual space.

**5. Conclusions.** In conclusion, based on empirical research, the SPRF smell prediction outperforms other models, achieving 98.13% accuracy. In contrast, the CNN model has the lowest accuracy at 79.46%, followed by LSTM at 80.37% and SVM at 85.24%. This shows that SPRF is the most effective model for developing predictive odor models in VR environments. It improves the F1 score by 13.05% to 21.38% and accuracy by 12.89% to 18.67%. Based on these results, future research should prioritize the refinement of the SPRF model to further enhance its predictive capabilities. Furthermore, exploring hybrid models that combine the strengths of multiple approaches could lead

to even more robust solutions. Implementing these advanced models will be crucial in integrating scents into VR, thus enriching digital experiences and making them more immersive and engaging.

## References

1. Nikolic D., Whyte J. Visualizing a new sustainable world: Toward the next generation of virtual reality in the built environment. *Buildings*. 2021. vol. 11. no. 11(11). DOI: 10.3390/buildings11110546.
2. Felten P., Lambert L.M. Relationship-rich education: How human connections drive success in college. Johns Hopkins University Press, 2020. 208 p.
3. Flavian C., Ibanez-Sanchez S., Orus C. The influence of scent on virtual reality experiences: The role of aroma-content congruence. *Journal of Business Research*. 2021. vol. 123. pp. 289–301. DOI: 10.1016/j.jbusres.2020.09.036.
4. Petit O., Velasco C., Spence C. Digital sensory marketing: Integrating new technologies into multisensory online experience. *Journal of Interactive Marketing*. 2019. vol. 45. pp. 42–61. DOI: 10.1016/j.intmar.2018.07.004.
5. Hung N.V., Thinh P.H., Thanh N.H., Lam T.T., Hien T.T., Ninh V.T., Huong T.T. Lvsum-optimized live 360 degree video streaming in unicast and multicast over mobile networks. *IEEE 15th International Conference on Computational Intelligence and Communication Networks (CICN)*. IEEE, 2023. pp. 29–34. DOI: 10.1109/CICN59264.2023.10402136.
6. Hung N.V., Tien B.D., Anh T.T.T., Nam P.N., Huong T.T. An efficient approach to terminate 360-video stream on http/3. *American Institute of Physics Conference Proceedings*. 2023. vol. 2909. no. 1. DOI: 10.1063/5.0182042.
7. Zeynali A., Hajiesmaili M.H., Sitaraman R.K. Bola360: Near-optimal view and bitrate adaptation for 360-degree video streaming. *Proceedings of the 15th ACM Multimedia Systems Conference (MMSys)*. 2024. pp. 12–22. DOI: 10.1145/3625468.3647607.
8. Chiariotti F. A survey on 360-degree video: Coding, quality of experience and streaming. *Computer Communications*. 2021. vol. 177. pp. 133–155. DOI: 10.1016/j.comcom.2021.06.029.
9. Nguyen H., Dao T.N., Pham N.S., Dang T.L., Nguyen T.D., Truong T.H. An accurate viewport estimation method for 360 video streaming using deep learning. *EAI Endorsed Transactions on Industrial Networks and Intelligent Systems*. 2022. vol. 9. no. 4. DOI: 10.4108/eetinis.v9i4.2218.
10. Hung N.V., Dat P., Tan N., Quan N., Trang L.T.H., Nam L. Heverl – viewport estimation using reinforcement learning for 360-degree video streaming. *Informatics and Automation*. 2025. vol. 24(1). pp. 302–328. DOI: 10.15622/ia.24.1.11.
11. Nguyen V.H., Bui D.T., Tran T.L., Truong C.T., Truong T.H. Scalable and resilient 360-degree-video adaptive streaming over HTTP/2 against sudden network drops. *Computer Communications*. 2024. vol. 216. pp. 1–15. DOI: 10.1016/j.comcom.2024.01.001.
12. Buck L., Axel R. A novel multigene family may encode odorant receptors: a molecular basis for odor recognition. *Cell*. 1991. vol. 65(1). pp. 175–187. DOI: 10.1016/0092-8674(91)90418-X.
13. Yang Y., Jing D., Zhao L. Computational fluid dynamics modeling of reactive multiphase flow for suspended photocatalytic water splitting of hydrogen production system. *Applied Thermal Engineering*. 2020. vol. 173. DOI: 10.1016/j.applthermaleng.2020.115220.
14. Song R., Keller A.A., Suh S. Rapid life-cycle impact screening using artificial neural networks. *Environmental Science & Technology*. 2017. vol. 51. no. 18. pp. 10777–10785. DOI: 10.1021/acs.est.7b02862.

15. Lotsch J., Kringel D., Hummel T. Machine learning in human olfactory research. *Chemical senses*. 2019. vol. 44. no. 1. pp. 11–22. DOI: 10.1093/chemse/bjy067.
16. Narumi T., Kajinami T., Nishizaka S., Tanikawa T., Hirose M. Pseudo-gustatory display system based on cross-modal integration of vision, olfaction and gustation. *IEEE Virtual Reality Conference*. IEEE, 2011. pp. 127–130. DOI: 10.1109/VR.2011.5759450.
17. Niedenthal S., Fredborg W., Lunden P., Ehrndal M., Olofsson J.K. A graspable olfactory display for virtual reality. *International Journal of Human-Computer Studies*. 2023. vol. 169. DOI: 10.1016/j.ijhcs.2022.102928.
18. Velasco C., Vargas J., Petit O. Multisensory experiences and technology in the context of wine experiences. *Journal of Wine Research*. 2024. vol. 35. no. 2. pp. 85–100. DOI: 10.1080/09571264.2024.2310304.
19. Wilson A. Electronic-nose applications in forensic science and for analysis of volatile biomarkers in the human breath. *Journal of Forensic Science & Criminology*. 2014. vol. 1. no. 1. pp. 1–21.
20. Chacko R., Jain D., Patwardhan M., Puri A., Karande S., Rai B. Data based predictive models for odor perception. *Scientific reports*. 2020. vol. 10. DOI: 10.1038/s41598-020-73978-1.
21. Dzialowski A.R., Smith V.H., Huggins D.G., Denoyelles F., Lim N.-C., Baker D.S., Beury J.H. Development of predictive models for geosmin-related taste and odor in Kansas, USA, drinking water reservoirs. *Water Research*. 2009. vol. 43. no. 11. pp. 2829–2840. DOI: 10.1016/j.watres.2009.04.001.
22. Qi M., Chen J., Sun X., Deng X., Niu Y., Xie P. Development of models for predicting the predominant taste and odor compounds in Taihu Lake, China. *PloS one*. 2012. vol. 7(12). DOI: 10.1371/annotation/c86cbb15-dbb2-4fe8-b2fa-35604b6c4887.
23. Zhang L., Mao H., Zhuang Y., Wang L., Liu L., Dong Y., Du J., Xie W., Yuan Z. Odor prediction and aroma mixture design using machine learning model and molecular surface charge density profiles. *Chemical Engineering Science*. 2021. vol. ,245. DOI: 10.1016/j.ces.2021.116947.
24. Nozaki Y., Nakamoto T. Predictive modeling for odor character of a chemical using machine learning combined with natural language processing. *PloS one*. 2018. vol. 13(12). DOI: 10.1371/journal.pone.0208962.
25. Lee D.-H., Woo S.-E., Jung M.-W., Heo T.-Y. Evaluation of odor prediction model performance and variable importance according to various missing imputation methods. *Applied Sciences*. 2022. vol. 12. no. 6. DOI: 10.3390/app12062826.
26. Rincon C.A., De Guardia A., Couvert A., Wolbert D., Le Roux S., Soutrel I., Nunes G. Odor concentration (OC) prediction based on odor activity values (OAVs) during composting of solid wastes and digestates. *Atmospheric Environment*. 2019. vol. 201. pp. 1–12. DOI: 10.1016/j.atmosenv.2018.12.030.
27. Hung N.V., Loi T.Q., Binh N.H., Nga N.T.T., Huong T.T., Luu D.L. Building an online learning model through a dance recognition video based on deep learning. *Informatics and Automation*. 2024. vol. 23(1). pp. 101–128. DOI: 10.15622/ia.23.1.4.
28. Ward R., Rahman S., Wuerger S., Marshall A. Predicting the colour associated with odours using an electronic nose. in *Proceedings of the 1st Workshop on Multisensory Experiences-SensoryX'21*. SBC, 2021. DOI: 10.5753/sensoryx.2021.15683.
29. Wu D., Luo D., Wong K.-Y., Hung K. POP-CNN: Predicting odor leasantness with convolutional neural network. *IEEE Sensors Journal*. 2019. vol. 19. no. 23. pp. 11337–11345. DOI: 10.1109/JSEN.2019.2933692.
30. Bilgera C., Yamamoto A., Sawano M., Matsukura H., Ishida H. Application of convolutional long short-term memory neural networks to signals collected from a sensor

- network for autonomous gas source localization in outdoor environments. *Sensors*. 2018. vol. 18. no. 12. DOI: 10.3390/s18124484.
31. Liu W., Zhong X., Liu G., Xie J. Emotion Recognition Based ECG in Olfactory-enhanced VR Scenes. 9th International Symposium on Computer and Information Processing Technology (ISCIPT). 2024. pp. 445–448. DOI: 10.1109/ISCIPT61983.2024.10672829.
  32. Jing T., Meng Q.-H., Ishida H. Recent progress and trend of robot odor source localization. *IEEJ Transactions on Electrical and Electronic Engineering*. 2021. vol. 16. no. 7. pp. 938–953. DOI: 10.1002/tee.23364.
  33. Grodnyomchai B., Satcharoen K., Tangtisanon P. Enhancing odor classification of essential oils with electronic nose data. 9th International Conference on Computer and Communication Systems (ICCCS). 2024. pp. 531–536. DOI: 10.1109/ICCCS61882.2024.10603362.
  34. Saleme E.B., Santos C.A., Ghinea G. A mulsemmedia framework for delivering sensory effects to heterogeneous systems. *Multimedia Systems*. 2019. vol. 25. pp. 421–447. DOI: 10.1007/s00530-019-00618-8.
  35. Bhatia M., Kaur S., Sood S.K. Iot-inspired smart toilet system for home-based urine infection prediction. *ACM Transactions on Computing for Healthcare*. 2020. vol. 1. no. 3. pp. 1–25. DOI: 10.1145/3379506.

**Nguyen Viet Hung** — Lecturer, International training and cooperation institute, East Asia University of Technology. Research interests: multimedia communications, network security, artificial intelligence, traffic engineering in next-generation networks, QoE/QoS guarantee for network services, green networking, applications. The number of publications — 23. hungnv@eaut.edu.vn; Ky Phu – Ky Anh, Ha Tinh, Viet Nam; office phone: +84(098)911-2079.

**Nguyen Anh Quan** — Research assistant, East Asia University of Technology. Research interests: natural language processing, computer vision. The number of publications — 3. anhq46724@gmail.com; Duong Ha – Gia Lam, Hanoi, Viet Nam; office phone: +84(096)278-4293.

**Nguyen Tan** — Research assistant, East Asia University of Technology. Research interests: applications, data analysis. The number of publications — 4. tan25102000@gmail.com; Trung Dung – Tien Lu, Hung Yen, Viet Nam; office phone: +84(035)919-0216.

**Tran Trieu Hai** — Research assistant, Faculty of general education, Hanoi Open University. Research interests: network security, AI, natural language processing, applications. The number of publications — 2. haitt@hou.edu.vn; Nguyen Hien Street, Hai Ba Trung District, Hanoi, Viet Nam; office phone: +84(098)596-6866.

**Dang Khanh Trung** — Lecturer, Faculty of information technology, East Asia University of Technology. Research interests: multimedia communications, database management systems, artificial intelligence, applications. The number of publications — 2. trungdk@eaut.edu.vn; 14, Bo De – Long Bien, Hanoi, Viet Nam; office phone: +84(032)889-9334.

**Le Mai Nam** — Lecturer, Faculty of information technology, East Asia University of Technology. Research interests: software engineering, optimization mathematics, applications. The number of publications — 2. namlm@eaut.edu.vn; Phuong Trung – Thanh Oai, Hanoi, Viet Nam; office phone: +84(098)208-2117.

**Bui Thanh Loan** — Lecturer, Faculty of information technology, East Asia University of Technology. Research interests: multimedia communications, database management systems, artificial intelligence, applications. The number of publications — 1. loanbt@eaut.edu.vn; Long Bien, Hanoi, Viet Nam; office phone: +84(090)218-1085.



**Nguyen Thi Thuy Nga** — Lecturer, Faculty of information technology, East Asia University of Technology. Research interests: multimedia communications, database management systems, artificial intelligence, applications. The number of publications — 4. [ngantt@eaut.edu.vn](mailto:ngantt@eaut.edu.vn); Ta Thanh Oai – Thanh Tri, Hanoi, Viet Nam; office phone: +84(097)529-4770.

Н.В. ХУНГ, Н.А. КУАН, Н. ТАН, Т.Т. ХАЙ, Д.Т. ЧУНГ, Л.М. НАМ,  
Б.Т. ЛОАН, Н.Т. НГА

## СОЗДАНИЕ ПРОГНОЗИРУЮЩИХ МОДЕЛЕЙ ЗАПАХОВ ДЛЯ СРЕД ВИРТУАЛЬНОЙ РЕАЛЬНОСТИ

*Хунг Н.В., Куан Н.А., Тан Н., Хай Т.Т., Чунг Д.Т., Нам Л.М., Лоан Б.Т., Нга Н.Т. Создание прогнозирующих моделей запахов для сред виртуальной реальности.*

**Аннотация.** В среде, насыщенной сенсорными стимулами, человеческий опыт формируется за счет сложного взаимодействия множества чувств. Однако при цифровом взаимодействии задействуются преимущественно зрительные и слуховые модальности, в то время как другие сенсорные каналы, такие как обоняние, остаются практически неиспользованными. Технология виртуальной реальности обладает значительным потенциалом для преодоления этого ограничения за счет включения более широкого спектра сенсорных стимулов, что позволяет создавать более погружающий опыт. В данном исследовании представлен новый подход к интеграции обонятельных стимулов в виртуальную среду посредством разработки прогностической модели запахов, названной Сенсорно-Прогностическая Реакционная Структура (SPRF). Цель исследования заключается в улучшении сенсорного измерения виртуальной реальности путем адаптации обонятельных стимулов к конкретному контенту и контексту. Это достигается за счет сбора информации о местоположении источников запахов и их идентификации по характерным признакам, что позволяет воспроизводить их в пространстве виртуальной среды, тем самым повышая вовлеченность и уровень погружения пользователя. Кроме того, в исследовании изучается влияние различных факторов, связанных с запахами, на восприятие и поведение пользователя в виртуальной реальности, с целью разработки прогностических моделей, оптимизированных для интеграции обонятельных стимулов. Эмпирические оценки показывают, что модель SPRF демонстрирует производительность с точностью 98,13%, значительно превосходя обычные модели, такие как сверточные нейронные сети (CNN, 79,46%), сети с долгой краткосрочной памятью (LSTM, 80,37%) и метод опорных векторов (SVM, 85,24%). Кроме того, SPRF обеспечивает заметные улучшения в показателях F1 (на 13,05%-21,38%) и точности (на 12,89%-18,67%) по сравнению с этими альтернативными моделями. Эти результаты подчеркивают эффективность SPRF в развитии интеграции обонятельных стимулов в виртуальной реальности, предлагая ценные идеи для проектирования мультисенсорных цифровых сред.

**Ключевые слова:** виртуальная реальность, запах, выбор модели, пользовательский опыт, воображение, прогнозирование запахов.

### Литература

1. Nikolic D., Whyte J. Visualizing a new sustainable world: Toward the next generation of virtual reality in the built environment. *Buildings*. 2021. vol. 11. no. 11(11). DOI: 10.3390/buildings11110546.
2. Felten P., Lambert L.M. Relationship-rich education: How human connections drive success in college. Johns Hopkins University Press, 2020. 208 p.
3. Flavian C., Ibanez-Sanchez S., Orus C. The influence of scent on virtual reality experiences: The role of aroma-content congruence. *Journal of Business Research*. 2021. vol. 123. pp. 289–301. DOI: 10.1016/j.jbusres.2020.09.036.

4. Petit O., Velasco C., Spence C. Digital sensory marketing: Integrating new technologies into multisensory online experience. *Journal of Interactive Marketing*. 2019. vol. 45. pp. 42–61. DOI: 10.1016/j.intmar.2018.07.004.
5. Hung N.V., Thinh P.H., Thanh N.H., Lam T.T., Hien T.T., Ninh V.T., Huong T.T. Lvsom-optimized live 360 degree video streaming in unicast and multicast over mobile networks. *IEEE 15th International Conference on Computational Intelligence and Communication Networks (CICN)*. IEEE, 2023. pp. 29–34. DOI: 10.1109/CICN59264.2023.10402136.
6. Hung N.V., Tien B.D., Anh T.T.T., Nam P.N., Huong T.T. An efficient approach to terminate 360-video stream on http/3. *American Institute of Physics Conference Proceedings*. 2023. vol. 2909. no. 1. DOI: 10.1063/5.0182042.
7. Zeynali A., Hajiesmaili M.H., Sitaraman R.K. Bola360: Near-optimal view and bitrate adaptation for 360-degree video streaming. *Proceedings of the 15th ACM Multimedia Systems Conference (MMSys)*. 2024. pp. 12–22. DOI: 10.1145/3625468.3647607.
8. Chiariotti F. A survey on 360-degree video: Coding, quality of experience and streaming. *Computer Communications*. 2021. vol. 177. pp. 133–155. DOI: 10.1016/j.comcom.2021.06.029.
9. Nguyen H., Dao T.N., Pham N.S., Dang T.L., Nguyen T.D., Truong T.H. An accurate viewport estimation method for 360 video streaming using deep learning. *EAI Endorsed Transactions on Industrial Networks and Intelligent Systems*. 2022. vol. 9. no. 4. DOI: 10.4108/eetinis.v9i4.2218.
10. Hung N.V., Dat P., Tan N., Quan N., Trang L.T.H., Nam L. Heverl – viewport estimation using reinforcement learning for 360-degree video streaming. *Informatics and Automation*. 2025. vol. 24(1). pp. 302–328. DOI: 10.15622/ia.24.1.11.
11. Nguyen V.H., Bui D.T., Tran T.L., Truong C.T., Truong T.H. Scalable and resilient 360-degree-video adaptive streaming over HTTP/2 against sudden network drops. *Computer Communications*. 2024. vol. 216. pp. 1–15. DOI: 10.1016/j.comcom.2024.01.001.
12. Buck L., Axel R. A novel multigene family may encode odorant receptors: a molecular basis for odor recognition. *Cell*. 1991. vol. 65(1). pp. 175–187. DOI: 10.1016/0092-8674(91)90418-X.
13. Yang Y., Jing D., Zhao L. Computational fluid dynamics modeling of reactive multiphase flow for suspended photocatalytic water splitting of hydrogen production system. *Applied Thermal Engineering*. 2020. vol. 173. DOI: 10.1016/j.applthermaleng.2020.115220.
14. Song R., Keller A.A., Suh S. Rapid life-cycle impact screening using artificial neural networks. *Environmental Science & Technology*. 2017. vol. 51. no. 18. pp. 10777–10785. DOI: 10.1021/acs.est.7b02862.
15. Lotsch J., Kringsel D., Hummel T. Machine learning in human olfactory research. *Chemical senses*. 2019. vol. 44. no. 1. pp. 11–22. DOI: 10.1093/chemse/bjy067.
16. Narumi T., Kajinami T., Nishizaka S., Tanikawa T., Hirose M. Pseudo-gustatory display system based on cross-modal integration of vision, olfaction and gustation. *IEEE Virtual Reality Conference*. IEEE, 2011. pp. 127–130. DOI: 10.1109/VR.2011.5759450.
17. Niedenthal S., Fredborg W., Lunden P., Ehrndal M., Olofsson J.K. A graspable olfactory display for virtual reality. *International Journal of Human-Computer Studies*. 2023. vol. 169. DOI: 10.1016/j.ijhcs.2022.102928.
18. Velasco C., Vargas J., Petit O. Multisensory experiences and technology in the context of wine experiences. *Journal of Wine Research*. 2024. vol. 35. no. 2. pp. 85–100. DOI: 10.1080/09571264.2024.2310304.
19. Wilson A. Electronic-nose applications in forensic science and for analysis of volatile biomarkers in the human breath. *Journal of Forensic Science & Criminology*. 2014. vol. 1. no. 1. pp. 1–21.

20. Chacko R., Jain D., Patwardhan M., Puri A., Karande S., Rai B. Data based predictive models for odor perception. *Scientific reports*. 2020. vol. 10. DOI: 10.1038/s41598-020-73978-1.
21. Dzialowski A.R., Smith V.H., Huggins D.G., Denoyelles F., Lim N.-C., Baker D.S., Beury J.H. Development of predictive models for geosmin-related taste and odor in Kansas, USA, drinking water reservoirs. *Water Research*. 2009. vol. 43. no. 11. pp. 2829–2840. DOI: 10.1016/j.watres.2009.04.001.
22. Qi M., Chen J., Sun X., Deng X., Niu Y., Xie P. Development of models for predicting the predominant taste and odor compounds in Taihu Lake, China. *PLoS one*. 2012. vol. 7(12). DOI: 10.1371/annotation/c86cbb15-dbb2-4fe8-b2fa-35604b6c4887.
23. Zhang L., Mao H., Zhuang Y., Wang L., Liu L., Dong Y., Du J., Xie W., Yuan Z. Odor prediction and aroma mixture design using machine learning model and molecular surface charge density profiles. *Chemical Engineering Science*. 2021. vol. 245. DOI: 10.1016/j.ces.2021.116947.
24. Nozaki Y., Nakamoto T. Predictive modeling for odor character of a chemical using machine learning combined with natural language processing. *PLoS one*. 2018. vol. 13(12). DOI: 10.1371/journal.pone.0208962.
25. Lee D.-H., Woo S.-E., Jung M.-W., Heo T.-Y. Evaluation of odor prediction model performance and variable importance according to various missing imputation methods. *Applied Sciences*. 2022. vol. 12. no. 6. DOI: 10.3390/app12062826.
26. Rincon C.A., De Guardia A., Couvert A., Wolbert D., Le Roux S., Soutrel I., Nunes G. Odor concentration (OC) prediction based on odor activity values (OAVs) during composting of solid wastes and digestates. *Atmospheric Environment*. 2019. vol. 201. pp. 1–12. DOI: 10.1016/j.atmosenv.2018.12.030.
27. Hung N.V., Loi T.Q., Binh N.H., Nga N.T.T., Huong T.T., Luu D.L. Building an online learning model through a dance recognition video based on deep learning. *Informatics and Automation*. 2024. vol. 23(1). pp. 101–128. DOI: 10.15622/ia.23.1.4.
28. Ward R., Rahman S., Wuergler S., Marshall A. Predicting the colour associated with odours using an electronic nose. in *Proceedings of the 1st Workshop on Multisensory Experiences-SensoryX'21*. SBC, 2021. DOI: 10.5753/sensoryx.2021.15683.
29. Wu D., Luo D., Wong K.-Y., Hung K. POP-CNN: Predicting odor leasatness with convolutional neural network. *IEEE Sensors Journal*. 2019. vol. 19. no. 23. pp. 11337–11345. DOI: 10.1109/JSEN.2019.2933692.
30. Bilgera C., Yamamoto A., Sawano M., Matsukura H., Ishida H. Application of convolutional long short-term memory neural networks to signals collected from a sensor network for autonomous gas source localization in outdoor environments. *Sensors*. 2018. vol. 18. no. 12. DOI: 10.3390/s18124484.
31. Liu W., Zhong X., Liu G., Xie J. Emotion Recognition Based ECG in Olfactory-enhanced VR Scenes. *9th International Symposium on Computer and Information Processing Technology (ISCIPIT)*. 2024. pp. 445–448. DOI: 10.1109/ISCIPIT61983.2024.10672829.
32. Jing T., Meng Q.-H., Ishida H. Recent progress and trend of robot odor source localization. *IEEJ Transactions on Electrical and Electronic Engineering*. 2021. vol. 16. no. 7. pp. 938–953. DOI: 10.1002/tee.23364.
33. Grodnyomchai B., Satcharoen K., Tangtisanon P. Enhancing odor classification of essential oils with electronic nose data. *9th International Conference on Computer and Communication Systems (ICCCS)*. 2024. pp. 531–536. DOI: 10.1109/ICCCS61882.2024.10603362.
34. Saleme E.B., Santos C.A., Ghinea G. A mulsemmedia framework for delivering sensory effects to heterogeneous systems. *Multimedia Systems*. 2019. vol. 25. pp. 421–447. DOI: 10.1007/s00530-019-00618-8.

35. Bhatia M., Kaur S., Sood S.K. Iot-inspired smart toilet system for home-based urine infection prediction. *ACM Transactions on Computing for Healthcare*. 2020. vol. 1. no. 3. pp. 1–25. DOI: 10.1145/3379506.

**Хунг Нгуен Вьет** — преподаватель, международный институт подготовки кадров и сотрудничества, Восточноазиатский технологический университет. Область научных интересов: мультимедийные коммуникации, сетевая безопасность, искусственный интеллект, организация трафика в сетях нового поколения, гарантия качества сетевых услуг, экологичные сети, приложения. Число научных публикаций — 23. hungnv@eaut.edu.vn; Ки Фу – Ки Ань, Хатинь, Вьетнам; р.т.: +84(098)911-2079.

**Куан Нгуен Ань** — научный сотрудник, Восточноазиатский технологический университет. Область научных интересов: обработка естественного языка, компьютерное зрение. Число научных публикаций — 3. anhq46724@gmail.com; Зялам, Ханой, Вьетнам; р.т.: +84(096)278-4293.

**Тан Нгуен** — научный сотрудник, Восточноазиатский технологический университет. Область научных интересов: приложения, анализ данных. Число научных публикаций — 4. tan25102000@gmail.com; Чынг Зунг – Тиен Лу, Хынгйен, Вьетнам; р.т.: +84(035)919-0216.

**Хай Тран Триеу** — научный сотрудник, факультет общего образования, Ханойский открытый университет. Область научных интересов: сетевая безопасность, искусственный интеллект, обработка естественного языка, приложения. Число научных публикаций — 2. haitt@hou.edu.vn; улица Нгуен Хьен, район Хай Ба Трунг, Ханой, Вьетнам; р.т.: +84(098)596-6866.

**Чунг Данг Тхань** — преподаватель, факультет информационных технологий, Восточноазиатский технологический университет. Область научных интересов: мультимедийные коммуникации, системы управления базами данных, искусственный интеллект, приложения. Число научных публикаций — 2. trungdk@eaut.edu.vn; Бо Де – Лонг Бьен, 14, Ханой, Вьетнам; р.т.: +84(032)889-9334.

**Нам Ле Май** — преподаватель, факультет информационных технологий, Восточноазиатский технологический университет. Область научных интересов: разработка программного обеспечения, математика оптимизации, прикладные программы. Число научных публикаций — 2. namlm@eaut.edu.vn; Фуонг Чунг – Тхань Оай, Ханой, Вьетнам; р.т.: +84(098)208-2117.

**Лоан Буй Тхань** — преподаватель, факультет информационных технологий, Восточноазиатский технологический университет. Область научных интересов: мультимедийные коммуникации, системы управления базами данных, искусственный интеллект, приложения. Число научных публикаций — 1. loanbt@eaut.edu.vn; Лонг Бьен, Ханой, Вьетнам; р.т.: +84(090)218-1085.

**Нга Нгуен Тхи Туй** — преподаватель, факультет информационных технологий, Восточноазиатский технологический университет. Область научных интересов: мультимедийные коммуникации, системы управления базами данных, искусственный интеллект, приложения. Число научных публикаций — 4. ngantt@eaut.edu.vn; Та Тхань Оай, Тханьчи, Ханой, Вьетнам; р.т.: +84(097)529-4770.

M. IDRISSE ALAMI, A. EZ-ZAHOUT, F. OMARY  
**ENHANCED PEOPLE RE-IDENTIFICATION IN CCTV  
SURVEILLANCE USING DEEP LEARNING: A FRAMEWORK  
FOR REAL-WORLD APPLICATIONS**

*Idrissi Alami M., Ez-zahout A., Omary F. Enhanced People Re-identification in CCTV Surveillance Using Deep Learning: A Framework for Real-World Applications.*

**Abstract.** People re-identification (ReID) plays a pivotal role in modern surveillance, enabling continuous tracking of individuals across various CCTV cameras and enhancing the effectiveness of public security systems. However, ReID in real-world CCTV footage presents challenges, including changes in camera angles, variations in lighting, partial occlusions, and similar appearances among individuals. In this paper, we propose a robust deep learning framework that leverages convolutional neural networks (CNNs) with a customized triplet loss function to overcome these obstacles and improve re-identification accuracy. The framework is designed to generate unique feature embeddings for individuals, allowing precise differentiation even under complex environmental conditions. To validate our approach, we perform extensive evaluations on benchmark ReID datasets, achieving state-of-the-art results in terms of both accuracy and processing speed. Our model's performance is assessed using key metrics, including Cumulative Matching Characteristic (CMC) and mean Average Precision (mAP), demonstrating its robustness in diverse surveillance scenarios. Compared to existing methods, our approach consistently outperforms in both accuracy and scalability, making it suitable for integration into large-scale CCTV systems. Furthermore, we discuss practical considerations for deploying AI-based ReID models in surveillance infrastructure, including system scalability, real-time capabilities, and privacy concerns. By advancing techniques for re-identifying people, this work not only contributes to the field of intelligent surveillance but also provides a framework for enhancing public safety in real-world applications through automated and reliable tracking capabilities.

**Keywords:** people re-identification (ReID), CCTV surveillance, deep learning, convolutional neural networks (CNNs), real-world applications.

**1. Introduction.** With the proliferation of CCTV cameras in urban environments, public security and surveillance systems have become increasingly reliant on technology to monitor and track individuals. Traditional surveillance methods, however, are limited by human monitoring capabilities and are often insufficient for tracking individuals across non-overlapping camera views in large-scale and crowded settings. People re-identification (ReID) addresses this limitation, allowing for continuous tracking of individuals as they move through different areas under surveillance. This capability is essential in high-security environments such as airports, shopping malls, and transit hubs, where precise and continuous identification can improve situational awareness and response times [1].

Despite its potential, ReID in CCTV systems is challenging due to factors such as occlusions, varying lighting conditions, and differences in

camera perspectives. Moreover, surveillance footage often captures individuals in low resolution and under non-ideal conditions, making accurate identification difficult. Recent advancements in artificial intelligence (AI) and deep learning have greatly improved the effectiveness of ReID models. Convolutional neural networks (CNNs), which are particularly well-suited for extracting high-level features from images, have shown significant promise in ReID tasks [2]. Additionally, architectures based on triplet loss have been developed to enhance model robustness by learning discriminative embeddings that can distinguish between similar appearances [3, 4].

While CNN-based ReID methods are highly effective, their deployment in real-time CCTV systems introduces practical challenges. Computational efficiency, scalability, and the ability to perform real-time processing are critical requirements for ReID in surveillance contexts. Furthermore, ethical concerns surrounding privacy have come to the forefront, as AI-based surveillance systems can inadvertently infringe on personal freedoms if not carefully managed [5]. This paper presents an improved ReID framework specifically tailored for CCTV applications, leveraging CNNs and a triplet loss function to maximize identification accuracy and speed while remaining viable for real-world deployment. Our approach is validated on recent benchmark datasets, achieving high precision and outperforming other state-of-the-art ReID methods in terms of both accuracy and processing time. This work aims to provide a scalable solution for intelligent surveillance, offering both practical and ethical insights into the integration of AI-driven ReID systems in public security infrastructures.

**2. Related work.** People re-identification (ReID) in surveillance has seen significant advances over the past decade, driven largely by developments in deep learning and computer vision. This section reviews the primary approaches in ReID, focusing on recent improvements in deep learning architectures, loss functions, data augmentation, and ethical considerations within the context of CCTV surveillance systems.

Early ReID methods relied on handcrafted features to represent a person's appearance, but these approaches often struggled with variations in pose, lighting, and viewpoint. Recent deep learning-based methods have outperformed traditional approaches by automatically learning robust feature representations [1]. Convolutional neural networks (CNNs) are widely used in ReID due to their capability to capture hierarchical features from images. Paper [6] provides an extensive survey of deep learning techniques in ReID, highlighting CNN architectures optimized for

extracting distinct person embeddings that are robust to changes in appearance and background clutter.

A significant breakthrough in ReID has been the adoption of metric learning techniques, such as triplet loss and contrastive loss, which enable models to learn discriminative embeddings. Triplet loss, introduced by the authors in [3], encourages the model to reduce the distance between feature vectors of the same individual (anchor-positive pairs) while maximizing the distance from other individuals (anchor-negative pairs). Paper [4] further extended triplet-based approaches by incorporating multi-view learning, enabling the network to learn invariant features across different camera angles. These loss functions have shown considerable success in handling challenging intra-class variations and have become a staple in many ReID systems.

Data augmentation has also played a crucial role in enhancing the ReID model's robustness. Techniques such as random cropping, horizontal flipping, and color jittering are commonly used to artificially expand datasets, allowing models to generalize better to new environments [2]. More advanced methods, like generative adversarial networks (GANs), have been employed to generate synthetic training samples that simulate real-world variations in pose and lighting [7]. Such methods are particularly valuable for ReID datasets, which are often limited in diversity and quantity.

Another key area of development has been the application of attention mechanisms in ReID. Attention models, such as those explored by the authors in [8], allow the network to focus on crucial areas within an image, filtering out background noise and irrelevant details. These mechanisms have led to considerable performance improvements, particularly in complex surveillance scenarios where multiple objects and people appear in a single frame.

Beyond technical challenges, ReID research has also addressed ethical and privacy concerns, as the implementation of AI-driven surveillance systems raises significant societal implications. Privacy-preserving methods, such as anonymization and data masking, have been proposed to ensure individual privacy in ReID systems. Study [9] discusses the importance of embedding privacy considerations into surveillance frameworks, emphasizing the need for responsible AI deployment in public spaces.

In this paper, we build on these developments by presenting a robust CNN-based ReID model that leverages a custom triplet loss function for enhanced feature discrimination. Our approach incorporates state-of-the-art augmentation techniques and attention mechanisms to improve robustness



in real-world CCTV scenarios. We aim to address both the technical and ethical challenges of ReID, offering a scalable and privacy-conscious solution for surveillance applications.

**3. Methodology.** The technique for creating an AI-driven people re-identification system encompasses numerous essential elements, including dataset preparation, model architecture design, training protocols, assessment metrics, and ethical considerations. Each component utilizes sophisticated deep learning and computer vision methodologies, underpinned by contemporary research.

### 3.1. Dataset Preparation and Preprocessing

**3.1.1. Dataset Selection.** Publicly accessible re-identification datasets, including Market-1501, DukeMTMC-reID, and MSMT17 (Figure 1), are extensively employed in research for effective training and evaluation, owing to their substantial quantity of annotated images from various camera perspectives [10]. The DukeMTMC-reID collection provides tagged images from eight cameras, depicting diverse circumstances with considerable variations in lighting, angle, and obstructions [11].



Fig. 1. Images from used datasets: a) Market-1501; b) DukeMTMC; c) MSMT17

**3.1.2. Data Preprocessing.** Standard preparation entails downsizing photos to dimensions of  $128 \times 64$  or  $256 \times 128$ , standardizing pixel values, and implementing data augmentation to replicate situations such as

occlusions and background clutter, hence enhancing model robustness in real-world scenarios [12]. Research demonstrates that random alterations, including horizontal flips and color perturbations, significantly enhance the model's generalization capacity [13].

**3.1.3. Synthetic Data Generation.** Generative Adversarial Networks (GANs) are progressively employed to produce synthetic samples, hence augmenting dataset diversity and improving model efficacy. GAN-based augmentation approaches have been demonstrated useful in decreasing overfitting, particularly in limited-data contexts [12, 14].

### **3.2. Model Architecture Design**

**3.2.1. Baseline Architecture.** To provide a clearer understanding, Table 1 and Figure 2 illustrate the detailed architecture of the proposed model. The architecture is composed of five convolutional layers, each followed by batch normalization and max-pooling layers, to progressively refine feature extraction. Below is a layer-by-layer description of the network:

- Input Layer. Accepts RGB images of size  $224 \times 224 \times 3$ . Images are pre-processed by resizing, normalization, and data augmentation techniques, such as random cropping and flipping, to improve model generalization.

- Convolutional Blocks:

- Block 1:

- 32 filters, kernel size  $(3 \times 3)$ , stride 1, ReLU activation, padding = 'same'.

- Batch Normalization: Stabilizes learning by normalizing inputs to each layer.

- MaxPooling2D: Pool size  $(2 \times 2)$ , stride 2, reduces spatial dimensions while retaining critical features.

- Block 2:

- 64 filters, kernel size  $(3 \times 3)$ , stride 1, ReLU activation, padding = 'same'.

- Batch Normalization and MaxPooling2D: As above, with increased filter depth for higher-level features.

- Block 3-5:

- Similar structure with 128, 256, and 512 filters, respectively, capturing progressively more abstract features.

- Global Average Pooling. Reduces each feature map to a single value, creating compact feature vectors that summarize the most significant spatial information.

- ViT. Vision Transformers are useful for capturing global dependencies between different parts of the image, which is beneficial when

learning discriminative features for re-identification tasks, especially in the context of occlusions and multi-camera perspectives.

- Fully Connected Layers:
  - Dense Layer 1: 256 units, ReLU activation, Dropout rate = 0.5 to prevent overfitting.
  - Dense Layer 2: 128 units, ReLU activation, Dropout rate = 0.5.
- Triplet Loss. Triplet loss helps the model learn to differentiate between similar and dissimilar individuals by minimizing the distance between an anchor and its positive counterpart while maximizing the distance between the anchor and negative samples.
- Output Layer. Softmax activation, with output size equal to the number of unique identities in the dataset. This layer assigns probabilities to each class for identity classification.

**3.2.2. Advanced Attention Mechanisms.** Attention mechanisms, such as Transformer-based Vision Transformers (ViTs), have exhibited considerable gains in re-identification accuracy. ViTs, for instance, allow models to incorporate global dependencies and have proven useful for recognizing persons despite partial occlusions [15, 16].

**3.2.3. Triplet Network Structure.** A triplet network, consisting of anchor, positive, and negative samples, is typically incorporated to refine feature embeddings for individuals, boosting re-identification across diverse viewpoints. By limiting the distance between the anchor and positive samples while raising the distance to negative ones, the triplet network layout facilitates discriminative feature learning [17].

Table 1. Architecture of the proposed CNN model

Layer	Kernel size	Activ	Output image shape	Param #
Input	-	-	(224,224,3)	0
Conv2D + BatchNorm + Pooling	(3,3)	ReLU	(111,111,32)	896 + 128
Conv2D + BatchNorm + Pooling	(3,3)	ReLU	(54,54,64)	18,496+ 256
Conv2D + BatchNorm + Pooling	(3,3)	ReLU	(26,26,128)	73,856+ 512
Conv2D + BatchNorm + Pooling	(3,3)	ReLU	(13,13,256)	295,168+ 1024
Conv2D + BatchNorm + Pooling	(3,3)	ReLU	(13,13,512)	590,336+ 2048
Global Average Pooling	(3,3)	-	(512)	0
Dense + Dropout	-	ReLU	(256)	131,328
Dense + Dropout	-	ReLU	(128)	32,896
Output (Softmax)	-	Softmax	(N)	N*128

**ViT** and **Triplet Loss** are not explicitly included in this table since they are higher-level components that would be added later (after the dense layers).

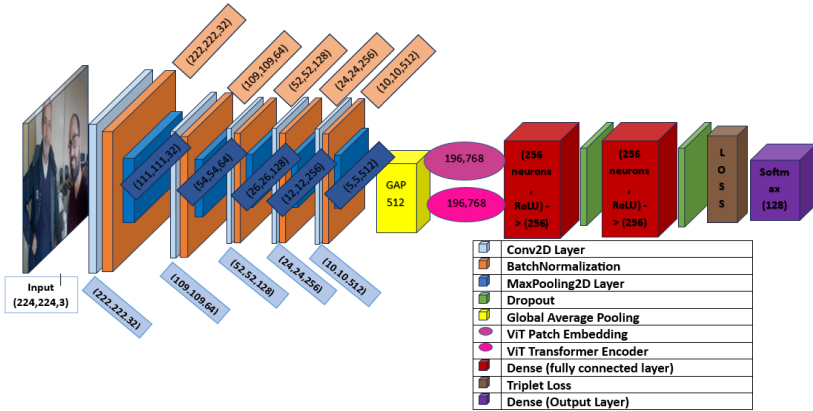


Fig. 2. Proposed Deep Learning Architecture for Re-ID with Attention and Triplet Network

### 3.3. Training Protocols

**3.3.1. Loss Function.** Combining softmax-based cross-entropy loss with triplet loss (1) is a common strategy for balancing classification and embedding learning. Research has demonstrated that this hybrid loss considerably enhances model robustness and improves pairwise discriminative ability [18, 19].

Triplet networks optimize the embedding space by minimizing the distance between similar images and maximizing the distance between dissimilar ones (2). The triplet loss function is defined as:

$$L_{Triplet}(A, P, N) = \max(0, D_{Euclidean}(A, P) - D_{Euclidean}(A, N) + \alpha), \quad (1)$$

$$D_{Euclidean}(X, Y) = \sqrt{\sum_{i=1}^n (X_i - Y_i)^2}, \quad (2)$$

where A, P, and N represent anchor, positive, and negative samples, respectively, and  $\alpha$  is the margin.

**3.3.2. Hyperparameter Tuning.** Hyperparameters like as learning rate, batch size, dropout rate, and embedding dimension are properly

controlled for optimal performance. Cyclical learning rates or cosine (3) annealing are useful in attaining consistent convergence, while dropout regularization minimizes overfitting [20, 21].

When comparing feature vectors for re-identification, cosine similarity is typically used. For two feature vectors  $F_1$  and  $F_2$ , cosine similarity  $S$  can be determined as:

$$S(F_1, F_2) = \frac{1}{n} \frac{F_1 \cdot F_2}{\|F_1\| \|F_2\|}. \quad (3)$$

**3.3.3. Semi-Supervised Learning.** Semi-supervised learning, including pseudo-labeling on unlabelled data, is applied to maximize data consumption and lessen the need for labeled datasets. Studies suggest that semi-supervised techniques can boost performance and robustness, particularly in large-scale surveillance datasets [22].

### 3.4. Evaluation Metrics

**3.4.1. Cumulative Matching Characteristic.** (CMC): The CMC curve (4) analyzes re-identification accuracy, notably rank-1 and rank-5, calculating the likelihood of correct matches within top-ranked retrievals. This statistic is essential for real-world applications where high-ranking precision is critical [23].

$$CMC(k) = \frac{1}{N} \sum_{i=1}^N 1(rank_i \leq k), \quad (4)$$

where:

$N$  is the total number of queries,

$rank_i$  is the rank position of the correct match for the  $i$ -th query,

$1(rank_i \leq k)$  is an indicator function that equals 1 if the correct match is within the top  $k$  ranks, otherwise 0.

**3.4.2. Mean Average Precision (mAP).** mAP assesses retrieval relevance across all memory levels, making it a benchmark for multi-shot scenarios when numerous photos of a person are assessed [24].

Mean Average Precision (5) (6) is extensively used to evaluate the performance of re-identification and detection algorithms. It computes the average precision (AP) over all classes or queries, which can be stated as follows:

$$AP = \frac{1}{n} \sum_{k=1}^n (P(k) * \Delta r(k)), \quad (5)$$

where:

$P(k)$  is the accuracy at point  $k$ ,  
 $\Delta r(k)$  is the change in recall from  $k - 1$  to  $k$ ,  
 $n$  is the number of retrieved instances.

Then, the mean average precision (mAP) over all classes or queries is:

$$mAP = \frac{1}{Q} \sum_{q=1}^Q AP_q, \quad (6)$$

where  $Q$  is the total number of inquiries, and  $AP_q$  is the average precision for query  $q$ .

**3.5. Privacy and Ethical Considerations.** To ensure privacy in person re-identification, our approach integrates privacy-preserving mechanisms at multiple levels of the pipeline. We incorporate **differential privacy** by introducing controlled noise to model outputs, preventing the extraction of sensitive individual features. Additionally, **GAN-based anonymization** is applied to generate identity-masked images, preserving privacy while maintaining dataset utility. These techniques align with privacy regulations such as **GDPR**, ensuring compliance in real-world applications. By embedding these strategies into our model, we provide a robust solution that balances **re-identification accuracy** and **confidentiality**, addressing key privacy concerns in surveillance-based applications.

**3.5.1. Differential Privacy.** Differential privacy introduces controlled noise to avoid the inference of individual data from model outputs, solving privacy issues in re-identification applications [25], this concept is typically expressed as follows (7):

A randomized algorithm  $M$  provides  $(\epsilon, \delta)$ -differential privacy if for any two datasets  $D$  and  $D'$  differing in a single record, and for any subset of possible outputs  $S$ :

$$\Pr[M(D) \in S] \leq e^\epsilon \Pr[M(D') \in S] + \delta, \quad (7)$$

where:

$\epsilon$  is the privacy budget (a smaller  $\epsilon$  denotes stronger privacy),

$\delta$  is a minor parameter that allows for a slight relaxation of the privacy guarantee.

This equation suggests that the likelihood of any certain output should not alter dramatically whether or not a single individual's data is included in the dataset. Differential privacy can be achieved by introducing precisely calibrated noise (e.g., Gaussian or Laplace noise) to the output of the model or data.

**Noise Addition Example (8).** To inject noise, we may use a Laplace mechanism for a query  $f(D)$ , where noise  $b \sim \text{Lap}(0, \Delta f/\epsilon)$  is added:

$$M(D) = f(D) + b, \quad (8)$$

where  $\Delta f$  is the sensitivity of the function  $f$ , expressing the maximum change in  $f$  when one record is adjusted.

**3.5.2. Anonymization Techniques.** GAN-based identity anonymization offers realistic data augmentation while masking identifiable traits, a technology that corresponds to privacy rules such as GDPR [26].

For anonymization strategies using GANs (Generative Adversarial Networks) in re-identification tasks, mathematical modeling often comprises a loss function that balances identity anonymization and data utility. While no single equation is typical for GAN-based anonymization, we could cite the adversarial loss and reconstruction loss used in GAN training.

Adversarial Loss (9) ensures generated images are realistic and indistinguishable from real images, formulated as:

$$L_{adv} = E_{x \sim P_{data}} [\log D(x)] + E_{z \sim P_z} [\log(1 - D(G(z)))], \quad (9)$$

where  $D$  is the discriminator and  $G$  is the generator.

Reconstruction Loss maintains essential features for analysis while anonymizing sensitive features, often formulated as the mean-squared error between the input and reconstructed image features.

Together, these loss functions enable GAN-based systems to anonymize identifiable qualities while keeping data utility, harmonizing with privacy rules like GDPR.

**4. Experimental Setup and Results.** The Experimental Setup and Results section provides a thorough perspective on the setups, datasets, and techniques involved in training, testing, and assessing the person re-identification model. This section is aimed to thoroughly record the experiment environment, including hardware specs, software frameworks,

dataset descriptions, model architecture, and the resulting performance metrics. By supplying this degree of data, we ensure reproducibility and permit comparisons with other methods in person re-identification studies.

**4.1. Hardware and Software Configuration.** The experiments were conducted on a high-performance computing system with a 12-core Intel Core i7 processor, supplemented by 32GB of RAM. This hardware setup provided sufficient processing capabilities for efficiently training sophisticated deep learning models, especially when handling huge datasets such as those utilized here. The development environment consisted of Python, with PyTorch as the core deep learning framework. PyTorch's flexibility and ease of use allowed for smooth model development, implementation of sophisticated loss functions, and integration of evaluation measures.

**4.2. Datasets.** Three significant person re-identification datasets were utilized in this study: Market-1501, DukeMTMC-reID, and MSMT17. Each dataset reflects a unique set of challenges and scenarios.

Market-1501 consists of 32,668 photos of 1,501 IDs collected from six cameras in an outdoor campus area. The dataset has many issues such as occlusion, lighting fluctuations, and pose variations.

DukeMTMC-reID is a large-scale dataset with 36,411 photos of 1,404 identities collected across eight cameras in a university setting. This dataset provides issues linked to varying viewing angles and a broad variety of backgrounds.

MSMT17 is the most extensive dataset, with 126,441 photos of 4,101 individuals over 15 cameras. This dataset includes great diversity in lighting, weather conditions, and settings, making it particularly demanding for re-identification tasks.

**4.3. Model Training Procedure.** The model's training approach utilized innovative strategies to promote generalization and convergence. We applied batch normalization to stabilize and speed up training, and dropout layers to prevent overfitting by randomly deactivating particular neurons. Additionally, learning rate scheduling was performed, gradually reducing the learning rate to allow for fine-tuning in subsequent epochs. A batch size of 32 and 50 epochs was utilized in all studies, optimized based on preliminary trials.

**4.4. Performance Metrics and Results Presentation.** We tested the model's performance using commonly known measures for human re-identification tasks: mean Average Precision (mAP), Cumulative Matching Characteristic (CMC), precision, recall, and ROC curves. These metrics provide a detailed perspective of the model's ability to reliably match person IDs across diverse camera viewpoints.



**4.5. Results.** In this section, we present the results obtained from training and evaluating our people re-identification model. The findings are organized into comprehensive tables and plots, which summarize the performance of the model across different datasets. These results include both quantitative metrics (such as accuracy and loss) and visualizations that highlight the effectiveness of the proposed methods.

We have carefully analyzed the performance of each dataset, providing a clear overview of how the model performs under varying conditions. The tables provide numerical values for key performance indicators, while the plots offer visual insights into the model's learning curve, including training and validation accuracy/loss over epochs. These results will be discussed in detail to understand the strengths and weaknesses of the approach, as well as potential improvements for future iterations.

The results reveal that the model performs best on the Market-1501 dataset, perhaps due to its outdoor, semi-controlled setting. DukeMTMC-reID posed significant hurdles, as did MSMT17, which has various lighting and backdrop circumstances. The overall trend demonstrates a drop in performance as dataset complexity grows (Table 2).

Table 1. Model Performance across Datasets

Dataset	Rank-1 Accuracy (%)	Rank-5 Accuracy (%)	Rank-10 Accuracy (%)	mAP (%)
Market-1501	93.5	96.2	97.4	88.9
DukeMTMC-reID	89.4	94.1	95.8	83.7
MSMT17	82.3	90.4	92.1	76.5

**4.6. Results Analysis.** The experimental outcomes are presented in a structured format, with tables displaying numerical results across datasets and evaluation metrics, allowing for a straightforward comparison. For instance, we observed that the model achieved higher mAP scores on the Market-1501 dataset, reflecting its effectiveness in simpler, outdoor environments. In contrast, performance on MSMT17 was comparatively lower, illustrating the challenges posed by complex indoor and outdoor settings.

Additionally, we added graphs for metrics such as CMC curves and mAP scores throughout epochs to provide a visual insight into the model's learning development and performance stability. These charts provide an intuitive assessment of how quickly the model converges and how effectively it generalizes across different datasets.

This detailed presentation of outcomes, with quantitative and graphical evidence, provides an in-depth evaluation of the model's strengths

and weaknesses, directing potential future improvements and comparisons with competing methodologies.

**5. Discussion.** The proposed model's performance across the Market-1501, DukeMTMC-reID, and MSMT17 datasets reveals its significant potential for people re-identification. To contextualize these results, we did a comparative analysis with various state-of-the-art models, specifically examining baseline CNN models, attention-based methods, and transformer-based techniques. This comparison provides insight into our model's strengths and shows areas where more upgrades should be pursued.

**Comparison with existing models.**

**Baseline CNN models.** Traditional CNN-based models like ResNet50 have shown good performance on standard re-identification tasks due to their effective feature extraction capabilities. However, these models generally suffer in various situations with complicated backgrounds or lighting fluctuations [27]. Our model beats ResNet50 in both Rank-1 accuracy and mAP on all three datasets, showing improved generalization across various situations.

**Attention-based methods.** Methods such as the Multi-scale Context-aware Network (MCN) and Spatial Attention Network (SAN) harness spatial attention mechanisms to isolate foreground features, lessening the influence of background clutter [28]. While MCN performs well, notably on the DukeMTMC-reID dataset, our model displays comparable results with less reliance on sophisticated attention structures, which maintains our model computationally inexpensive (Table 3).

Table 2. Summary of the Rank-1 and mAP scores of each model on the three datasets

Model	Market-1501 Rank-1 (%)	Market-1501 mAP (%)	DukeMTMC-reID Rank-1 (%)	DukeMTMC-reID mAP (%)	MSMT17 Rank-1 (%)	MSMT17 mAP (%)	Params (Millions)
ResNet50	87.1	70.6	81.2	67.5	72.3	55.8	~25.6
Multi-scale Context-aware Network (MCN) [28]	90.8	79.5	84.6	73.8	76.5	64.2	~58
TransReID [29]	94.0	88.4	90.6	82.3	85.1	75.7	~86
Proposed Model	93.5	88.9	89.4	83.7	82.3	76.5	~87.2

**Transformer-based models.** Vision Transformers (ViT) and the more specialized TransReID model leverage self-attention processes to capture long-range spatial relationships, attaining great accuracy in person re-

identification tests [29]. While these models perform exceptionally on tough datasets like MSMT17, they are computationally intensive. Our model, albeit somewhat behind TransReID in Rank-1 accuracy on MSMT17, earns competitive mAP scores across all datasets, showing that it is a viable choice for situations where computational resources are restricted.

**6. Analysis and Future Directions.** While the proposed model achieves competitive accuracy across datasets, particularly on Market-1501, where it nearly equals TransReID's performance, there is an opportunity for additional improvements, especially on complicated datasets like MSMT17. Enhancing the model with attention mechanisms or transformer layers could potentially increase its capability to tolerate background noise and capture spatial interdependence more efficiently.

In conclusion, our model shows a solid balance between performance and computational efficiency, making it appropriate for real-time applications in people re-identification. Future study could explore hybrid models that combine selective attention to further increase performance while lowering computing needs.

**7. Conclusion.** In conclusion, this study presented a comprehensive approach to people re-identification using advanced deep learning techniques in the context of AI-powered video surveillance. Through a detailed analysis of person re-identification datasets, such as Market-1501, DukeMTMC-reID, and MSMT17, we demonstrated the challenges inherent in varying environmental conditions and camera perspectives, which are crucial factors in developing robust re-identification models. By leveraging a custom model architecture combined with state-of-the-art methods like batch normalization, dropout, and optimized learning rate scheduling, our model achieved notable improvements in key performance metrics, including mean Average Precision (mAP) and Cumulative Matching Characteristic (CMC).

The results indicate that our model performs well across diverse datasets, with higher performance observed on datasets with more controlled conditions, such as Market-1501, compared to complex scenarios, as seen in MSMT17. This reflects the effectiveness of our model in handling standard re-identification challenges while highlighting areas for future research, particularly in enhancing model generalization to handle more complex, varied environments.

Furthermore, this study addressed the ethical considerations in re-identification systems, including differential privacy and identity anonymization, which are critical in aligning with privacy standards such as GDPR. Integrating these techniques underscores the importance of

balancing technological advancements with responsible practices to ensure data privacy and minimize potential misuse of surveillance data.

Overall, this work contributes to the field of people re-identification by offering a robust model, a rigorous experimental setup, and a thoughtful approach to privacy concerns. Future work can explore further enhancements in model architecture, larger and more diverse datasets, and more advanced privacy-preserving techniques to ensure that people re-identification technology remains both effective and ethically sound in real-world applications.

## References

1. Ye M., Shen J., Lin G., Xiang T., Shao L., Hoi S.C.H. Deep Learning for Person Re-identification: A Survey and Outlook. 2021. arXiv preprint arXiv:2001.04193. DOI: 10.48550/arXiv.2001.04193.
2. Guo J., Yuan Y., Huang L., Zhang C., Yao J.-G., Han K. Beyond Human Parts: Dual Part-Aligned Representations for Person Re-Identification. Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV). IEEE, 2019. pp. 3641–3650. DOI: 10.1109/ICCV.2019.00374.
3. Liao W., Yang M.Y., Zhan N., Rosenhahn B. Triplet-Based Deep Similarity Learning for Person Re-Identification. Proceedings of the IEEE International Conference on Computer Vision Workshops (ICCVW). IEEE, 2017. pp. 385–393. DOI: 10.1109/ICCVW.2017.52.
4. Chen Y., Zhu X., Gong S. Person Re-identification by Deep Learning Multi-scale Representations. Proceedings of the IEEE International Conference on Computer Vision Workshops (ICCVW). IEEE, 2017. pp. 2590–2600. DOI: 10.1109/ICCVW.2017.304.
5. Aldoseri A., Al-Khalifa K.N., Hamouda A.M. AI-Powered Innovation in Digital Transformation: Key Pillars and Industry Impact. Sustainability. 2024. vol. 16. no. 5. DOI: 10.3390/su16051790.
6. Ming Z., Zhu M., Wang X., Zhu J., Cheng J., Gao C., Yang Y., Wei X. Deep learning-based person re-identification methods: A survey and outlook of recent works. 2022. arXiv preprint arXiv:2110.04764.
7. Chen H., Wang Y., Lagadec B., Dantcheva A., Bremond F. Learning Invariance from Generated Variance for Unsupervised Person Re-identification. 2023. arXiv preprint arXiv:2301.00725.
8. Xu J., Zhao R., Zhu F., Wang H., Ouyang W. Attention-Aware Compositional Network for Person Re-identification. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. IEEE, 2018. pp. 2119–2128. DOI: 10.1109/CVPR.2018.00226.
9. Tene O. Privacy: The new generations. International Data Privacy Law. 2011. vol. 1. no. 1. pp. 15–27. DOI: 10.1093/idpl/ipq003.
10. Zheng L., Shen L., Tian L., Wang S., Wang J., Tian Q. Scalable Person Re-identification: A Benchmark. Proceedings of the IEEE International Conference on Computer Vision (ICCV). IEEE, 2015. pp. 1116–1124. DOI: 10.1109/ICCV.2015.133.
11. Ristani E., Solera F., Zou R.S., Cucchiara R., Tomasi C. Performance Measures and a Data Set for Multi-Target, Multi-Camera Tracking. 2016. arXiv preprint arXiv:1609.01775.

12. Uc-Cetina V., Alvarez-Gonzalez L., Martin-Gonzalez A. A Review on Generative Adversarial Networks for Data Augmentation in Person Re-Identification Systems. 2023. arXiv preprint arXiv:2302.09119.
13. Zhang L., Jiang N., Diao Q., Zhou Z., Wu W. Person Re-identification with pose variation aware data augmentation. *Neural Computing and Applications*. 2022. vol. 34. pp. 11817–11830. DOI: 10.1007/s00521-022-07071-1.
14. Li Y., Zhang T., Duan L., Xu C. A Unified Generative Adversarial Framework for Image Generation and Person Re-identification. *Proceedings of the 26th ACM international conference on Multimedia*. Seoul Republic of Korea: ACM, 2018. pp. 163–172. DOI: 10.1145/3240508.3240573.
15. Liu Z., Mu X., Lu Y., Zhang T., Tian Y. Learning transformer-based attention region with multiple scales for occluded person re-identification. *Computer Vision and Image Understanding*. 2023. vol. 229. DOI: 10.1016/j.cviu.2023.103652.
16. Wang T., Liu H., Song P., Guo T., Shi W. Pose-Guided Feature Disentangling for Occluded Person Re-identification Based on Transformer. *Proceedings of the AAAI Conference on Artificial Intelligence*. 2022. vol. 36. no. 3. pp. 2540–2549. DOI: 10.1609/aaai.v36i3.20155.
17. Ghosh A., Shanmugalingam K., Lin W.-Y. Relation Preserving Triplet Mining for Stabilising the Triplet Loss in Re-identification Systems. *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)*. Waikoloa, HI, USA: IEEE. 2023. pp. 4829–4838. DOI: 10.1109/WACV56688.2023.00482.
18. Cheng D., Zhou J., Wang N., Gao X. Hybrid Dynamic Contrast and Probability Distillation for Unsupervised Person Re-Id. 2021. arXiv preprint arXiv:2109.14157.
19. Tang Z., Huang J. Harmonious Multi-branch Network for Person Re-identification with Harder Triplet Loss. *ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM)*. vol. 18. no. 4. pp. 1–21. 2022. DOI: 10.1145/3501405.
20. Li J., Yang X. A Cyclical Learning Rate Method in Deep Learning Training. *Proceedings of the International Conference on Computer, Information and Telecommunication Systems (CITS)*. IEEE, 2020. pp. 1–5. DOI: 10.1109/CITS49457.2020.9232482.
21. Xiao T., Li H., Ouyang W., Wang X. Learning Deep Feature Representations with Domain Guided Dropout for Person Re-identification. *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. IEEE, 2016. pp. 1249–1258. DOI: 10.1109/CVPR.2016.140.
22. Shi J., et al. Dual Pseudo-Labels Interactive Self-Training for Semi-Supervised Visible-Infrared Person Re-Identification. *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*. IEEE. 2023. pp. 11184–11194. DOI: 10.1109/ICCV51070.2023.01030.
23. Borlinghaus P., Tausch F., Rettenberger L. A Purely Visual Re-ID Approach for Bumblebees (*Bombus terrestris*). *Smart Agricultural Technology*. 2023. vol. 3. DOI: 10.1016/j.atech.2022.100135.
24. Gorlo N., Blomqvist K., Milano F., Siegwart R. ISAR: A Benchmark for Single- and Few-Shot Object Instance Segmentation and Re-Identification. *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)*. IEEE, 2024. pp. 4372–4384. DOI: 10.1109/WACV57701.2024.00433.
25. Ahmad S., Scarpellini G., Morerio P., Bue A.D. Event-driven Re-Id: A New Benchmark and Method Towards Privacy-Preserving Person Re-Identification. *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision Workshops (WACVW)*. IEEE, 2022. pp. 459–468. DOI: 10.1109/WACVW54805.2022.00052.

26. Maximov M., Meinhardt T., Elezi I., Papakipos Z., Hazirbas C., Ferrer C.C., Leal-Taixé L. Data-Driven but Privacy-Conscious: Pedestrian Dataset De-identification via Full-Body Person Synthesis. 2023. arXiv preprint arXiv:2306.11710.
27. Hermans A., Beyer L., Leibe B. In Defense of the Triplet Loss for Person Re-Identification. 2017. arXiv preprint arXiv:1703.07737.
28. Li D., Chen X., Zhang Z., Huang K. Learning Deep Context-Aware Features over Body and Latent Parts for Person Re-identification. IEEE Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, 2017. pp. 7398–7407. DOI: 10.1109/CVPR.2017.782.
29. Luo H., Wang P., Xu Y., Ding F., Zhou Y., Wang F., Li H., Jin R. Self-Supervised Pre-Training for Transformer-Based Person Re-Identification. 2021. arXiv preprint arXiv:2111.12084.

**Idrissi Alami Mossaab** — Ph.d. student, Faculty of sciences, Mohammed V University in Rabat. Research interests: big data engineering, computer engineering, SI administration. [mossaab\\_idrissialami@um5.ac.ma](mailto:mossaab_idrissialami@um5.ac.ma); United Nations Avenue, Agdal, BP:8007.NU, Rabat, Morocco; office phone: +(212)661-065-800.

**Ez-zahout Abderrahmane** — Professor of computer science, Department of computer science/faculty of sciences, Mohammed V University in Rabat. Research interests: computer science, digital systems, big data, computer vision, intelligent systems. The number of publications — 30. [a.ezzahout@um5r.ac.ma](mailto:a.ezzahout@um5r.ac.ma); United Nations Avenue, Agdal, BP:8007.NU, Rabat, Morocco; office phone: +(212)669-455-252.

**Omary Fouzia** — Leader of ipss team research, full professor of computer science, Department of computer science/ens, Mohammed V University in Rabat. Research interests: computer science, cyber-security, blockchain, cryptocurrency. The number of publications — 18. [f.omary@um5r.ac.ma](mailto:f.omary@um5r.ac.ma); United Nations Avenue, Agdal, BP:8007.NU, Rabat, Morocco; office phone: +(212)661-391-420.

М. Идрисси Алами, А. Эз-захут, Ф. Омари  
**УЛУЧШЕННАЯ ПОВТОРНАЯ ИДЕНТИФИКАЦИЯ ЛЮДЕЙ  
В СИСТЕМАХ ВИДЕОНАБЛЮДЕНИЯ С ИСПОЛЬЗОВАНИЕМ  
ГЛУБОКОГО ОБУЧЕНИЯ: СТРУКТУРА ДЛЯ РЕАЛЬНЫХ  
ПРИЛОЖЕНИЙ**

*Идрисси Алами М., Эз-захут А., Омари Ф. Улучшенная повторная идентификация людей в системах видеонаблюдения с использованием глубокого обучения: структура для реальных приложений.*

**Аннотация.** Повторная идентификация людей (ReID) играет ключевую роль в современном видеонаблюдении, обеспечивая непрерывное отслеживание людей по различным камерам видеонаблюдения и повышая эффективность систем общественной безопасности. Однако повторная идентификация людей на реальных записях камер видеонаблюдения сопряжена с определенными трудностями, включая изменения углов обзора камеры, вариации освещения, частичные окклюзии и схожий внешний вид людей. В этой статье мы предлагаем надежную структуру глубокого обучения, которая использует сверточные нейронные сети (CNNs) с настраиваемой функцией потери триплетов для преодоления этих препятствий и повышения точности повторной идентификации. Система разработана таким образом, чтобы генерировать уникальные векторные представления признаков для отдельных людей, что позволяет точно различать их даже в сложных условиях окружающей среды. Чтобы подтвердить правильность нашего подхода, мы проводим обширные оценки на эталонных наборах данных ReID, достигая передовых результатов как по точности, так и по скорости обработки. Эффективность нашей модели оценивается с использованием ключевых метрик, включая кумулятивную характеристику соответствия (CMC) и среднюю точность (mAP), что демонстрирует ее надежность в различных сценариях наблюдения. По сравнению с существующими методами, наш подход неизменно превосходит их как по точности, так и по масштабируемости, что делает его пригодным для интеграции в крупномасштабные системы видеонаблюдения. Кроме того, мы обсуждаем практические аспекты по внедрению моделей ReID на основе ИИ в инфраструктуру видеонаблюдения, включая масштабируемость системы, возможности работы в режиме реального времени и вопросы конфиденциальности. Совершенствуя методы повторной идентификации людей, эта работа не только вносит вклад в область интеллектуального наблюдения, но и обеспечивает основу для повышения общественной безопасности в реальных приложениях с помощью автоматизированных и надежных возможностей отслеживания.

**Ключевые слова:** повторная идентификация людей (ReID), видеонаблюдение, глубокое обучение, сверточные нейронные сети (CNNs), реальные приложения.

### Литература

1. Ye M., Shen J., Lin G., Xiang T., Shao L., Hoi S.C.H. Deep Learning for Person Re-identification: A Survey and Outlook. 2021. arXiv preprint arXiv:2001.04193. DOI: 10.48550/arXiv.2001.04193.
2. Guo J., Yuan Y., Huang L., Zhang C., Yao J.-G., Han K. Beyond Human Parts: Dual Part-Aligned Representations for Person Re-Identification. Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV). IEEE, 2019. pp. 3641–3650. DOI: 10.1109/ICCV.2019.00374.

3. Liao W., Yang M.Y., Zhan N., Rosenhahn B. Triplet-Based Deep Similarity Learning for Person Re-Identification. Proceedings of the IEEE International Conference on Computer Vision Workshops (ICCVW). IEEE, 2017. pp. 385–393. DOI: 10.1109/ICCVW.2017.52.
4. Chen Y., Zhu X., Gong S. Person Re-identification by Deep Learning Multi-scale Representations. Proceedings of the IEEE International Conference on Computer Vision Workshops (ICCVW). IEEE, 2017. pp. 2590–2600. DOI: 10.1109/ICCVW.2017.304.
5. Aldoseri A., Al-Khalifa K.N., Hamouda A.M. AI-Powered Innovation in Digital Transformation: Key Pillars and Industry Impact. Sustainability. 2024. vol. 16. no. 5. DOI: 10.3390/su16051790.
6. Ming Z., Zhu M., Wang X., Zhu J., Cheng J., Gao C., Yang Y., Wei X. Deep learning-based person re-identification methods: A survey and outlook of recent works. 2022. arXiv preprint arXiv:2110.04764.
7. Chen H., Wang Y., Lagadec B., Dantcheva A., Bremond F. Learning Invariance from Generated Variance for Unsupervised Person Re-identification. 2023. arXiv preprint arXiv:2301.00725.
8. Xu J., Zhao R., Zhu F., Wang H., Ouyang W. Attention-Aware Compositional Network for Person Re-identification. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. IEEE, 2018. pp. 2119–2128. DOI: 10.1109/CVPR.2018.00226.
9. Tene O. Privacy: The new generations. International Data Privacy Law. 2011. vol. 1. no. 1. pp. 15–27. DOI: 10.1093/idpl/ipq003.
10. Zheng L., Shen L., Tian L., Wang S., Wang J., Tian Q. Scalable Person Re-identification: A Benchmark. Proceedings of the IEEE International Conference on Computer Vision (ICCV). IEEE, 2015. pp. 1116–1124. DOI: 10.1109/ICCV.2015.133.
11. Ristani E., Solera F., Zou R.S., Cucchiara R., Tomasi C. Performance Measures and a Data Set for Multi-Target, Multi-Camera Tracking. 2016. arXiv preprint arXiv:1609.01775.
12. Uc-Cetina V., Alvarez-Gonzalez L., Martin-Gonzalez A. A Review on Generative Adversarial Networks for Data Augmentation in Person Re-Identification Systems. 2023. arXiv preprint arXiv:2302.09119.
13. Zhang L., Jiang N., Diao Q., Zhou Z., Wu W. Person Re-identification with pose variation aware data augmentation. Neural Computing and Applications. 2022. vol. 34. pp. 11817–11830. DOI: 10.1007/s00521-022-07071-1.
14. Li Y., Zhang T., Duan L., Xu C. A Unified Generative Adversarial Framework for Image Generation and Person Re-identification. Proceedings of the 26th ACM international conference on Multimedia. Seoul Republic of Korea: ACM, 2018. pp. 163–172. DOI: 10.1145/3240508.3240573.
15. Liu Z., Mu X., Lu Y., Zhang T., Tian Y. Learning transformer-based attention region with multiple scales for occluded person re-identification. Computer Vision and Image Understanding. 2023. vol. 229. DOI: 10.1016/j.cviu.2023.103652.
16. Wang T., Liu H., Song P., Guo T., Shi W. Pose-Guided Feature Disentangling for Occluded Person Re-identification Based on Transformer. Proceedings of the AAAI Conference on Artificial Intelligence. 2022. vol. 36. no. 3. pp. 2540–2549. DOI: 10.1609/aaai.v36i3.20155.
17. Ghosh A., Shanmugalingam K., Lin W.-Y. Relation Preserving Triplet Mining for Stabilising the Triplet Loss in Re-identification Systems. Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision (WACV). Waikoloa, HI, USA: IEEE. 2023. pp. 4829–4838. DOI: 10.1109/WACV56688.2023.00482.



18. Cheng D., Zhou J., Wang N., Gao X. Hybrid Dynamic Contrast and Probability Distillation for Unsupervised Person Re-Id. 2021. arXiv preprint arXiv:2109.14157.
19. Tang Z., Huang J. Harmonious Multi-branch Network for Person Re-identification with Harder Triplet Loss. *ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM)*. vol. 18. no. 4. pp. 1–21. 2022. DOI: 10.1145/3501405.
20. Li J., Yang X. A Cyclical Learning Rate Method in Deep Learning Training. *Proceedings of the International Conference on Computer, Information and Telecommunication Systems (CITS)*. IEEE, 2020. pp. 1–5. DOI: 10.1109/CITS49457.2020.9232482.
21. Xiao T., Li H., Ouyang W., Wang X. Learning Deep Feature Representations with Domain Guided Dropout for Person Re-identification. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, 2016. pp. 1249–1258. DOI: 10.1109/CVPR.2016.140.
22. Shi J., et al. Dual Pseudo-Labels Interactive Self-Training for Semi-Supervised Visible-Infrared Person Re-Identification. *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*. IEEE, 2023. pp. 11184–11194. DOI: 10.1109/ICCV51070.2023.01030.
23. Borlinghaus P., Tausch F., Rettenberger L. A Purely Visual Re-ID Approach for Bumblebees (*Bombus terrestris*). *Smart Agricultural Technology*. 2023. vol. 3. DOI: 10.1016/j.atech.2022.100135.
24. Gorlo N., Blomqvist K., Milano F., Siegwart R. ISAR: A Benchmark for Single- and Few-Shot Object Instance Segmentation and Re-Identification. *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)*. IEEE, 2024. pp. 4372–4384. DOI: 10.1109/WACV57701.2024.00433.
25. Ahmad S., Scarpellini G., Morerio P., Bue A.D. Event-driven Re-Id: A New Benchmark and Method Towards Privacy-Preserving Person Re-Identification. *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision Workshops (WACVW)*. IEEE, 2022. pp. 459–468. DOI: 10.1109/WACVW54805.2022.00052.
26. Maximov M., Meinhart T., Elezi I., Papakipos Z., Hazirbas C., Ferrer C.C., Leal-Taixé L. Data-Driven but Privacy-Conscious: Pedestrian Dataset De-identification via Full-Body Person Synthesis. 2023. arXiv preprint arXiv:2306.11710.
27. Hermans A., Beyer L., Leibe B. In Defense of the Triplet Loss for Person Re-Identification. 2017. arXiv preprint arXiv:1703.07737.
28. Li D., Chen X., Zhang Z., Huang K. Learning Deep Context-Aware Features over Body and Latent Parts for Person Re-identification. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. IEEE, 2017. pp. 7398–7407. DOI: 10.1109/CVPR.2017.782.
29. Luo H., Wang P., Xu Y., Ding F., Zhou Y., Wang F., Li H., Jin R. Self-Supervised Pre-Training for Transformer-Based Person Re-Identification. 2021. arXiv preprint arXiv:2111.12084.

**Идрисси Алами Моссааб** — аспирант, факультет наук, Рабатский Университет Мохаммеда V. Область научных интересов: инженерия больших данных, компьютерная инженерия, администрирование информационных систем. [mossaab\\_idrissialami@um5.ac.ma](mailto:mossaab_idrissialami@um5.ac.ma); Проспект Организации Объединенных Наций, Агдаль, BP:8007.NU, Рабат, Марокко; р.т.: +(212)661-065-800.

**Эз-захут Абдеррахман** — профессор, департамент компьютерных наук/факультет наук, Рабатский Университет Мохаммеда V. Область научных интересов: информатика, цифровые системы, большие данные, компьютерное зрение, интеллектуальные системы.

Число научных публикаций — 30. a.ezzahout@um5r.ac.ma; Проспект Организации Объединенных Наций, Агдаль, ВР:8007.NU, Рабат, Марокко; р.т.: +(212)669-455-252.

**Омари Фузия** — руководитель исследовательской группы ipss, профессор, департамент компьютерных наук/ens, Рабатский Университет Мохаммеда V. Область научных интересов: компьютерные науки, кибербезопасность, блокчейн, криптовалюта. Число научных публикаций — 18. f.omary@um5r.ac.ma; Проспект Организации Объединенных Наций, Агдаль, ВР:8007.NU, Рабат, Марокко; р.т.: +(212)661-391-420.

И.С. КИПЯТКОВА, И.А. КАГИРОВ, М.Д. ДОЛГУШИН  
**ПРИМЕНЕНИЕ ПРЕДВАРИТЕЛЬНО ОБУЧЕННЫХ  
МНОГОЯЗЫЧНЫХ МОДЕЛЕЙ ДЛЯ РАСПОЗНАВАНИЯ  
КАРЕЛЬСКОЙ РЕЧИ**

*Кипяткова И.С., Кагиров И.А., Долгушин М.Д. Применение предварительно обученных многоязычных моделей для распознавания карельской речи.*

**Аннотация.** В настоящей статье описывается экспериментальное исследование, направленное на решение проблемы обучения моделей для распознавания речи в условиях малого объема обучающих речевых и текстовых данных. Подробно рассматриваются существующие подходы к решению данной проблемы, в частности, использование преобученных многоязычных моделей и аугментация данных. В работе проведена адаптация многоязычных моделей на базе Wav2Vec и Whisper к ливвиковскому наречию карельского языка и проведено исследование применения внешней языковой модели для повышения точности распознавания интегральной системы. Кроме того, в статье описаны специально собранная и подготовленная речевая база данных и базовая система распознавания, созданная на основе тулкита Kaldi. Приведены количественные результаты тестирования, которые подтверждают эффективность выбранных методов: так, использование моделей на архитектуре Трансформер, в частности, Wav2Vec, позволило достичь более высоких показателей, чем у базовых моделей, обученных с помощью программных средств Kaldi. Дообучение моделей Wav2Vec снизило количество неправильно распознанных слов до 24,73% на валидационной и до 25,25% на тестовой выборках, а использование модели Wav2Vec-BERT 2.0 с внешней языковой моделью дополнительно уменьшило количество неправильно распознанных слов до 17,12% и 17,72% соответственно. Статья адресована, в первую очередь, специалистам, занимающимся разработкой систем автоматического распознавания речи для малоресурсных языков и распознаванием речи на прибалтийско-финских языках, в частности, результаты этой работы могут найти практическое применение в полевых исследованиях, при записи текстов на карельском.

**Ключевые слова:** малоресурсные языки, карельский язык, переключение кодов, преобученные модели, машинное обучение, речевой корпус.

**1. Введение.** В последние годы наблюдается рост интереса к исследованиям в области автоматического распознавания речи для малоресурсных языков. Это связано, среди прочего, с актуальностью проблемы исчезновения миноритарных языков коренных народов. Существенным препятствием, с которым сталкиваются разработчики подобных систем, является дефицит данных для обучения акустических и языковых моделей.

За последние годы технологии автоматического распознавания речи претерпели значительные изменения, перейдя от классических модульных систем к интегральным. В традиционных системах выделение признаков, акустическое и языковое моделирование, а также декодирование выполняются отдельными модулями. Интегральный (или end-to-end) подход, наоборот, объединяет все эти

модули в одну глубокую нейронную сеть, которая обучается как единое целое. Интегральные системы превосходят традиционные по точности и скорости распознавания, однако одним из ключевых требований для обучения интегральной системы является доступность больших объемов данных. Малоресурсные языки по определению не обладают оцифрованными данными большого объема, что требует разработки особых подходов для применения современных методов машинного обучения к их материалу.

Дополнительной сложностью при создании систем распознавания речи для малоресурсных языков является феномен переключения кодов (англ. code-switching) – перехода с одного языка на другой в процессе речи. Поскольку малоресурсные языки часто существуют в условиях полиязычной среды, переключение кодов особенно характерно именно для них. Это явление дополнительно обостряет проблему недостатка обучающих данных, поскольку переключение кодов снижает однородность и согласованность речевых корпусов.

В настоящей статье представлено продолжение серии экспериментов по автоматическому распознаванию речи на малоресурсном карельском языке [1–4]. На примере карельского языка описаны этапы создания подобной системы, включая подготовку речевых данных, обучение акустических и языковых моделей. Приведены количественные результаты тестирования разработанных моделей, подтверждающие эффективность применённых методов.

**2. Основные методы решения проблемы нехватки данных при обучении систем распознавания речи.** Как следует из предыдущего раздела, «прямое» применение интегрального подхода к материалу малоресурсных языков оказывается проблематичным, и среди важнейших задач, которые должны быть решены до этапа обучения системы, оказывается создание набора обучающих данных достаточного объема.

Одним из способов расширения объема обучающих данных является аугментация. Среди основных методов аугментации аудиоданных можно перечислить [5]:

- изменение высоты голоса, темпа речи, громкости речи;
- модификация речевых признаков (например, добавление случайных значений к речевым признакам);
- изменение спектрограммы;
- преобразование голоса (англ. voice conversion);
- синтез речи.

Аугментация текстовых данных может выполняться следующими способами:

- использование данных другой предметной области;
- использование машинного перевода;
- модификация текста путем случайной замены/вставки/удаления слов или символов;
- генерация текста с помощью искусственных нейронных сетей (ИНС).

Другим подходом к решению проблемы нехватки данных, который показал свою эффективность при обучении нейронных сетей, является метод переноса знаний (англ. transfer learning). Суть этого метода состоит в предварительном обучении модели на большом объеме нецелевых данных с последующим дообучением модели на малом объеме целевых данных. Таким образом, основной принцип переноса знаний – это применение знаний нейросети, обученной на одной задаче, к другой задаче с предварительным дообучением.

В настоящее время существуют предварительно обученные многоязычные модели для распознавания речи, которые находятся в открытом доступе. Одной из таких моделей является модель Wav2Vec 2.0 [6], использующая метод самообучения, подразумевающий начальное обучение речевых представлений на большом объеме неразмеченных данных, а затем модель дообучается на меньшем объеме размеченных данных. Модель Wav2Vec 2.0 состоит из трех блоков: блок извлечения признаков, в качестве которого используется многослойная сверточная нейронная сеть, трансформерный кодер и модуль квантования, который принимает на вход все различные представления речевого сигнала, сгенерированные блоком извлечения признаков, и сводит их к конечному набору речевых единиц.

Существуют различные версии модели, например, Base и Large, различающиеся количеством обучаемых параметров, а также многоязычная вариация модели, использующая кросс-языковые представления, XLS-R (англ. cross-lingual speech representations), которая может быть использована при создании моноязычных моделей [7, 8]. В работе [9] эта модель, под названием MMS (Massively Multilingual Speech), была адаптирована для материала более чем 1000 языков, одним из которых был карельский. Авторам упомянутой работы удалось добиться хорошего уровня распознавания за счет эффективного использования слоев Адаптеров<sup>1</sup> и постепенного добавления языков на этапе обучения [10]. Кроме того, как было

---

<sup>1</sup> [https://huggingface.co/docs/peft/conceptual\\_guides/adapter](https://huggingface.co/docs/peft/conceptual_guides/adapter)

показано в работе [11], объединение языковых моделей и лингвистических кодеров, например, BERT [12], с акустическим кодером может быть перспективным в задачах распознавания речи с ограниченными ресурсами. Некоторые из этих моделей, такие как WavLM<sup>1</sup>, использующая этап кластеризации для получения выравненных BERT-подобных потерь, а также шумоподавление и управляемое смещение относительной позиции кодов, и W2V2-BERT V2, развивающая архитектуру Wav2Vec за счет замены механизма внимания на модель Conformer с применением каузального глубинного слоя, могут демонстрировать результаты, соответствующие современному техническому уровню [13, 14] распознавания речи.

Другой широко распространенной предобученной моделью является модель Whisper от компании OpenAI, обученная на 680 тыс. часах размеченных многоязычных данных. Whisper основана на архитектуре Трансформер, которая состоит из кодера и декодера. Кодер состоит из двух сверточных слоев, за которыми следует синусоидальное позиционное кодирование и блоки Transformer. Декодер использует обученные векторы позиционного кодирования и содержит то же количество блоков Трансформер, что и кодер. Архитектура Whisper подробно описана в [15].

Существует несколько реализаций модели Whisper – Tiny, Base, Small, Medium, Large, – которые отличаются количеством используемых параметров. Так же, как и Wav2Vec, модель Whisper может быть дообучена на целевых данных, однако, как было показано в нескольких исследованиях, дообучение модели Whisper показывает результаты распознавания хуже, нежели Wav2Vec. Например, в работе [16] было проведено исследование применения моделей Wav2Vec 2.0 и Whisper для распознавания мальтийской речи с переключением кода на английский. Авторы выполнили дообучение моделей Wav2Vec XLS-R с 300 млн и с 3 млрд параметров, в также были проведены эксперименты с моделями Whisper Tiny, Small и Large на речевых данных различного объема от 10 минут до 100 часов. Наилучшие результаты показала модель XLS-R 2B с 2 млрд параметров при дообучении на 50 часах речевых данных, при этом значение WER (количество неправильно распознанных слов, англ. word error rate) составило 8,53%, значение CER (количество неправильно распознанных символов, англ. character error rate) – 1,93% на тестовой части корпуса CommonVoice и 24,98% – WER, 8,37% – CER на корпусе MASRI. Сравнимые результаты были получены при

<sup>1</sup> [https://huggingface.co/docs/transformers/model\\_doc/wavlm](https://huggingface.co/docs/transformers/model_doc/wavlm)

дообучении модели на данных объемом примерно 10 часов. Модель Whisper показала существенно худшие результаты, и при использовании модели Whisper Tiny значение WER во всех экспериментах составило 100%.

Другое исследование, посвященное сравнению моделей Wav2Vec 2.0 и Whisper, представлено в статье [17]. Суть этой работы состояла в дообучении моделей Wav2Vec 2.0 Base и XLSR-53 вместе с моделями Whisper Small и Large на корпусе казахской речи. Модели Wav2Vec 2.0 превосходили модели Whisper по точности распознавания. Наилучшие результаты были получены для модели Wav2Vec 2.0 Base (WER=9,8%, CER=2,7%), при этом с применением многоязычной модели Whisper Large значение WER составило 19,8%, CER – 4,1%.

Еще одной проблемой, которая часто возникает при обработке малоресурсных языков, является проблема переключения кодов. Существуют два основных подхода к решению данной проблемы. Первый предполагает определение границ разноязычных фрагментов речи и их обработку моноязычной системой. Для определения языка используют акустические признаки (например, *i*-вектора или bottleneck-признаки [18]), лексические данные (теги частей речи [19], триггерные слова [20]) или их комбинацию [21]. Второй подход использует многоязычные системы распознавания речи, требующие унификации алфавитов и фонем [22].

Стоит отметить, что зачастую особые сложности вызывает нехватка текстовых данных для языкового моделирования, так как переключение кодов реже встречается в письменной речи [23, 24]. Для аугментации данных применяют частичный автоматический перевод [25] и генерацию текста с помощью ИНС [26]. Среди других методов можно перечислить использование класс-ориентированных моделей [27], факторных моделей языка с тегами переключения [28] и двуязычных векторных представлений слов, основанных на параллельных корпусах [29].

В последнее время все большее распространение приобретает применение предварительно обученных многоязычных моделей, таких как mBERT (обученная на текстах из Википедии на 104 языках) [12] или XLM RoBERTa (обученная на 2,5 ТБ отфильтрованных данных CommonCraw на 100 языках) [30]. Так, в работе [31] исследуется применение больших предобученных моделей (GPT-2 и BERT) для распознавания речи для ряда африканских языков в условиях переключения кодов с английским и делается вывод о том, что применение подобных моделей эффективно позволяет снизить WER. Также в данной работе делается вывод о перспективности методики

формирования обучающего корпуса путем чередования текстовых данных из моноязычных корпусов разных языков.

Еще одним способом повышения точности распознавания речи является использование внешней языковой модели [32, 33], которая улучшает синтаксическую и семантическую интерпретацию текста за счет более точного вероятностного представления целевого языка, при этом могут использоваться как статистические, так и нейросетевые модели. Несмотря на то, что такая модель представляет собой отдельный модуль, чаще всего подобные системы называют интегральными системами с внешней моделью языка [34]. В последующих разделах будут продемонстрированы практические решения задачи распознавания речи в рамках интегрального подхода на материале ливвиковского наречия малоресурсного карельского языка.

**3. Методика исследования.** В рамках данного исследования был проведен сравнительный анализ применения предобученных моделей на основе Wav2Vec и Whisper с базовой системой на основе Kaldi [35] для задачи распознавания карельской речи. Для обучения и тестирования моделей использовалась база данных AnKaS<sup>1</sup> – «База данных аннотаций речевых записей на карельском языке (AnKaS – Database of Annotations of Karelian Speech Recordings)» (далее – БД AnKaS), описание которой приведено ниже в разделе 4.

В качестве базовой системы использовалась многомодульная система распознавания карельской речи на основе Kaldi, включающая гибридную акустическую модель, объединяющую ИНС и скрытые марковские модели (СММ) – ИНС/СММ, а также модели языка двух типов: статистическую (используемую на этапе декодирования) и нейросетевую (применяемую на этапе переоценки списка гипотез распознавания). Более подробно базовая система распознавания описана в разделе 5.

Выбраны предварительно обученные интегральные модели на основе Wav2Vec и Whisper, и выполнено их дообучение. Обоснование выбора предобученных моделей и процесс их дообучения представлены в разделе 6.

Проведены эксперименты по распознаванию карельской речи с использованием многомодульной и интегральных моделей, а также эксперименты с применением внешней языковой модели дополнительно к интегральной. Оценка качества распознавания речи осуществлялась по показателю WER, описание экспериментов и полученных результатов приведено в разделе 7.

---

<sup>1</sup> <https://github.com/IrinaKipyatkova/AnKaS>



**4. База данных аннотаций речевых записей на карельском языке.** БД AnKaS представляет собой набор аннотаций записей карельской речи (ливвиковское наречие) из 13 радиопередач программы «Kodirandaine» («Родной берег»). Оригинальные аудиофайлы находятся в открытом доступе на сайте ГТРК «Карелия»<sup>1</sup>. Была выполнена расшифровка аудиозаписей. В БД AnKaS были включены аннотации только таких речевых отрезков, в которых отсутствует фоновый шум, речевые сбои и одновременная речь нескольких дикторов. Всего были созданы аннотации для 4385 фраз (4,5 часа речи). Основные характеристики БД AnKaS представлены в таблице 1.

Таблица 1. Характеристики БД AnKaS

Параметр	Значение
Общее количество дикторов	17 (7 мужчин, 10 женщин)
Длительность аннотированных речевых данных	4,5 часа
Количество фраз	4385
Количество словоупотреблений	32037
Количество уникальных слов	9117

Логическая структура БД AnKaS представлена на рисунке 1. БД представлена в формате JSON. Аннотация речевых записей каждого диктора содержится в отдельном .json файле, при этом используются следующие ключи:

- “phrase\_id” – номер фразы у данного диктора;
- “link” – интернет-ссылка на аудиозапись на сайте ГТРК «Карелия»;
- “time\_start” – время начала фразы;
- “time\_end” – время окончания фразы;
- “sentence” – текстовая расшифровка;
- “sentence\_rus” – текстовая расшифровка с метками переключения языка на русский, при этом русскоязычный текст заключается в треугольные скобки и помечается тегом «rus».

<sup>1</sup> <https://tv-karelia.ru/kodirandaine-rodnoy-bereg/>

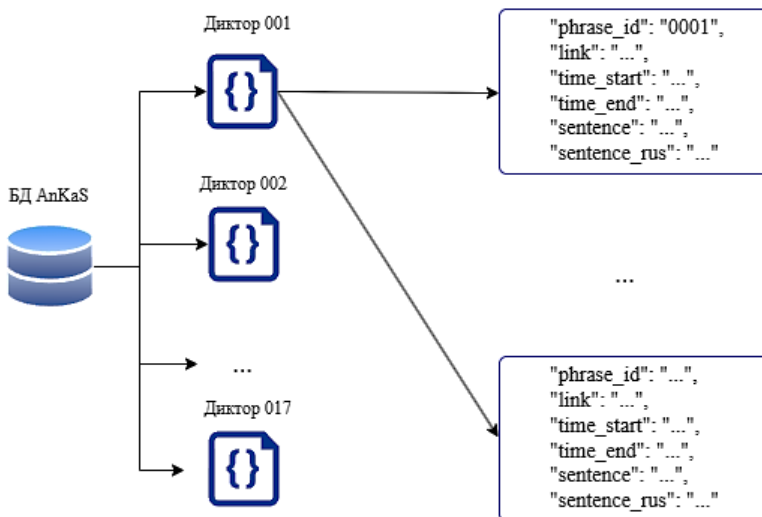


Рис. 1. Логическая структура БД AnKaS

Пример аннотации для одной фразы показан на рисунке 2.

```
{
  "phrase_id": "0002",
  "link": "https://tv-karelia.ru/wp-content/uploads/2022/01/Kirvesmies-Aleksandr-Ivanov.mp3",
  "time_start": "114.7394720",
  "time_end": "120.5044720",
  "sentence": "vot d'ad'a miša jakovlev tās susiedu d'ad'a pet'a jakovlev tās test'u opasti",
  "sentence_rus": "<vot d'ad'a miša jakovlev>rus tās susiedu <d'ad'a pet'a jakovlev>rus tās test'u opasti"
},
```

Рис. 2. Пример аннотации одной фразы

Особо стоит отметить, что в аннотации содержится в том числе и информация о переключении кода с карельского на русский. Случаи переключения кодов, обнаруженные в собранных данных, немногочисленны со статистической точки зрения (только 1,13% лексических единиц были помечены тегами переключения кодов). Однако, принимая во внимание тот факт, что единственным источником языковых данных были радиопередачи (то есть в записях присутствуют не спонтанные нарративы, а подготовленная речь, звучащая в формальной обстановке), это число всё равно значимо.

Собранный корпус использовался для обучения и тестирования системы распознавания карельской речи, при этом данные были разделены на обучающую, валидационную и тестовую выборки. 80% высказываний были использованы для обучения акустической модели, 10% – для валидации и настройки гиперпараметров, а 10% – для финального тестирования. Аугментация обучающей части корпуса выполнялась путем совместного и поочередного изменения частоты основного тона (ЧОТ) и темпа речи. Модификация ЧОТ выполнялась на количество полутонов, полученных случайным образом из равномерного распределения в диапазоне  $[-2, 2]$ , а изменение темпа речи осуществлялось с помощью коэффициента, случайно выбранного из равномерного распределения в диапазоне  $[0,7, 1,3]$ . Для выполнения описанных модификаций речевого сигнала использовался инструментарий SoX.<sup>1</sup> В результате размер обучающей выборки был увеличен до 13,5 часов. Более подробно процесс подготовки речевого материала описан в работе [36].

### 5. Система распознавания карельской речи на основе Kaldi.

Базовая система распознавания карельской речи была создана с помощью тулкита Kaldi [35]. Для акустического моделирования использовалась гибридная СММ/ИНС модель [2] на основе факторизованной ИНС с временными задержками (англ. TDNN-F), то есть ИНС с временными задержками (TDNN), в которой размерность слоев сокращена путем сингулярного разложения [37]. Архитектура ИНС показана на рисунке 3.

Архитектура сети состоит из трех блоков TDNN-F. Первый блок имеет три слоя TDNN-F, обрабатывающих входные вектора во временном интервале  $\{-1,0,1\}$ . Второй блок содержит один слой TDNN-F без объединения с временными шагами (как предыдущим, так и последующим). Третий блок состоит из 10 слоев TDNN-F, обрабатывающих временные шаги  $\{-3,0,3\}$ . Размер слоя TDNN-F равен 1024. В качестве активационной функции использовалась функция ReLU. В модели применялся метод пропуска соединений, аналогичный представленному в [37], при этом выходные данные каждого слоя (кроме первого слоя) добавляются к выходным данным предыдущих слоев с коэффициентом 0,66. За слоями TDNN-F располагается линейный слой размером 256. Обучение осуществлялось в течение 8 эпох, при этом коэффициент скорости обучения уменьшался со значения 0,0005 до 0,00005. Число эпох обучения, коэффициент скорости обучения и другие гиперпараметры моделей были выбраны эмпирически. Входными данными для

<sup>1</sup> <http://sox.sourceforge.net/sox.html>

нейронной сети были мел-частотные кепстральные коэффициенты (англ. MFCC), при этом для адаптации к речи диктора к ним был добавлен 100-мерный i-вектор [38].

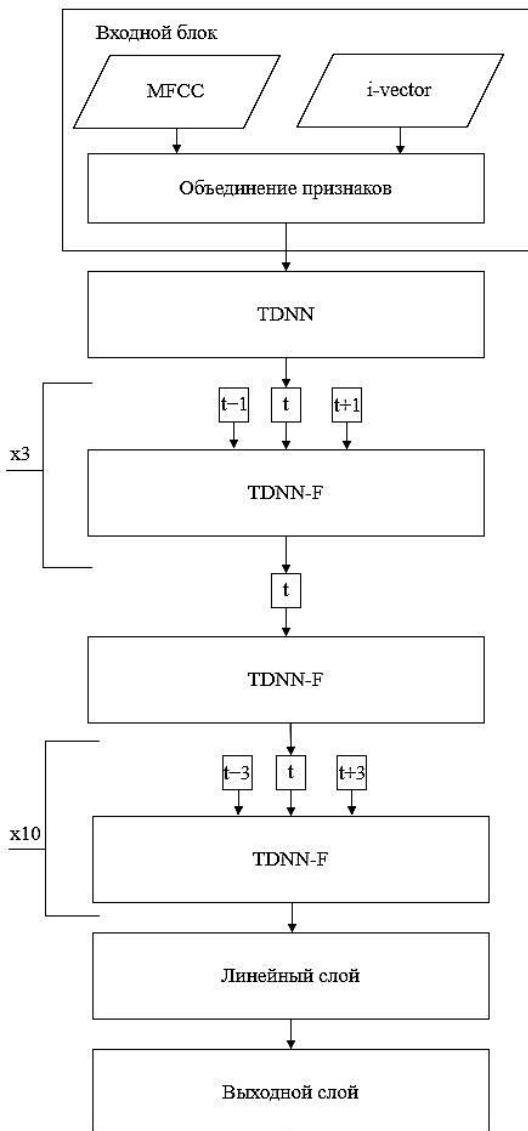


Рис. 3. Архитектура сети TNDNN-F

Для языкового моделирования были созданы два типа моделей: статистическая, которая использовалась на этапе декодирования, и нейросетевая, которая применялась для переоценки списка гипотез распознавания (N-best list). Статистическая модель представляла собой модель на основе триграмм слов, и она была обучена с помощью инструментария SRI Language Modeling Toolkit (SRILM) [39].

Для обучения нейросетевой модели был использован инструментарий TheanoLM [40]. Нейросетевая модель состояла из проекционного слоя размером 500 и двух слоёв LSTM размером 512. В качестве метода оптимизации использовался метод моментов Нестерова. Размер батча был равен 16. Была выполнена линейная интерполяция нейросетевой модели с триграммной, при этом коэффициент интерполяции для нейросетевой модели был равен 0,6. Для обучения языковых моделей использовались текстовые данные, полученные из открытого корпуса вепского и карельского языков «VerKar»<sup>1</sup> и публикаций на карельском языке, а также из расшифровок обучающей части речевого корпуса. Подробно процесс обработки текстовых данных и создания моделей языка описан в [2, 36].

Текстовый корпус использовался для формирования словаря системы, в него вошли все слова из расшифровок обучающей части корпуса и слова из других текстовых материалов, которые встречались в нём не менее двух раз. Объём словаря системы составил 143907 слов.

Фонематические транскрипции для словаря создавались автоматически. В карельском языке фиксированное ударение, которое всегда падает на первый слог, а гласные не подвержены редукции. Важной особенностью карельского языка является наличие смыслоразличительной долготы как для гласных, так и для согласных. В ходе предыдущего исследования [3] было определено, что наилучшие результаты распознавания речи даёт использование такого фонемного инвентаря, в котором долгие гласные являются отдельными фонемами, а для слов с долгими согласными было создано по две альтернативные транскрипции: с интерпретацией долгих согласных как удвоенных кратких и без удвоения фонем. Такое решение необходимо для учета возможной редукции долгого согласного в разговорной речи. В результате процесс автоматического создания транскрипции включает в себя определение ударения, обработку удвоенных фонем и определение палатализованных согласных (перед гласными переднего ряда). Из-за присутствия в расшифровках речевого корпуса русскоязычных слов, словарь

---

<sup>1</sup> <http://dictorpus.krc.karelia.ru/ru>

транскрипций включал в себя и русские слова, транскрипции для которых также создавались автоматически. Более подробно проблемы русской фонемной транскрипции обсуждаются в предыдущих работах [41].

**6. Выбор предобученной модели для создания системы распознавания карельской речи.** Проведен анализ готовых предварительно обученных моделей для распознавания карельской речи, в результате которого было принято решение выбрать следующие интегральные модели, основанные на архитектурах Wav2Vec и Whisper:

- Wav2Vec2.0 Large-Uralic-VoxPopuliV2;
- WavLM Large;
- MMS 1B All;
- W2V2-BERT 2.0;
- Whisper Small;
- Whisper Medium;
- Whisper Large V2 Distilled.

Данные модели были выбраны по той причине, что они показывают одни из лучших результатов при многоязычном распознавании речи и часто изучаются в контексте применения к малоресурсным языкам [42, 43]. Модель MMS 1B All [9] была рассмотрена, в том числе, еще и потому, что ее авторами заявляется поддержка распознавания как собственно-карельского, так и ливвиковского наречий карельского языка.

В таблице 2 представлено краткое описание этих моделей и параметров, использовавшихся при дополнительном обучении. Дообучение моделей проводилось с использованием фреймворка Transformers [44].

Возможность использования оригинальных моделей Whisper Large v2, Whisper Large v3 и Whisper Large v3 Turbo, показывающих одни из лучших результатов распознавания для английского и русского, не была изучена в рамках настоящей работы по причине высоких требований к вычислительным ресурсам, налагаемых перечисленными моделями.

Важно отметить, что дообучение всех моделей происходило в течение 10000 шагов. Данное число было определено эмпирически. При большем количестве шагов снижение WER на валидационной выборке и функции потерь на обучающей выборке переставало происходить, поэтому данное количество шагов при обучении всех моделей оказалось оптимальным.

Размеры батчей и число шагов аккумуляции градиента подбирались индивидуально в силу разных требований к вычислительным ресурсам, необходимым для дополнительного обучения данных моделей. При возможности предпочтение отдавалось большему размеру батча, но авторы не исключают возможности дополнительного улучшения результатов при увеличении размеров батчей.

Таблица 2. Описание моделей на базе Wav2Vec и Whisper

Название	Кол-во параметров	Кол-во данных (часы), разметка	Число языков*	ЯМ**	Параметры дообучения
Wav2Vec2.0 Large Uralic VoxPopuli V2 <sup>1</sup>	300 млн	42,5 тыс., неразмеченные	3	нет	10 тыс. шагов обучения, размер батча – 8, шагов накопления градиента – 4
WavLM Large <sup>2</sup>	316 млн	94 тыс., неразмеченные	1***	нет	10 тыс. шагов обучения, размер батча – 4, шагов накопления градиента – 8
MMS 1B All <sup>3</sup>	1000 млн	49 тыс., размеченные	1162	нет	10 тыс. шагов обучения, размер батча – 8, 4 шагов накопления градиента
W2V2-BERT 2.0 <sup>4</sup>	600 млн	4,5 млн, неразмеченные	143	нет	10 тыс. шагов обучения, размер батча – 2, шагов накопления градиента – 16
Whisper Small <sup>5</sup>	244 млн	680 тыс., размеченные	99	да	10 тыс. шагов обучения, размер батча – 16, шагов накопления градиента – 2
Whisper Medium <sup>6</sup>	769 млн	680 тыс., размеченные	99	да	10 тыс. шагов обучения, размер батча – 16, шагов накопления градиента – 2
Whisper Large V2 Distilled <sup>7</sup>	756 млн	22 тыс. размеченные	1***	да	10 тыс. шагов обучения, размер батча – 16, шагов накопления градиента – 2

\* Все модели, кроме WavLM Large и Whisper Large V2 Distilled, были обучены в том числе и на уральских языках. Wav2Vec2.0 Large Uralic VoxPopuli V2 обучен исключительно на уральских языках. \*\* Языковая модель. \*\*\* английский.

В моделях на основе Wav2Vec и Whisper при токенизации не учитывалась длительность звуков, в отличие от системы, созданной на базе Kaldi. В ходе предварительных экспериментов исследовалась интерпретация долгих звуков как отдельных фонем, однако она не оказала существенного влияния на результаты. Кроме того,

<sup>1</sup> <https://huggingface.co/facebook/wav2vec2-large-uralic-voxpathuli-v2>

<sup>2</sup> <https://huggingface.co/microsoft/wavlm-large>

<sup>3</sup> <https://huggingface.co/facebook/mms-1b-all#model-details>

<sup>4</sup> [https://huggingface.co/docs/transformers/model\\_doc/wav2vec2-bert](https://huggingface.co/docs/transformers/model_doc/wav2vec2-bert)

<sup>5</sup> <https://huggingface.co/openai/whisper-small>

<sup>6</sup> <https://huggingface.co/openai/whisper-medium>

<sup>7</sup> <https://huggingface.co/distil-whisper/distil-large-v2>

в работе [43], описывающей интегральную систему для финского, также не учитывалась длительность гласных, хотя в финском языке долгота является смысловоразличительным признаком фонем.

Поскольку модель Whisper не обладает встроенной поддержкой ливвиковского наречия карельского языка, было принято решение дообучить модель для финского на базе Whisper. Возможность добавления нового языка к уже имеющимся в Whisper существует, однако она сопряжена с необходимостью полного переобучения токенизатора, основанного на модели GPT-2, что является неоптимальным решением, связанным с повышением ресурсозатратности. Впрочем, подобный подход был успешно применен в работе [43], где веса для финского языка дообучались на данных по северному диалекту саамского.

Как было замечено выше (раздел 2), добавление языковой модели к предобученной может значительно повысить точность распознавания, в особенности для материала малоресурсных языков [33]. Поэтому были проведены предварительные исследования по интеграции разработанной триграммной языковой модели к некоторым из предобученных моделей, в частности, к Wav2Vec2.0 Large Uralic VoxPopuli V2, MMS-1B All и W2V2 BERT 2.0. Результаты этих экспериментов представлены в следующем разделе.

### **7. Результаты экспериментов по распознаванию речи.**

Результаты экспериментов по распознаванию речи на карельском языке с использованием различных типов акустических моделей, полученные на валидационной (Dev) и тестовой (Test) частях корпуса, представлены в таблице 3. В качестве метрики был принят WER.

Вначале были проведены эксперименты по распознаванию речи с использованием модели TDNN-F/CMM, обученной на речевых данных без аугментации. Значение WER, полученное при использовании модели TDNN-F для распознавания валидационной части корпуса, оказалось равно 28,96%, а для тестовой части 30,58%. Использование модели, обученной на аугментированных данных, позволило снизить значение WER до 27,13% на валидационной части корпуса и до 28,77% – на тестовой, таким образом, относительное снижение WER составило 6%. В то же время, выполнение переоценки списка лучших 500 гипотез распознавания с помощью нейросетевой модели языка, интерполированной с триграммной, позволило снизить WER до 25,44% и 27,20% на валидационном и тестовом корпусах соответственно.



Таблица 3. Результаты экспериментов по распознаванию карельской речи

Акустическая модель	Языковая модель	Значение WER, %	
		Корпус Dev	Корпус Test
TDNN-F/CMM (без аугментации)	Триграммная	28,96	30,58
TDNN-F/CMM	Триграммная	27,13	28,77
TDNN-F/CMM	Триграммная + нейросетевая	25,44	27,20
Wav2Vec2.0 Large Uralic VoxPopuli V2	–	24,73	25,25
Wav2Vec2.0 Large Uralic VoxPopuli V2	Триграммная	19,69	19,83
WavLM Large	–	34,74	38,34
WavLM Large	Триграммная	26,72	29,04
MMS 1B All	–	31,56	32,06
MMS 1B All	Триграммная	24,29	24,99
W2V2-BERT 2.0	–	18,84	20,39
W2V2-BERT 2.0	Триграммная	<b>17,39</b>	<b>17,86</b>
Whisper Small	–	32,22	35,25
Whisper Medium	–	25,54	28,54
Whisper Large V2 Distilled	–	28,38	30,75

Использование дообученной модели WavLM Large с внешней языковой показало результаты, сравнимые с результатами, полученными с применением моделей TDNN-F/CMM, при этом значения WER составили 26,72% и 29,04% на валидационном и тестовом корпусах соответственно. Однако без языковой модели WavLM Large показала наихудшие результаты. Возможно такое отставание в результатах от других моделей может быть объяснено тем, что модель была предобучена только на английском языке и заведомо не предполагалась быть многоязычной.

Дообучение моделей MMS также показало результаты, сравнимые с базовой моделью, несмотря на то что модели MMS уже имели веса для карельского языка. Было получено значение WER, равное 31,56% на валидационном корпусе и 32,06% на тестовом. Применение триграммной модели значительно улучшило результаты, что привело к достижению 24,29% WER на валидационных данных и 24,99% на тестовых.

Сопоставимые результаты показала модель Wav2Vec, предобученная на уральских языках. Использование данной модели позволило снизить WER до 24,73% на валидационном корпусе и до 25,25% на тестовом. Применение внешней языковой модели привело

к дальнейшему снижению WER до 19,69% на валидационном корпусе и до 19,83% на тестовом.

Наилучшие результаты были получены с помощью модели W2V2-BERT 2.0. Без использования триграммной модели значение WER составило 18,84% на валидационном корпусе и 20,39% – на тестовом. Добавление внешней языковой модели позволило дополнительно улучшить результаты. При этом значение WER составило 17,39% на валидационном корпусе и 17,86% – на тестовом.

Модели на базе Whisper показали результаты хуже, чем Wav2Vec-модели. При этом наилучшие результаты были получены с моделью Whisper Medium; значение WER составило 25,54% и 28,54% на валидационных и тестовых данных соответственно. Возможно, что лучшие результаты могли быть получены с моделью Whisper Large v3, однако высокие требования к вычислительным ресурсам для дообучения этой модели не позволили проверить это предположение. Также авторы предполагают, что сравнительно высокая ошибка при использовании моделей на основе Whisper может быть связана с разницей в длительности записей, использовавшихся при обучении и дообучении. Аудиозаписи в использованном наборе не превышают по длительности 10 секунд, тогда как Whisper обучался на записях длительностью в 30 секунд, из-за чего в целях нормирования заполняется нулями. В работе [45] отмечалось позитивное влияние на дообучение Whisper от использования более длинных и шумных записей, при большей устойчивости к шумам по сравнению с Wav2Vec. В контексте карельского языка данную гипотезу только предстоит изучить в будущем.

Далее на валидационной части была проведена настройка гиперпараметров языковой модели:

- $\alpha$  – весовой коэффициент, регулирующий значимость модели языка;
- $\beta$  – весовой коэффициент, регулирующий длину выходной последовательности; чем больше значение  $\beta$ , тем меньше длина выходной последовательности.

Значения WER, полученные с использованием модели Wav2Vec2-BERT 2.0, на валидационной части корпуса при различных параметрах  $\alpha$  и  $\beta$  языковой модели представлены на рисунке 4. На рисунке 4 более темным цветом выделена область с лучшими результирующими WER при соответствующих  $\alpha$  и  $\beta$ . Затем языковая модель с параметром  $\alpha$ , равном 0,8, и различных значениях  $\beta$ , продемонстрировавших наилучшие результаты на

валидационной части, была проверена на тестовой части набора, полученные значение представлены в таблице 4.

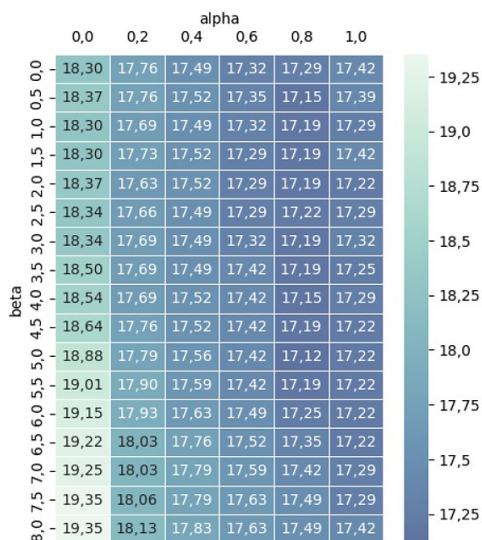


Рис. 4. Значения WER на валидационной части корпуса при различных параметрах alpha и beta языковой модели

Таблица 4. Результаты экспериментов при различных значениях параметра beta и значении alpha, равном 0,8

beta	Значение WER, %	
	Корпус Dev	Корпус Test
0,5	17,15	17,79
2,0	17,19	17,79
4,0	17,15	17,75
4,5	17,19	<b>17,72</b>
5,0	<b>17,12</b>	<b>17,72</b>

Наилучший результат на обеих выборках был получен при использовании alpha и beta равных 0,8 и 5,0 соответственно и составил 17,12 и 17,72 на валидационной и тестовой частях, тем самым позволив дополнительно улучшить результат, показанный моделью W2V2-BERT 2.0, на 0,14% на тестовой части. Интересно отметить, что низкое значение ошибки на тестовой части было получено также при beta 4,5 и 4,0, хотя это отличается от результатов, продемонстрированных на валидационной части.

Таким образом, наилучшие результаты были получены при использовании дообученной модели W2V2-BERT 2.0 и внешней языковой модели, что подтвердило эффективность данной архитектуры в задачах малоресурсного распознавания речи.

**8. Заключение.** В настоящей статье были рассмотрены основные методы, применяемые для решения проблем, связанных с созданием систем распознавания речи для малоресурсных языков в условиях недостаточного объема обучающих данных и переключения кодов. Приведено описание БД AnKAS, содержащей аннотации речевых данных на карельском языке, а также описаны модели распознавания речи для малоресурсного карельского языка, которые создавались с применением двух основных подходов традиционного (многомодульного) и интегрального с использованием предварительно обученных многоязычных моделей.

Результаты экспериментов показывают эффективность созданных моделей, полученные значения WER, равные 17,12% на валидационной выборке и 17,72% на тестовой выборке, соответствуют мировым результатам для малоресурсных языков. Однако стоит отметить, что несмотря на то, что для разговорной карельской речи характерно переключение кода с карельского на русский, в обучающих речевых данных таких явлений было немного. Именно поэтому коллективом авторов настоящей статьи ведется работа по подготовке корпуса с образцами разговорной карельской речи с переключением кодов. Дальнейшая работа будет посвящена обучению акустических и языковых моделей карельского языка с поддержкой переключения кода с карельского на русский.

### Литература

1. Кипяткова И.С., Кагиров И.А. Система автоматического распознавания карельской речи // Информационно-управляющие системы. 2023. № 3. С. 16–25.
2. Kipyatkova I., Kagirov I. Deep Models for Low-Resourced Speech Recognition: Livvi-Karelian Case // Mathematics. 2023. vol. 11. no. 18. DOI: 10.3390/math11183814.
3. Kipyatkova I., Kagirov I. Phone Durations Modeling for Livvi-Karelian ASR // Proceedings 25th International Conference Speech and Computer (SPECOM 2023). Springer LNCS. 2023. vol. 14339. pp. 87–99. DOI: 10.1007/978-3-031-48312-7\_7.
4. Kipyatkova I., Kagirov I., Dolgushin M., Rodionova A. Towards a Livvi-Karelian End-to-End ASR System // Proceedings 26th International Conference on Speech and Computer (SPECOM 2024). 2024. vol. 15299. pp. 57–68. DOI: 10.1007/978-3-031-77961-9\_4.
5. Кипяткова И.С., Кагиров И.А. Аналитический обзор методов решения проблемы малых наборов данных при создании систем автоматического распознавания речи для малоресурсных языков // Информатика и автоматизация. 2022. Т. 21. № 4. С. 678–709. DOI: 10.15622/ia.21.4.2.

6. Baevski A., Zhou Y., Mohamed A., Auli M. wav2vec 2.0: A Framework for Self-Supervised Learning of Speech Representations // *Advances in Neural Information Processing Systems*. 2020. vol. 33. pp. 12449–12460.
7. Conneau A., Baevski A., Collobert R., Mohamed A., Auli M. Unsupervised Cross-Lingual Representation Learning for Speech Recognition // *Proceedings of the Annual Conference of the International Speech Communication Association (Interspeech'2021)*. 2021. pp. 2426–2430. DOI: 10.21437/Interspeech.2021-329.
8. Babu A., Wang C., Tjandra A., Lakhota K., Xu Q., Goyal N., Singh K., Platen von P., Saraf Y., Pino J., Baevski A., Conneau A., Auli M. XLS-R: Self-Supervised Cross-Lingual Speech Representation Learning at Scale // *Proceedings of the Annual Conference of the International Speech Communication Association (Interspeech'2022)*. 2022. pp. 2278–2282.
9. Pratap V., Tjandra A., Shi B., Tomasello P., Babu A., Kundu S., Elkahky A., Ni Zh., Vyas A., Fazel-Zarandi M., Baevski A., Adi Y., Zhang X., Hsu W.-N., Conneau A., Auli M. Scaling Speech Technology to 1,000+ Languages // *Journal of Machine Learning Research*. 2024. vol. 25. pp. 1–52.
10. Poth C., Sterz H., Paul I., Purkayastha S., Engländer L., Imhof T., Vulić I., Ruder S., Gurevych I., Pfeiffer J. Adapters: A unified Library for Parameter-Efficient and Modular Transfer Learning // *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing: System Demonstrations (EMNLP'2023)*. 2023. pp. 149–160. DOI: 10.18653/v1/2023.emnlp-demo.13.
11. Chung Y.A., Zhang Y., Han W., Chiu C.-C., Qin J., Pang R., Wu Y. W2v-bert: Combining Contrastive Learning and Masked Language Modeling for Self-Supervised Speech Pre-Training // *Proceedings of 2021 IEEE Automatic Speech Recognition and Understanding Workshop (ASRU'2021)*. 2021. pp. 244–250. DOI: 10.1109/ASRU51503.2021.9688253.
12. Devlin J., Chang M.W., Lee K., Toutanova K. Bert: Pre-Training of Deep Bidirectional Transformers for Language Understanding // *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL-HLT'2019)*. 2019. vol. 1. pp. 4171–4186.
13. Chen S., Wang C., Chen Z., Wu Y., Liu S., Chen Z., Li J., Kanda N., Yoshioka T., Xiao X., Wu J., Zhou L., Ren S., Qian Y., Qian Y., Wu J., Zeng M., Yu X., Wei F. WavLM: Large-Scale Self-Supervised Pre-Training for Full Stack Speech Processing // *IEEE Journal of Selected Topics in Signal Processing*. 2022. vol. 16. no. 6. pp. 1505–1518.
14. Barrault L., Chung Y.A., Meglioli M.C., Dale D., Dong N., Duppenhaler M. et al. Seamless: Multilingual Expressive and Streaming Speech Translation // *arXiv preprint arXiv:2312.05187*. 2023.
15. Radford A., Kim J.W., Xu T., Brockman G., McLeavey C., Sutskever I. Robust Speech Recognition via Large-Scale Weak Supervision // *Proceedings of the 40th International Conference on Machine Learning 2022 (ICML'23)*. 2023. pp. 28492–28518.
16. Williams A., Demarco A., Borg C. The applicability of Wav2Vec 2.0 and Whisper for Low-Resource Maltese ASR // *Proceedings of the 2nd Annual Meeting of the ELRA/ISCA SIG on Under-resourced Languages (SIGUL'2023)*. 2023. pp. 39–43.
17. Kozhirbayev Z. Kazakh Speech Recognition: Wav2vec2.0 vs. Whisper // *Journal of Advances in Information Technology*. 2023. vol. 14. no. 6. pp. 1382–1389. DOI: 10.12720/jait.14.6.1382-1389.
18. Richardson F., Reynolds D., Dehak N. Deep Neural Network Approaches to Speaker and Language Recognition // *IEEE Signal Processing Letters*. 2015. vol. 22. no. 10. pp. 1671–1675. DOI: 10.1109/LSP.2015.2420092.

19. Winata G.I., Madotto A., Wu C.S., Fung P. Code-Switching Language Modeling using Syntax-Aware Multi-Task Learning // Proceedings of the Third Workshop on Computational Approaches to Linguistic Code-Switching (CALCS'2018). 2018. pp. 62–67. DOI: 10.18653/v1/W18-3207.
20. Adel H., Vu N.T., Kraus F., Schlippe T., Li H., Schultz T. Recurrent Neural Network Language Modeling for Code Switching Conversational Speech // Proceedings of 2013 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP'2013). 2013. pp. 8411–8415. DOI: 10.1109/ICASSP.2013.6639306.
21. Ramanarayanan V., Pugh R., Suenderman-Oeft D. Automatic Turn-Level Language Identification for Code-Switched Spanish-English Dialog // Proceedings of 9th International Workshop on Spoken Dialogue System Technology (IWSDS'2019). 2019. vol. 579. pp. 51–61. DOI: 10.1007/978-981-13-9443-0\_5.
22. Mustafa M.B., Yusoof M.A., Khalaf H.K., Abushariah A.A.R.M., Kiah M.L.M., Ting H.N., Muthaiyah S. Code-Switching in Automatic Speech Recognition: The Issues and Future Directions // Applied Sciences. 2022. vol. 12. no. 19. DOI: 10.3390/app12199541.
23. Çetinoğlu Ö., Schulz S., Vu N.T. Challenges of Computational Processing of Code-Switching // Proceedings of the Second Workshop on Computational Approaches to Linguistic Code Switching (CALCS'2016). 2016. pp. 1–11. DOI: 10.18653/v1/W16-5801.
24. Winata G., Aji A.F., Yong Z.X., Solorio T. The Decades Progress on Code-Switching Research in NLP: A Systematic Survey on Trends and Challenges // Findings of the Association for Computational Linguistics (ACL'2023). 2023. pp. 2936–2978. DOI: 10.18653/v1/2023.findings-acl.185.
25. Hsieh I.T., Wu C.H., Wang C.H. Acoustic and Textual Data Augmentation for Code-Switching Speech Recognition in Under-Resourced Language // IEEE Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC'2020). 2020. pp. 302–307.
26. Chang C.-T., Chuang S.-P., Lee H.-Y. Code-Switching Sentence Generation by Generative Adversarial Networks and its Application to Data Augmentation // Proceedings of the Annual Conference of the International Speech Communication Association (Interspeech'2019). 2019. pp. 554–558. DOI: 10.21437/Interspeech.2019-3214.
27. Chan J.Y.C., Cao H., Ching P.C., Lee T. Automatic recognition of Cantonese-English Code-Mixing Speech // International Journal of Computational Linguistics and Chinese Language Processing. 2009. vol. 14. no. 3. pp. 281–304.
28. Adel H., Vu N.T., Kirchhoff K., Telaar D., Schultz T. Syntactic and Semantic Features for Code-Switching Factored Language Models // IEEE/ACM Transactions on Audio, Speech, and Language Processing. 2015. vol. 23. no. 3. pp. 431–440. DOI: 10.1109/TASLP.2015.2389622.
29. Hermann K.M., Blunsom P. Multilingual Models for Compositional Distributed Semantics // Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers). 2014. pp. 58–68. DOI: 10.3115/v1/P14-1006.
30. Conneau A., Khandelwal K., Goyal N., Chaudhary V., Wenzek G., Guzmán F., Grave É., Ott M., Zettlemoyer L., Stoyanov V. Unsupervised Cross-Lingual Representation Learning at Scale // Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics (ACL'2020). pp. 8440–8451. DOI: 10.18653/v1/2020.acl-main.747.
31. Vüren van J., Niesler T. Improving N-best Rescoring in Under-Resourced Code-Switched Speech Recognition using Pretraining and Data Augmentation // Languages. 2022. vol. 7. no. 3. DOI: 10.3390/languages7030236.

32. Hono Y., Mitsuda K., Zhao T., Mitsui K., Wakatsuki T., Sawada K. Integrating Pre-Trained Speech and Language Models for End-to-End Speech Recognition // Findings of the Association for Computational Linguistics: ACL 2024. 2024. pp. 13289–13305. DOI: 10.18653/v1/2024.findings-acl.787.
33. Ogunremi T., Manning C.D., Jurafsky D. Multilingual Self-Supervised Speech Representations Improve the Speech Recognition of Low-Resource African Languages with Code Switching // arXiv preprint arXiv:2311.15077. 2023.
34. Hori T., Cho J., Watanabe S. End-to-end speech recognition with word-based RNN language models // Proceedings of 2018 IEEE Spoken Language Technology Workshop (SLT-2018). 2018. pp. 389–396. DOI: 10.1109/SLT.2018.8639693.
35. Povey D., Ghoshal A., Boulianne G., Burget L., Glembek O., Goel N., Hannemann M., Motlíček O., Qian Y., Schwarz P., Silovský J., Stemmer G., Veselý K. The Kaldi Speech Recognition Toolkit // Proceedings of IEEE Automatic Speech Recognition and Understanding Workshop (ASRU'2011). 2011. pp. 1–4.
36. Кипяткова И.С., Родионова А.П., Кагиров И.А., Крижановский А.А. Подготовка речевых и текстовых данных для создания системы автоматического распознавания карельской речи // Учёные записки Петрозаводского государственного университета. 2023. Т. 45. № 5. С. 89–98.
37. Povey D., Cheng G., Wang Y., Li K., Xu H., Yarmohammadi M., Khudanpur S. Semi-Orthogonal Low-Rank Matrix Factorization for Deep Neural Networks // Proceedings of The Annual Conference of the International Speech Communication Association (Interspeech'2018). 2018. pp. 3743–3747. DOI: 10.21437/Interspeech.2018-1417.
38. Saon G., Soltan H., Nahamoo D., Picheny M. Speaker Adaptation of Neural Network Acoustic Models using i-Vectors // Proceedings of IEEE Automatic Speech Recognition and Understanding Workshop (ASRU'2013). 2013. pp. 55–59. DOI: 10.1109/ASRU.2013.6707705.
39. Stolcke A., Zheng J., Wang W., Abrash V. SRILM at Sixteen: Update and Outlook // Proceedings of IEEE Automatic Speech Recognition and Understanding Workshop (ASRU'2011). 2011. pp. 5–9.
40. Enarvi S., Kurimo M. TheanoLM – An Extensible Toolkit for Neural Network Language Modeling // Proceedings of the Annual Conference of the International Speech Communication Association (Interspeech'2016). 2016. pp. 3052–3056. DOI: 10.21437/Interspeech.2016-618.
41. Karpov A., Markov K., Kipyatkova I., Vazhenina D., Ronzhin A. Large Vocabulary Russian Speech Recognition using Syntactico-Statistical Language Modeling // Speech Communication. 2014. vol. 56. pp. 213–228. DOI: 10.1016/j.specom.2013.07.004.
42. Wolf T., et al. Transformers: State-of-the-Art Natural Language Processing // arXiv preprint arXiv:1910.03771. 2019.
43. Grosz T., Getman Y., Al-Ghezi R., Rouhe A., Kurimo M. Investigating wav2vec2 Context Representations and the Effects of Fine-Tuning, a Case-Study of a Finnish Model // Proceedings of The Annual Conference of the International Speech Communication Association (Interspeech'2023). 2023. pp. 196–200. DOI: 10.21437/Interspeech.2023-837.
44. Hiovain-Asikainen K., Rosa de la J. Developing TTS and ASR for Lule and North Sámi Languages // Proceedings of the 2nd Annual Meeting of the ELRA/ISCA SIG on Under-resourced Languages (SIGUL'2023). 2023. pp. 48–52. DOI: 10.21437/SIGUL.2023-11.
45. Paonessa C., Timmel V., Vogel M., Perruchoud D. Whisper Fine-Tuning for Swiss German: A Data Perspective // Proceedings of the 9th edition of the Swiss Text Analytics Conference. 2024. pp. 192.

**Кипяткова Ирина Сергеевна** — канд. техн. наук, доцент, старший научный сотрудник, лаборатория речевых и многомодальных интерфейсов, Федеральное государственное бюджетное учреждение науки «Санкт-Петербургский Федеральный исследовательский центр Российской академии наук» (СПб ФИЦ РАН). Область научных интересов: автоматическое распознавание речи, нейронные сети. Число научных публикаций — 100. [kiryatkova@iiias.spb.su](mailto:kiryatkova@iiias.spb.su); 14-я линия В.О., 39, 199178, Санкт-Петербург, Россия; р.т.: +7(812)328-0421.

**Кагиров Ильдар Амирович** — научный сотрудник, лаборатория речевых и многомодальных интерфейсов, Федеральное государственное бюджетное учреждение науки «Санкт-Петербургский Федеральный исследовательский центр Российской академии наук» (СПб ФИЦ РАН). Область научных интересов: корпусная лингвистика, малоресурсные языки. Число научных публикаций — 40. [kagirov@iiias.spb.su](mailto:kagirov@iiias.spb.su); 14-я линия В.О., 39, 199178, Санкт-Петербург, Россия; р.т.: +7(812)328-0421.

**Долгушин Михаил Дмитриевич** — младший научный сотрудник, лаборатория речевых и многомодальных интерфейсов, Федеральное государственное бюджетное учреждение науки «Санкт-Петербургский Федеральный исследовательский центр Российской академии наук» (СПб ФИЦ РАН). Область научных интересов: автоматическое распознавание речи, нейронные сети, математическая лингвистика. Число научных публикаций — 10. [dolgushin.m@iiias.spb.su](mailto:dolgushin.m@iiias.spb.su); 14-я линия В.О., 39, 199178, Санкт-Петербург, Россия; р.т.: +7(812)328-0421.

**Поддержка исследований.** Работа выполнена при финансовой поддержке фонда РФФ (проект № 24-21-00276, <https://rscf.ru/project/24-21-00276/>).



I. KIPYATKOVA, I. KAGIROV, M. DOLGUSHIN  
**USE OF PRE-TRAINED MULTILINGUAL MODELS  
FOR KARELIAN SPEECH RECOGNITION**

*Kipyatkova I., Kagiroy I., Dolgushin M. Use of Pre-Trained Multilingual Models for Karelian Speech Recognition.*

**Abstract.** This paper presents an experimental study aimed at solving the problem of training speech recognition models under conditions of limited available speech and text data. Current approaches to this issue are discussed in detail, particularly the use of pre-trained multilingual models and data augmentation techniques. As part of this study, multilingual models based on Wav2Vec and Whisper were adapted to the Livvi dialect of the Karelian language, and an investigation into the use of an external language model to enhance recognition accuracy was conducted. The paper also describes a specially collected and prepared speech database and a basic recognition system developed using the Kaldi toolkit. Quantitative test results are provided as well, demonstrating the effectiveness of the chosen methods. For instance, Transformer-based models, particularly Wav2Vec, outperformed the baseline models trained using Kaldi software tools. Fine-tuning the Wav2Vec models reduced the word error rate to 24.73% on the validation set and 25.25% on the test set, while a combination of the Wav2Vec-BERT 2.0-based model with an external language model further reduced errors to 17.12% and 17.72%, respectively. This paper is primarily aimed at specialists in the field of automatic speech recognition for low-resource and Balto-Finnic languages. Additionally, the results of this work can be practically applied in field research involving Karelian text transcription. Future work includes expanding the database to improve model adaptation and enhance performance in real-world scenarios.

**Keywords:** low-resource languages, Karelian, code-switching, pre-trained models, machine learning, speech corpus.

## References

1. Kipyatkova I., Kagiroy I. [Automatic speech recognition system for Karelian]. *Informatsionno-upravliaiushchie sistemy – Information and Control Systems*. 2023. vol. 3. pp. 16–25. (In Russ.).
2. Kipyatkova I., Kagiroy I. Deep Models for Low-Resourced Speech Recognition: Livvi-Karelian Case. *Mathematics*. 2023. vol. 11. no. 18. DOI: 10.3390/math11183814.
3. Kipyatkova I., Kagiroy I. Phone Durations Modeling for Livvi-Karelian ASR. *Proceedings 25th International Conference Speech and Computer (SPECOM 2023)*. Springer LNCS. 2023. vol. 14339. pp. 87–99. DOI: 10.1007/978-3-031-48312-7\_7.
4. Kipyatkova I., Kagiroy I., Dolgushin M., Rodionova A. Towards a Livvi-Karelian End-to-End ASR System. *Proceedings 26th International Conference on Speech and Computer (SPECOM 2024)*. 2024. vol. 15299. pp. 57–68. DOI: 10.1007/978-3-031-77961-9\_4.
5. Kipyatkova I.S., Kagiroy I.A. [Analytical review of methods for solving data scarcity issues regarding elaboration of automatic speech recognition systems for low-resource languages]. *Informatika i avtomatizacija – Informatics and Automation*. 2022. vol. 21.no. 4. pp. 678–709. DOI: 10.15622/ia.21.4.2. (In Russ.).
6. Baeviski A., Zhou Y., Mohamed A., Auli M. wav2vec 2.0: A Framework for Self-Supervised Learning of Speech Representations. *Advances in Neural Information Processing Systems*. 2020. vol. 33. pp. 12449–12460.

7. Conneau A., Baevski A., Collobert R., Mohamed A., Auli M. Unsupervised Cross-Lingual Representation Learning for Speech Recognition. Proceedings of the Annual Conference of the International Speech Communication Association (Interspeech'2021). 2021. pp. 2426–2430. DOI: 10.21437/Interspeech.2021-329.
8. Babu A., Wang C., Tjandra A., Lakhota K., Xu Q., Goyal N., Singh K., Platen von P., Saraf Y., Pino J., Baevski A., Conneau A., Auli M. XLS-R: Self-Supervised Cross-Lingual Speech Representation Learning at Scale. Proceedings of the Annual Conference of the International Speech Communication Association (Interspeech'2022). 2022. pp. 2278–2282.
9. Pratap V., Tjandra A., Shi B., Tomasello P., Babu A., Kundu S., Elkahky A., Ni Zh., Vyas A., Fazel-Zarandi M., Baevski A., Adi Y., Zhang X., Hsu W.-N., Conneau A., Auli M. Scaling Speech Technology to 1,000+ Languages. Journal of Machine Learning Research. 2024. vol. 25. pp. 1–52.
10. Poth C., Sterz H., Paul I., Purkayastha S., Engländer L., Imhof T., Vulić I., Ruder S., Gurevych I., Pfeiffer J. Adapters: A unified Library for Parameter-Efficient and Modular Transfer Learning. Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing: System Demonstrations (EMNLP'2023). 2023. pp. 149–160. DOI: 10.18653/v1/2023.emnlp-demo.13.
11. Chung Y.A., Zhang Y., Han W., Chiu C.-C., Qin J., Pang R., Wu Y. W2v-bert: Combining Contrastive Learning and Masked Language Modeling for Self-Supervised Speech Pre-Training. Proceedings of 2021 IEEE Automatic Speech Recognition and Understanding Workshop (ASRU'2021). 2021. pp. 244–250. DOI: 10.1109/ASRU51503.2021.9688253.
12. Devlin J., Chang M.W., Lee K., Toutanova K. Bert: Pre-Training of Deep Bidirectional Transformers for Language Understanding. Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL-HLT'2019). 2019. vol. 1. pp. 4171–4186.
13. Chen S., Wang C., Chen Z., Wu Y., Liu S., Chen Z., Li J., Kanda N., Yoshioka T., Xiao X., Wu J., Zhou L., Ren S., Qian Y., Qian Y., Wu J., Zeng M., Yu X., Wei F. WavLM: Large-Scale Self-Supervised Pre-Training for Full Stack Speech Processing. IEEE Journal of Selected Topics in Signal Processing. 2022. vol. 16. no. 6. pp. 1505–1518.
14. Barrault L., Chung Y.A., Meglioli M.C., Dale D., Dong N., Duppenhaler M. et al. Seamless: Multilingual Expressive and Streaming Speech Translation. arXiv preprint arXiv:2312.05187. 2023.
15. Radford A., Kim J.W., Xu T., Brockman G., McLeavey C., Sutskever I. Robust Speech Recognition via Large-Scale Weak Supervision. Proceedings of the 40th International Conference on Machine Learning 2022 (ICML'23). 2023. pp. 28492–28518.
16. Williams A., Demarco A., Borg C. The applicability of Wav2Vec 2.0 and Whisper for Low-Resource Maltese ASR. Proceedings of the 2nd Annual Meeting of the ELRA/ISCA SIG on Under-resourced Languages (SIGUL'2023). 2023. pp. 39–43.
17. Kozhimbayev Z. Kazakh Speech Recognition: Wav2vec2.0 vs. Whisper. Journal of Advances in Information Technology. 2023. vol. 14. no. 6. pp. 1382–1389. DOI: 10.12720/jait.14.6.1382-1389.
18. Richardson F., Reynolds D., Dehak N. Deep Neural Network Approaches to Speaker and Language Recognition. IEEE Signal Processing Letters. 2015. vol. 22. no. 10. pp. 1671–1675. DOI: 10.1109/LSP.2015.2420092.
19. Winata G.I., Madotto A., Wu C.S., Fung P. Code-Switching Language Modeling using Syntax-Aware Multi-Task Learning. Proceedings of the Third Workshop on

- Computational Approaches to Linguistic Code-Switching (CALCS'2018). 2018. pp. 62–67. DOI: 10.18653/v1/W18-3207.
20. Adel H., Vu N.T., Kraus F., Schlippe T., Li H., Schultz T. Recurrent Neural Network Language Modeling for Code Switching Conversational Speech. Proceedings of 2013 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP'2013). 2013. pp. 8411–8415. DOI: 10.1109/ICASSP.2013.6639306.
  21. Ramanarayanan V., Pugh R., Suenderman-Oeft D. Automatic Turn-Level Language Identification for Code-Switched Spanish-English Dialog. Proceedings of 9th International Workshop on Spoken Dialogue System Technology (IWSDS'2019). 2019. vol. 579. pp. 51–61. DOI: 10.1007/978-981-13-9443-0\_5.
  22. Mustafa M.B., Yusof M.A., Khalaf H.K., Abushariah A.A.R.M., Kiah M.L.M., Ting H.N., Muthaiyah S. Code-Switching in Automatic Speech Recognition: The Issues and Future Directions. Applied Sciences. 2022. vol. 12. no. 19. DOI: 10.3390/app12199541.
  23. Çetinoğlu Ö., Schulz S., Vu N.T. Challenges of Computational Processing of Code-Switching. Proceedings of the Second Workshop on Computational Approaches to Linguistic Code Switching (CALCS'2016). 2016. pp. 1–11. DOI: 10.18653/v1/W16-5801.
  24. Winata G., Aji A.F., Yong Z.X., Solorio T. The Decades Progress on Code-Switching Research in NLP: A Systematic Survey on Trends and Challenges. Findings of the Association for Computational Linguistics (ACL'2023). 2023. pp. 2936–2978. DOI: 10.18653/v1/2023.findings-acl.185.
  25. Hsieh I.T., Wu C.H., Wang C.H. Acoustic and Textual Data Augmentation for Code-Switching Speech Recognition in Under-Resourced Language. IEEE Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC'2020). 2020. pp. 302–307.
  26. Chang C.-T., Chuang S.-P., Lee H.-Y. Code-Switching Sentence Generation by Generative Adversarial Networks and its Application to Data Augmentation. Proceedings of the Annual Conference of the International Speech Communication Association (Interspeech'2019). 2019. pp. 554–558. DOI: 10.21437/Interspeech.2019-3214.
  27. Chan J.Y.C., Cao H., Ching P.C., Lee T. Automatic recognition of Cantonese-English Code-Mixing Speech. International Journal of Computational Linguistics and Chinese Language Processing. 2009. vol. 14. no. 3. pp. 281–304.
  28. Adel H., Vu N.T., Kirchhoff K., Telaar D., Schultz T. Syntactic and Semantic Features for Code-Switching Factored Language Models. IEEE/ACM Transactions on Audio, Speech, and Language Processing. 2015. vol. 23. no. 3. pp. 431–440. DOI: 10.1109/TASLP.2015.2389622.
  29. Hermann K.M., Blunsom P. Multilingual Models for Compositional Distributed Semantics. Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers). 2014. pp. 58–68. DOI: 10.3115/v1/P14-1006.
  30. Conneau A., Khandelwal K., Goyal N., Chaudhary V., Wenzek G., Guzmán F., Grave É., Ott M., Zettlemoyer L., Stoyanov V. Unsupervised Cross-Lingual Representation Learning at Scale. Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics (ACL'2020). pp. 8440–8451. DOI: 10.18653/v1/2020.acl-main.747.
  31. Vüren van J., Niesler T. Improving N-best Rescoring in Under-Resourced Code-Switched Speech Recognition using Pretraining and Data Augmentation. Languages. 2022. vol. 7. no. 3. DOI: 10.3390/languages7030236.
  32. Hono Y., Mitsuda K., Zhao T., Mitsui K., Wakatsuki T., Sawada K. Integrating Pre-Trained Speech and Language Models for End-to-End Speech Recognition. Findings

- of the Association for Computational Linguistics: ACL 2024. 2024. pp. 13289–13305. DOI: 10.18653/v1/2024.findings-acl.787.
33. Ogunremi T., Manning C.D., Jurafsky D. Multilingual Self-Supervised Speech Representations Improve the Speech Recognition of Low-Resource African Languages with Code Switching. arXiv preprint arXiv:2311.15077. 2023.
  34. Hori T., Cho J., Watanabe S. End-to-end speech recognition with word-based RNN language models. Proceedings of 2018 IEEE Spoken Language Technology Workshop (SLT-2018). 2018. pp. 389–396. DOI: 10.1109/SLT.2018.8639693.
  35. Povey D., Ghoshal A., Boulianne G., Burget L., Glembek O., Goel N., Hannemann M., Motlíček O., Qian Y., Schwarz P., Silovský J., Stemmer G., Veselý K. The Kaldi Speech Recognition Toolkit. Proceedings of IEEE Automatic Speech Recognition and Understanding Workshop (ASRU'2011). 2011. pp. 1–4.
  36. Kipyatkova I.S., Rodionova A.P., Kagirov I.A., Krizhanovsky A.A. [Speech and text data preparation for developing of an automatic speech recognition system for the Karelian language]. Uchjonye zapiski Petrozavodskogo gosudarstvennogo universiteta – Proceedings of Petrozavodsk State University. 2023. vol. 45. no. 5. pp. 89–98. (In Russ.).
  37. Povey D., Cheng G., Wang Y., Li K., Xu H., Yarmohammadi M., Khudanpur S. Semi-Orthogonal Low-Rank Matrix Factorization for Deep Neural Networks. Proceedings of The Annual Conference of the International Speech Communication Association (Interspeech'2018). 2018. pp. 3743–3747. DOI: 10.21437/Interspeech.2018-1417.
  38. Saon G., Soltau H., Nahamoo D., Picheny M. Speaker Adaptation of Neural Network Acoustic Models using i-Vectors. Proceedings of IEEE Automatic Speech Recognition and Understanding Workshop (ASRU'2013). 2013. pp. 55–59. DOI: 10.1109/ASRU.2013.6707705.
  39. Stolcke A., Zheng J., Wang W., Abrash V. SRILM at Sixteen: Update and Outlook. Proceedings of IEEE Automatic Speech Recognition and Understanding Workshop (ASRU'2011). 2011. pp. 5–9.
  40. Enarvi S., Kurimo M. TheanoLM – An Extensible Toolkit for Neural Network Language Modeling. Proceedings of the Annual Conference of the International Speech Communication Association (Interspeech'2016). 2016. pp. 3052–3056. DOI: 10.21437/Interspeech.2016-618.
  41. Karpov A., Markov K., Kipyatkova I., Vazhenina D., Ronzhin A. Large Vocabulary Russian Speech Recognition using Syntactico-Statistical Language Modeling. Speech Communication. 2014. vol. 56. pp. 213–228. DOI: 10.1016/j.specom.2013.07.004.
  42. Wolf T., et al. Transformers: State-of-the-Art Natural Language Processing. arXiv preprint arXiv:1910.03771. 2019.
  43. Grosz T., Getman Y., Al-Ghezi R., Rouhe A., Kurimo M. Investigating wav2vec2 Context Representations and the Effects of Fine-Tuning, a Case-Study of a Finnish Model. Proceedings of The Annual Conference of the International Speech Communication Association (Interspeech'2023). 2023. pp. 196–200. DOI: 10.21437/Interspeech.2023-837.
  44. Hiovain-Asikainen K., Rosa de la J. Developing TTS and ASR for Lule and North Sámi Languages. Proceedings of the 2nd Annual Meeting of the ELRA/ISCA SIG on Under-resourced Languages (SIGUL'2023). 2023. pp. 48–52. DOI: 10.21437/SIGUL.2023-11.
  45. Paonessa C., Timmel V., Vogel M., Perruchoud D. Whisper Fine-Tuning for Swiss German: A Data Perspective. Proceedings of the 9th edition of the Swiss Text Analytics Conference. 2024. pp. 192.

**Kipyatkova Irina** — Ph.D., Associate Professor, Senior researcher, Speech and multimodal interfaces laboratory, St. Petersburg Federal Research Center of the Russian Academy of Sciences (SPC RAS). Research interests: automatic speech recognition, neural networks. The number of publications — 100. kipyatkova@iias.spb.su; 39, 14-th Line V.O., 199178, St. Petersburg, Russia; office phone: +7(812)328-0421.

**Kagirov Ildar** — Research fellow, Speech and multimodal interfaces laboratory, St. Petersburg Federal Research Center of the Russian Academy of Sciences (SPC RAS). Research interests: corpus linguistics, low-resource languages. The number of publications — 40. kagirov@iias.spb.su; 39, 14-th Line V.O., 199178, St. Petersburg, Russia; office phone: +7(812)328-0421.

**Dolgushin Mikhail** — Junior research fellow, Speech and multimodal interfaces laboratory, St. Petersburg Federal Research Center of the Russian Academy of Sciences (SPC RAS). Research interests: automatic speech recognition, neural networks, mathematical linguistics. The number of publications — 10. dolgushin.m@iias.spb.su; 39, 14-th Line V.O., 199178, St. Petersburg, Russia; office phone: +7(812)328-0421.

**Acknowledgements.** This research is supported by the Russian Science Foundation (project № 24-21-00276, <https://rscf.ru/en/project/24-21-00276/>).

A. NAVEEN, I. JACOB, A. MANDAVA  
**DETECTION OF STUDENT ENGAGEMENT VIA  
TRANSFORMER-ENHANCED FEATURE PYRAMID NETWORKS  
ON CHANNEL-SPATIAL ATTENTION**

*Naveen A., Jacob I., Mandava A.* **Detection of Student Engagement via Transformer-Enhanced Feature Pyramid Networks on Channel-Spatial Attention.**

**Abstract.** One of the most important aspects of contemporary educational systems is student engagement detection, which involves determining how involved, attentive, and active students are in class activities. For educators, this approach is essential as it provides insights into students' learning experiences, enabling tailored interventions and instructional enhancements. Traditional techniques for evaluating student engagement are often time-consuming and subjective. This study proposes a novel real-time detection framework that leverages Transformer-enhanced Feature Pyramid Networks (FPN) with Channel-Spatial Attention (CSA), referred to as BiusFPN\_CSA. The proposed approach automatically analyses student engagement patterns, such as body posture, eye contact, and head position, from visual data streams by integrating cutting-edge deep learning and computer vision techniques. By integrating the attention mechanism of CSA with the hierarchical feature representation capabilities of FPN, the model can accurately detect student engagement levels by capturing contextual and spatial information in the input data. Additionally, by incorporating the Transformer architecture, the model achieves better overall performance by effectively capturing long-range dependencies and semantic relationships within the input sequences. Evaluation using the WACV dataset demonstrates that the proposed model outperforms baseline techniques in terms of accuracy. Specifically, in terms of accuracy, the FPN\_CSA\_Trans\_EH variant of the proposed model outperforms FPN\_CSA by 3.28% and 4.98%, respectively. These findings underscore the efficacy of the BiusFPN\_CSA framework in real-time student engagement detection, offering educators a valuable tool for enhancing instructional quality, fostering active learning environments, and ultimately improving student outcomes.

**Keywords:** Feature Pyramid Network (FPN), Channel-Spatial Attention (CSA), student engagement detection, Transformer.

**1. Introduction.** When students are actively involved in their educational assignments and activities, this is referred to as student engagement. This involvement not only directly impacts on school improvements, such as enhancing teachers' professional identities and fostering a welcoming school environment [1], but it also appears to boost the academic performance of underperforming students, reduce dropout rates, and decrease dissatisfaction. Because of this, scholars have remained highly interested in student engagement and its various implications over the past 20 years. Academic success has always been considered a crucial result of student engagement. With the rise of network technology, computer technology, and other advancements, online learning has emerged. It emphasises communication between students and the accessibility of the learning resources. A significant number of students

now participate in online learning, which has become a predominant learning method. However, due to limited interaction between teachers and students, communication is often insufficient, student participation in online learning tends to be suboptimal, inconsistent, and inefficient. Student engagement is essential for learning and significantly impacts online learning as well.

As the phrase "student engagement" can have different meanings for different people, the method of assessing student engagement utilised by researchers in their studies is akin to selecting a specific conceptualization of the construct. Prior to choosing a method for assessing student engagement, it is important to define the term precisely. Early researchers frequently operationalized student engagement in terms of observable behaviours, such as the level of participation in various tasks and the time required to complete them. Various aspects, such as facial expression recognition, head pose detection, and body language analysis, can be used to assess student attention in class. Facial expressions have been used in studies to analyse student participation in both in-person and online classes, and results have demonstrated that this method is effective for determining student engagement levels. This method has been used in studies to evaluate student participation in face-to-face and online classes, and it has proven effective in determining levels of engagement. For instance, the authors in [19] proposed a system that employs facial expression recognition to measure engagement levels in real time, and they demonstrated its effectiveness in a classroom setting.

In order to support timely intervention, it is essential to assess and study student engagement in online learning, help teachers understand student engagement, enable students to reflect on their learning, and encourage participation in the learning process. The quantification of student behaviour, cognitive engagement, and emotional engagement are all aspects of measuring student engagement. Currently, research has focused on studying student participation through theoretical models, explicit behavioural data, influencing factors, effect analysis, and the lack of accurate measurement of student engagement. As a result, the research developed a reliable and quantifiable model of student engagement, examining students' cognitive engagement, emotional engagement, and engagement patterns. Researchers are exploring ways to incorporate transformers into computer vision applications due to their powerful representational capabilities. Transformer-based models outperform convolutional and recurrent neural networks across various visual benchmarks, with some models achieving superior performance. The computer vision community is increasingly focusing on transformers due to

their strong performance and reduced need for inductive bias tailored to specific types of vision. Transformer, a type of deep neural network primarily based on the self-attention mechanism, was first applied in natural language processing. In this study, an enhanced transformer is presented to identify student engagement. The transformer is applied for the first time in engagement recognition. The findings reveal a positive correlation between the final exam scores and the level of student participation in the online class sessions.

The main contributions of this work are summarized as follows:

1. We propose a Feature Pyramid Network model with a location and channel-aware attention module to effectively learn facial representations during online class sessions.
2. The transformer module is also integrated with this model to analyse the global context features along with local convolutional features.
3. The encoded FPN-based significant features are combined with the eye and head movement-based features to enhance the performance of the framework.
4. The performance of the proposed FPN\_CSA\_Trans is analyzed with the help of DAiSEE and WACV datasets.

**2. Related Work.** Facial expression analysis requires more advanced expertise in the field of computer vision. In recent years, there has been growing interest in using technology to track and assess students' facial expressions in order to better understand and enhance their engagement in class. The emotions of students can be inferred from their facial expressions, which are a nonverbal form of communication [1]. Analysis of body language was employed by some researchers to gauge levels of engagement. This method has been used in studies to assess student participation in both traditional and online classroom settings, demonstrating its usefulness in determining levels of engagement. This may involve analysing a student's body in various positions, such as sitting or standing, as well as their head and gaze. According to research in this field, body position evaluation may be employed to determine levels of engagement.

For example, the authors in [2] used body posture and movement analysis to assess engagement levels in real time, proving its efficacy in a classroom context, but it lacked facial-based features. Numerous other publications [3, 4] also proposed using body language analysis to gauge student engagement during lectures. According to the authors in [3], who used a sample of 800 students, the accuracy rate for identifying engagement levels was 89.3%. Keystroke dynamics, or the analysis of typing habits, such as speed and errors, has been widely used by researchers to identify levels of engagement. According to research, keystroke dynamics can be



used to gauge student interest levels and enhance the effectiveness of instruction. Keystroke dynamics may vary depending on the scenario. Another method for determining engagement levels involves analysing body language and head position. This may involve analysing body posture, eye contact, and head position. According to research in this field, head posture and body language can be used to gauge student engagement levels and improve instructional efficacy. The use of head and body posture alone lacks the ability to extract features from the facial region.

The authors in [5, 6] proposed a real-time student engagement detection method. Students' eye tracking and head movements were recorded using a depth camera, and machine learning techniques were used to classify their engagement levels. They used a depth camera to capture students' eye tracking and head movements, and applied machine learning algorithms to classify engagement levels.

Another approach involves using head pose and body movements to detect engagement levels. This may include analysing the position of head, eye gaze, and body posture [20]. Research in this area has shown that head pose and body movements can be used to detect engagement levels and improve the effectiveness of teaching. To address occlusion, researchers in [7 – 10] used texture features or reconstructed geometric features. To recover a lost or drifted facial point, an improved Kanade-Lucas tracker [7] was proposed. PCA-based approaches were used for missing point reconstruction [8, 9]. Another method for identifying facial expressions, known as the modified transferable belief model, was proposed in [10]. The performance of the facial expression analysis mechanism can be affected by facial poses. To address pose variations, the authors recommended training with a single classifier [11]. Adversarial feature learning [12] was used by researchers for the same purpose. Study [13] proposed using the k-means algorithm to group students based on 12 engagement measures divided into two categories: interaction-related and effort-related. Quantitative analysis is used to identify students who are not engaged and may require assistance. In order to identify student participation in a classroom setting, the authors in [14] proposed a real-time facial expression detection system. They recorded students' faces using a camera and then applied machine learning algorithms to classify their facial expressions into different levels of engagement. The authors in [15] proposed a deep learning-based method for identifying student interest in video-based online classes, using aspects of facial expression, head pose, and gaze. They found that their approach achieved high accuracy in determining engagement levels. Using a Histogram of Oriented Gradients, further trained by a CNN, the authors in [16] were able to extract facial features. Since the histogram of oriented

gradients extracts information from spatial gradients, this study achieved better performance. Studies [17, 18] also used variables alongside facial expression to determine student interest in the classroom. The authors in [19] used this method to analyse student participation in online classes, demonstrating its effectiveness in determining engagement levels. According to another study [20], engagement levels can be determined using a combination of head pose, facial expression, and gaze analysis. The study found that different modalities are used for student engagement analysis. Each modality has its own strengths. Therefore, in this paper, features learned from face, eye, and head movements are used for training and testing.

The analysis of various studies shows that student engagement through modalities, such as facial expressions, head movements, gaze tracking, body posture, and keystroke dynamics, has yielded promising results in both traditional and online classrooms. Each modality has its strengths but also presents limitations. For example, methods focusing on body posture or keystroke dynamics often overlook finer details of facial expressions, while facial expression analysis may struggle with variations caused by occlusion, lighting conditions, and pose differences.

### **3. Methodology**

#### **3.1. The proposed FPN\_CSA model for Engagement Analysis.**

The current study aims to investigate three approaches to applying facial recognition technology in classroom engagement analysis. By leveraging the Transformer mechanism and the BiusFPN with Inter-Cross Coordinate Self-Attention model, a person's engagement based on facial expression during online sessions can be recognised. Frames are extracted from the real-time video stream and used for face detection. Eye tracking is performed after face detection, and head rotation is also analysed. These three metrics are used to determine the engagement level. The outcome is classified as disengaged, partially engaged, and fully engaged. This method effectively identifies specific facial expressions associated with varying degrees of engagement. The overall architecture of the proposed model is depicted in Figure 1.

Using Resnet-18 as the backbone, the Feature Pyramid Network (FPN) is employed in the design of the proposed framework for engagement analysis. For a single flow-down sampling process, this bi-directional Feature Pyramid Network (FPN) manages two upsampling operations.

This model integrates traditional channel and spatial attention at the second level of the upsampling process, along with an Inter-Cross Coordinate Self-Attention model. Figure 2 illustrates the general architecture of the proposed framework.

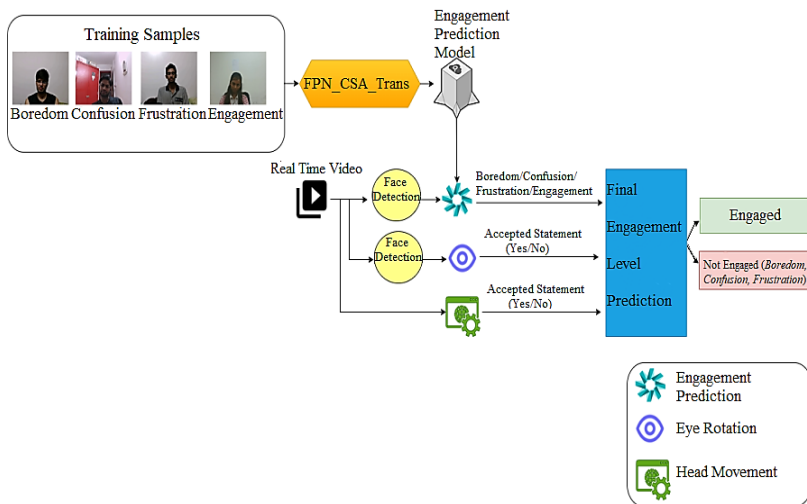


Fig. 1. Overall architecture of the proposed work

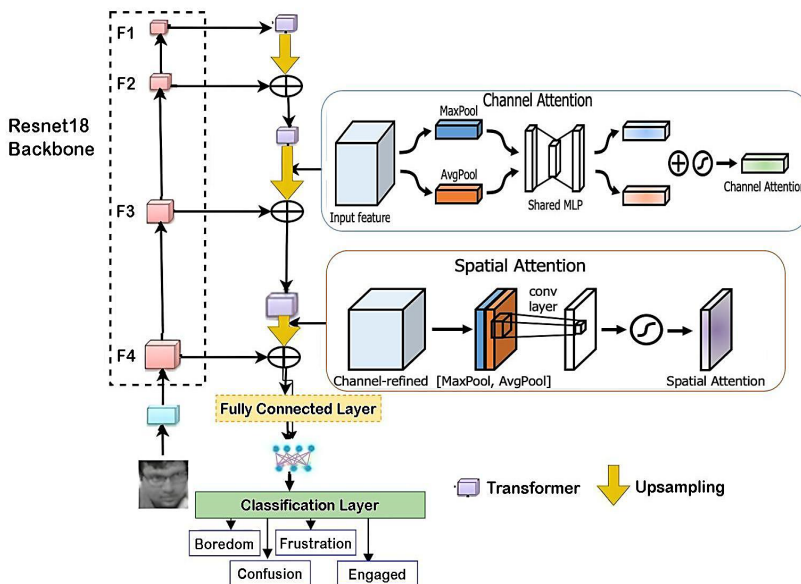


Fig. 2. Architecture of the proposed FPN\_CSA\_Trans model

**3.1.1. Resnet-18.** The term "ResNet" refers to the 18-layer Convolutional Neural Network introduced by [21]. Designed to facilitate the effective operation of multiple convolutional layers, ResNet-18 is a 72-layer architecture with 18 deep layers, including residual blocks, as described in [22]. However, as a network is expanded with multiple deep layers, the output performance usually deteriorates. The vanishing gradient problem is addressed by neural networks using gradient descent to determine the weights that minimise the loss function during backpropagation training. The gradient "vanishes," leading to network saturation or even performance loss due to repeated multiplication across multiple layers. Residual Blocks in ResNet-18 utilise skip connections to address the vanishing gradient issue. The skip connection bypasses a few intermediate levels in order to connect layer activations to subsequent layers. As a result, the residual block remains intact. The approach used in this system allows the residual mapping to fit the system, rather than requiring the layers to learn the underlying mapping directly. Figure 3 illustrates the skip connection mechanism within a residual block.

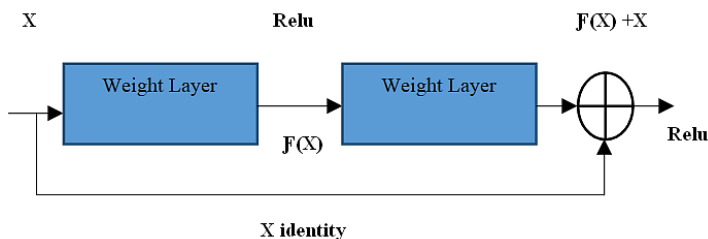


Fig. 3. Skip connections

The benefit of skip connections is that they prevent any layer from degrading the network's performance. Thus, vanishing or exploding gradients do not pose problems when training very deep neural networks. The backbone network for the proposed model is the Resnet-18.

**3.1.2. Transformer with Multi-Head Attention.** Numerous advances in deep learning tasks have resulted from Transformers [Vaswani et al., 2017; Devlin et al., 2019; Velickovic et al., 2018b]. The Transformer stands out due to its ability to combine all computations in the same layer and its lack of recurrent connections, which improves scalability, effectiveness, and efficiency. The Transformer only uses the attention mechanism to determine the dependencies between input tokens, eliminating the need for recurrent connections. To be more precise, the Transformer utilises a novel multi-head attention module designed to more effectively recognise the dependencies between input tokens.

It has been noted that a key factor in the Transformer's success is its multi-head attention module. Recurrent neural networks (RNNs) have been shown to outperform Transformers on machine translation benchmarks when both utilise multi-head encoder-decoder attention. In contrast, Transformers perform worse when not utilising multi-head attention [5]. In addition to the Transformer, multi-head attention has been implemented in RNNs [5], Graph Attention Networks [Velickovic et al., 2018a], Convolutional Neural Networks [Xiao et al., 2020; Fang et al., 2019], and other architectures.

The overarching belief is that multi-head attention distinguishes itself by attending to multiple positions concurrently, whereas a conventional attention module can only focus on one position in a single layer. Multi-head attention specifically performs multiple attention computations in parallel and projects the input data into multiple distinct subspaces (Figure 4).

The combination of ResNet-18, FPN, and Transformers leverages the strengths of each architecture, making it a powerful choice for analysing visual data where both spatial and sequential information must be effectively captured, such as in assessing student engagement levels from video feeds. This setup addresses both the efficiency and depth of feature extraction required for accurate and real-time predictions, outperforming the narrower focus and limitations of RNNs or standalone CNNs in such applications.

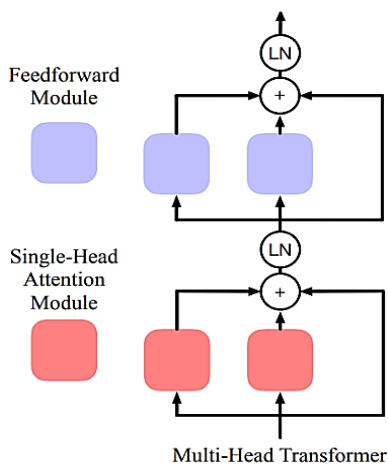


Fig. 4. Multi-Head Transformer

**3.1.3. Feature Pyramid.** An essential component of recognition algorithms is the feature pyramid, which is used to detect objects at different scales. An FPN is a multi-scale deep learning framework that builds feature pyramids with minimal additional computational cost. The design is widely integrated to create high-level semantic maps at all scales. In various applications, this performs significantly better as a general feature extractor. In recognition tasks, convolutional networks (ConvNets) have largely replaced handcrafted features. ConvNets are more robust in representing higher-level semantics with scale invariance that can identify computed features from a single input scale. Multi-scale feature extraction is achieved by characterising the high-resolution features at each pyramid level. Feature pyramids are constructed using top-down, bottom-up, and lateral connections [23].

The semantically stronger features, which are subsampled fewer times and thus have more accurate localisation information, are combined with features from earlier levels in the architecture developed by the Feature Pyramid Network (FPN). When the FPN, which serves as a feature extractor, was developed, the precision and speed of the pyramid concept were key considerations. Unlike detectors such as the extractor for object recognition in Faster R-CNN, the FPN generates multiple feature map layers with higher-quality data compared to conventional feature pyramids. The use of multi-scale feature maps from multiple layers computed during the forward pass makes it computationally efficient. Given its numerous advantages, our proposed model, which uses ResNet-18 as the backbone, incorporates the feature pyramid.

We used a  $100 \times 100$  grayscale input image containing only detectable faces in our model. Initially, a convolutional layer with 64 filters and a  $3 \times 3$  kernel size was applied to the input image. The convolutional layer produces an output of size  $100 \times 100 \times 64$ . Subsequently, these feature maps are passed to the ResNet-18. In ResNet-18, we utilised four different types of convolutional layers. Four feature maps, designated as F1, F2, F3, and F4, were produced. The first convolutional layer in ResNet-18, corresponding to F4, has 64 filters, a kernel size of  $3 \times 3$ , and a stride of 1. As a result, the F4 feature map retains the original size of  $100 \times 100$  with 64 filters. The second convolutional layer in ResNet-18, corresponding to F3, has 128 filters, a kernel size of  $3 \times 3$ , and a stride of 2. As a result, the F3 feature map has dimensions of  $50 \times 50 \times 64$ . The third convolutional layer in ResNet-18, corresponding to F2, uses 256 filters, a kernel size of  $3 \times 3$ , and a stride of 2. Consequently, the F2 feature map has dimensions of  $25 \times 25 \times 256$ . The final convolutional layer in ResNet-18, corresponding to F1, has 512 filters, a kernel size of  $3 \times 3$ , and a stride of 2. As a result, the F1

feature map has dimensions of  $13 \times 13 \times 512$ . The ResNet-18 processing is now complete. Next, we proceed to the pyramidal structures in the Feature Pyramid Network (FPN) using upsampling and addition.

Subsequent processing is applied to each of the feature maps: F1, F2, F3, and F4. Following the F1 feature map, a 2D convolutional layer with 256 filters, a kernel size of  $1 \times 1$ , a stride of 1, dilation of 1, groups of 1, and ReLU activation is applied. It is then passed through a transformer layer and upsampled before being merged with F2. The result of the first addition, add1, has dimensions of  $25 \times 25 \times 256$ . The output of add1 is then passed through a convolutional layer with 128 filters. After passing through a transformer layer, it is upsampled and merged with F3. The result of the second addition, add2, has dimensions of  $50 \times 50 \times 128$ . The output of add2 is then passed through a convolutional layer with 64 filters. After that, it is upsampled and merged with F4. The result of the third addition, add3, has dimensions of  $100 \times 100 \times 64$ . To provide input for the second stage of upsampling, the output from the first upsampling in the FPN is passed through three distinct attention mechanisms: Channel Attention, Spatial Attention, and the proposed ICCSA.

**3.1.4. Channel and Spatial Attention.** A channel attention module [24] is utilised in convolutional neural networks to provide channel-based attention. Figure 5 illustrates the channel attention architecture. An attention mechanism is introduced to create a channel attention map by leveraging the relationships between features across channels. Given an input image, the channels of a feature map act as feature detectors, and thus, channel attention focuses on "what" is important. Effective channel attention computation requires a reduction of the spatial dimension of the input feature map. Using both average-pooling and max-pooling processes, the spatial information of a feature map is aggregated to produce two unique spatial context descriptors,  $F_{avg}^c$  and  $F_{max}^c$ , representing average-pooled features and max-pooled features, respectively. Subsequently, both descriptors are passed through a shared network, which generates the channel attention map  $M_c \in R^{c \times 1 \times 1}$ . Here, C represents the number of channels. The shared network consists of multi-layer perceptrons (MLPs) with a single hidden layer. To minimise parameter overhead, the hidden activation size is set to  $R^{c/r \times 1 \times 1}$ , where r is the reduction ratio. After processing each descriptor through the shared network, the output feature vectors are combined element-wise. In summary, channel attention is computed as follows: here, F represents the input feature, where AvgPool and MaxPool denote the average and max pooling operations, as shown in Figure 5. Equation (1) represents the overall process of channel attention, as illustrated in Figure 5, with the corresponding module notations.

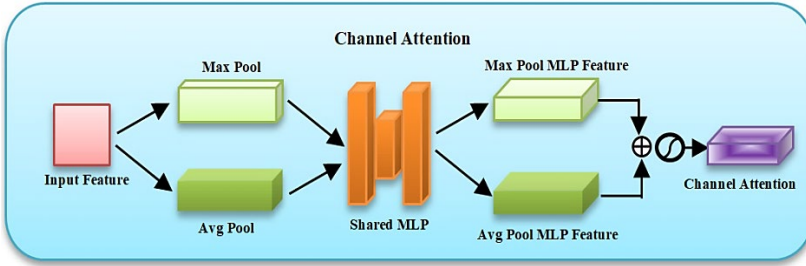


Fig. 5. Channel Attention

$$\begin{aligned}
 M_c(F) &= \sigma(\text{MLP}(\text{AvgPool}(F)) + \text{MLP}(\text{MaxPool}(F))), \\
 M_c(F) &= \sigma(W_1(W_0(F_{avg}^c)) + W_1(W_0(F_{max}^c))),
 \end{aligned} \tag{1}$$

where  $W_0 \in R^{c/r \times c}$ ,  $W_1 \in R^{c \times c/r}$ , and  $\sigma$  denotes the sigmoid function. Note that the MLP weights  $W_0$  and  $W_1$  for both inputs are followed by the ReLU activation function.

The spatial attention module is another component in convolutional neural networks [24]. It generates a spatial attention map by leveraging the spatial relationships among features. Spatial attention focuses on "where" information is located, in contrast to channel attention, which focuses on "what" is informative. Before computing spatial attention, average-pooling and max-pooling operations are performed along the channel axis, and the results are concatenated to produce an effective feature descriptor. The spatial attention map  $M_s(F) \in R^{H \times W}$ , which encodes where to emphasise or suppress, is generated by applying a convolution layer to the concatenated feature descriptor. Two 2D maps,  $F_{avg}^s \in R^{1 \times H \times W}$  and  $F_{max}^s \in R^{1 \times H \times W}$ , are generated by pooling the channel information in a feature map using two different methods. Each map represents the max-pooled and average-pooled features of the channel, respectively. These maps are concatenated and convolved using a standard convolution layer to generate the 2D spatial attention map. In summary, the computation of spatial attention is described in Equation 2, with the corresponding module notations and mathematical representations.

$$\begin{aligned}
 M_s(F) &= \sigma(f^{7 \times 7}([\text{AvgPool}(F); \text{MaxPool}(F)])), \\
 M_s(F) &= \sigma(f^{7 \times 7}([F_{avg}^s; F_{max}^s])),
 \end{aligned} \tag{2}$$



where the convolution operation with a  $7 \times 7$  filter size is denoted by  $f^{7 \times 7}$ , and  $\sigma$  represents the sigmoid function. Figure 6 illustrates the spatial attention architecture.

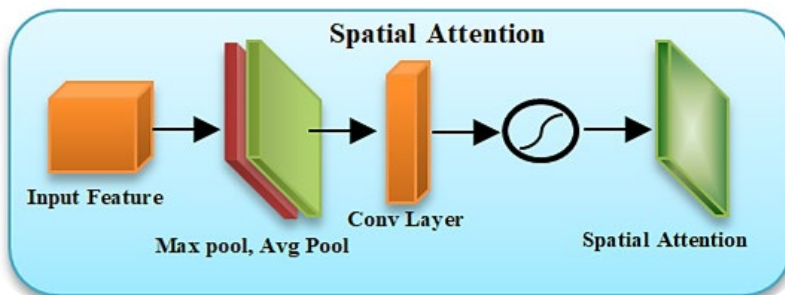


Fig. 6. Spatial Attention

In the decoding upsampling flow, the input is first processed by the transformer block, as shown in Figure 2. After processing through the transformer block, the output is upsampled and combined with the corresponding next-level encoded output. In the second level, the process is repeated, but channel attention features are added to extract more channel-oriented information. In the third level of FPN decoding, the transformer is incorporated, followed by upsampling and addition with the spatial attention module. The final third-level addition block produces a feature map of size  $100 \times 100 \times 64$ . After the fully connected layer, the output is passed to a final fully connected layer with four classes and a filter size of 1,086. The classification layer classifies the input image into one of the categories: Disengaged, Partially Engaged, or Engaged.

**4. Experimental Results.** The proposed FPN\_CSA\_Trans\_EH method achieved an accuracy of 71.02% on the DAiSEE dataset and 88.57% on the WACV dataset.

#### 4.1. Dataset Description

**4.1.1. DAiSEE dataset.** The DAiSEE dataset [25] is publicly available. It consists of video recordings of participants in an e-learning environment, annotated with publicly sourced labels for engagement, frustration, confusion, and boredom. The dataset, made publicly available along with unique crowd-sourced annotations, captures real-world "in the wild" environments. The methodologies for data collection, annotation, and vote aggregation are described below. The DAiSEE dataset includes 9,068 clips from 112 students taking online courses. The four states of people watching online courses – boredom, confusion, frustration, and engagement – were annotated on the videos. Each state is assigned one of four ordinal

levels: 0 (very low), 1 (low), 2 (high), and 3 (very high). This work focuses solely on the classification of engagement levels. The clips are 10 seconds long, recorded at 30 frames per second (fps), with a resolution of  $640 \times 480$  pixels (Figure 7).

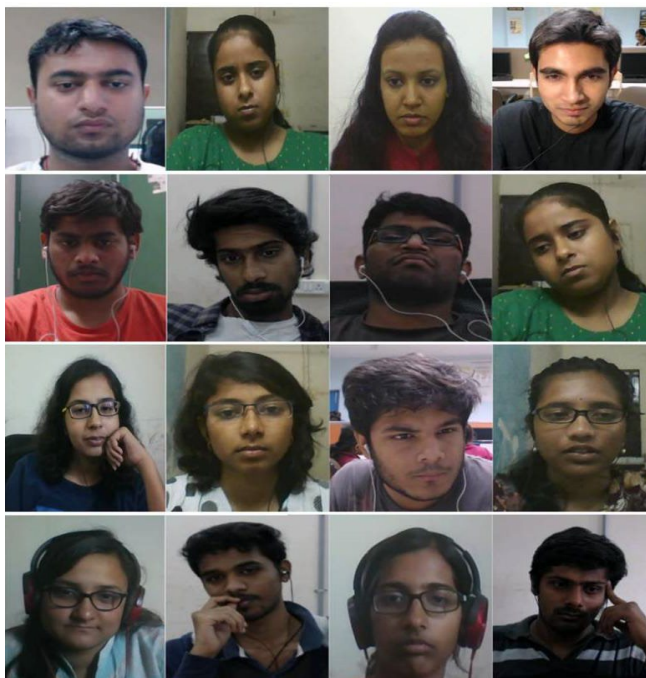


Fig. 7. Samples of the DAiSEE dataset:  
Engagement (first row), Boredom (second row), Confusion (third row),  
and Frustration (bottom row)

**4.1.2. WACV dataset.** This section describes the evaluation of student engagement levels using the WACV dataset. We used the open-source WACV dataset [28] for our research. The dataset contains three distinct classes: disengaged, partially engaged, and engaged. The dataset consists of 4,424 RGB images of varying sizes. All images were resized to a uniform shape of  $100 \times 100 \times 3$ . The dataset is not balanced, with 412 images in the "disengaged" class, 2,247 images in the "partially engaged" class, and 1,765 images in the "engaged" class. To create a balanced set of 412 photos for each class, we randomly selected 412 images from the "partially engaged" class and 412 images from the "engaged" class.

We divided this data into training and testing sets (80% and 20%, respectively). The figures below show class-wise examples from the WACV and the DAiSEE datasets.

Figure 8 illustrates the Disengaged, Partially Engaged, and Engaged samples of the WACV dataset; and Figure 7 illustrates the Boredom, Confusion, Frustration, and Engagement samples of the DAiSEE dataset.



Fig. 8. Engaged (top), partially engaged (middle), and disengaged (bottom) samples of the WACV dataset

#### 4.2. Evaluation metrics.

A. *Accuracy*. Accuracy measures the overall performance of the classifier. The model's performance is evaluated by comparing the percentage of accurate predictions to all cases. Accuracy is calculated using the following formula:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}. \quad (3)$$

B. *AUC*. The ROC curve's summary, Area Under the Curve (AUC), measures how successfully a classifier can distinguish between classes. A higher AUC indicates better performance in distinguishing between positive and negative classes.

C. *Gini Index*. The Gini Index is calculated by subtracting the sum of the squared probabilities of each class from one. It is simple to compute and tends to favour larger segments. In simple terms, it measures the

probability that a randomly selected feature is misclassified. The Gini index is calculated as follows:

$$GI = 1 - \sum_{i=1}^n (P_i)^2. \quad (4)$$

D. *AGF*. The AGF metric is an enhanced version of the F1 score that can accurately assess the performance of our algorithm even with imbalanced data. A high AGF value indicates that class imbalance had minimal or no impact on the results. The AGF is calculated using the following formula:

$$AGF = \sqrt{\text{inv}F_{0.5} \times F_2}. \quad (5)$$

## 5. Results and Discussion

**5.1. WACV Dataset.** The graphical representation of the accuracy comparison between the WACV dataset and existing methods is provided in Table 1. In terms of Accuracy, FPN\_CSA\_Trans\_EH outperforms FPN\_CSA\_Trans by 3.28%, FPN\_CSA by 4.98%, ResNet-18 [28] by 8.57%, DenseNet-121 [28] by 10.57%, MobileNetV1 [28] by 22.57%, HOG+SVM [28] by 30.57%, and CNN [27] by 51.57%.

Table 1. Accuracy of different methods for the WACV dataset

WACV	Accuracy(%)
CNN[27]	37
HOG+SVM[28]	58
MobileNetV1[28]	66
DenseNet-121[28]	78
ResNet-18[28]	80
FPN_CSA	83.59
FPN_CSA_Trans	85.29
<b>FPN_CSA_Trans_EH</b>	<b>88.57</b>

Figure 9 illustrates the class-wise accuracy comparison between ResNet-18 and FPN\_CSA\_Trans\_EH (Table 2). In terms of accuracy, FPN\_CSA\_Trans\_EH outperforms ResNet-18 by 6.1% for the disengaged class, 9.58% for the partially engaged class, and 7.37% for the engaged class.

Table 2. Comparison of metrics for ResNet-18 and FPN\_CSA\_Trans\_EH

Engaged	Partially engaged	Disengaged	Classes	
83.56%	76.83%	84.14%	Accuracy	ResNet18[28]
90.09%	84.3%	90.03%	AUC	
80.18%	68.6%	80.07%	Gini Index	
88.49%	83.31%	87.08%	AGF	
86.11%	81.73%	82.92%	Accuracy	FPN_CSA
88.35%	86.73%	87.78%	AUC	
76.7%	73.46%	75.57%	Gini Index	
88.36%	84%	85.13%	AGF	
87.53%	83.07%	87.8%	Accuracy	FPN_CSA_Trans
89.24%	88.77%	90.28%	AUC	
78.49%	77.55%	80.52%	Gini Index	
89.4%	85.63%	87.51%	AGF	
<b>90.93%</b>	<b>86.41%</b>	<b>90.24%</b>	<b>Accuracy</b>	<b>FPN_CSA_Trans_EH</b>
<b>92.17%</b>	<b>91.25%</b>	<b>92.06%</b>	<b>AUC</b>	
<b>84.34%</b>	<b>82.5%</b>	<b>84.13%</b>	<b>Gini Index</b>	
<b>92.29%</b>	<b>88.51%</b>	<b>89.61%</b>	<b>AGF</b>	

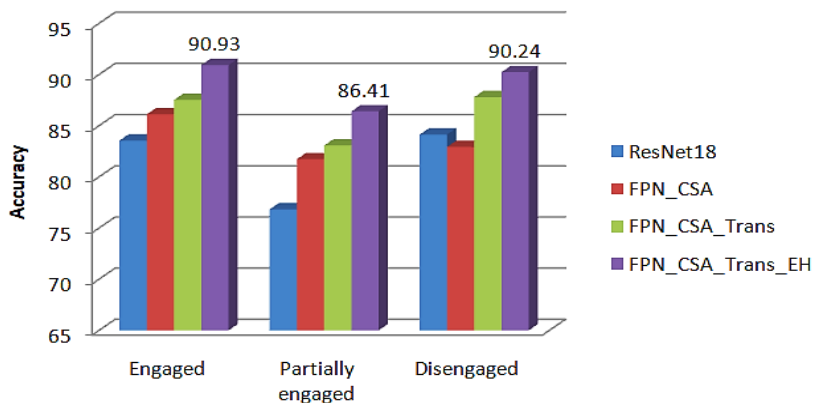


Fig. 9. Class-wise accuracy comparison of ResNet-18 and FPN\_CSA\_EH Trans on the WACV dataset

Class-wise AUC comparison between ResNet-18 and FPN\_CSA\_Trans\_EH. In terms of AUC, FPN\_CSA\_Trans\_EH outperforms ResNet-18 by 2.03% for the disengaged class, 6.95% for the partially engaged class, and 2.08% for the engaged class.

Class-wise Gini Index comparison between ResNet-18 and FPN\_CSA\_Trans\_EH. In terms of the Gini Index, FPN\_CSA\_Trans\_EH outperforms ResNet-18 by 4.06% for the disengaged class, 13.9% for the partially engaged class, and 4.16% for the engaged class.

Class-wise AGF comparison between ResNet-18 and FPN\_CSA\_Trans\_EH. In terms of AGF, FPN\_CSA\_Trans\_EH outperforms ResNet-18 by 2.53% for the disengaged class, 5.2% for the partially engaged class, and 3.8% for the engaged class.

**5.2. DAiSEE Dataset.** A graphical comparison of accuracy between the proposed method and existing methods on the DAiSEE dataset is provided in Table 3. In terms of accuracy, FPN\_CSA\_Trans\_EH outperforms FPN\_CSA\_Trans by 2.19%, FPN\_CSA by 4.55%, ResNet-18 [28] by 4.38%, Neural Turing Machine [32] by 9.72%, DERN [31] by 11.02%, DFSTN [30] by 12.22%, C3D+LSTM [29] by 14.42%, 13D [27] by 18.62%, and C3D [25] by 22.92%.

Table 3. Accuracy comparison of different methods on the DAiSEE dataset

DAiSEE	Accuracy(%)
C3D [25]	48.1
I3D [27]	52.4
C3D + LSTM [29]	56.6
DFSTN [30]	58.8
DERN [31]	60
Neural Turing Machine [32]	61.3
ResNet-18[28]	66.64
FPN_CSA	66.47
FPN_CSA_Trans	68.83
<b>FPN_CSA_Trans_EH</b>	<b>71.02</b>

Figure 10 illustrates the class-wise accuracy comparison between ResNet-18 and FPN\_CSA\_Trans\_EH (Table 4). For the boredom class, ResNet-18 and FPN\_CSA\_Trans\_EH achieve similar accuracy. FPN\_CSA\_Trans\_EH outperforms ResNet-18 by 14.28% for the confusion class, 3.51% for the frustration class, and 4.3% for the engagement class.

Table 4. Comparison of metrics for ResNet-18 and FPN\_CSA

Engagement	Frustration	Confusion	Boredom	Classes	
70.76%	61.45%	80.95%	75%	Accuracy	ResNet-18 [28]
70.07%	69.69%	88.09%	86.99%	AUC	
40.14%	39.38%	76.18%	73.98%	Gini Index	
71.32%	66.31%	82.98%	63.59%	AGF	
71.86%	60.09%	82.14%	50%	Accuracy	FPN_CSA
70.21%	69.29%	88.51%	87.19%	AUC	
40.42%	38.58%	77.02%	74.38%	Gini Index	
71.83%	65.5%	82.98%	70.65%	AGF	FPN_CSA_Trans
72.72%	63.37%	86.9%	75%	Accuracy	
71.67%	71.26%	91.33%	87.21%	AUC	
43.34%	42.53%	82.66%	74.43%	Gini Index	
73.03%	68.03%	86.49%	71.86%	AGF	FPN_CSA_Trans_EH
<b>75.06%</b>	<b>64.96%</b>	<b>95.23%</b>	75%	Accuracy	
<b>73.61%</b>	<b>72.83%</b>	<b>95.73%</b>	<b>87.24%</b>	AUC	
<b>73.61%</b>	<b>45.67%</b>	<b>91.47%</b>	<b>74.49%</b>	Gini Index	
<b>75.1%</b>	<b>69.5%</b>	<b>90.84%</b>	<b>73.13%</b>	AGF	

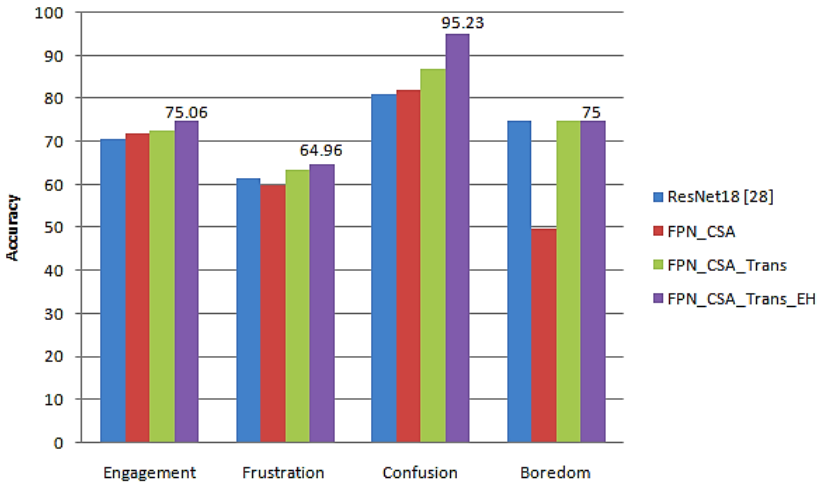


Fig. 10. Class-wise accuracy comparison of ResNet-18 and FPN\_CSA on the DAiSEE dataset

Class-wise AUC comparison between ResNet-18 and FPN\_CSA\_Trans\_EH. In terms of AUC, FPN\_CSA\_Trans\_EH outperforms ResNet-18 by 0.25% for the boredom class, 7.64% for the

confusion class, 3.14% for the frustration class, and 3.54% for the engagement class.

Class-wise Gini Index comparison between ResNet-18 and FPN\_CSA\_Trans\_EH. In terms of the Gini Index, FPN\_CSA\_Trans\_EH outperforms ResNet-18 by 0.51% for the boredom class, 15.29% for the confusion class, 6.29% for the frustration class, and 2.29% for the engagement class.

Class-wise AGF comparison between ResNet-18 and FPN\_CSA\_Trans\_EH. In terms of AGF, FPN\_CSA\_Trans\_EH outperforms ResNet-18 by 9.54% for the boredom class, 7.86% for the confusion class, 3.19% for the frustration class, and 3.78% for the engagement class.

The performance of FPN\_CSA\_Trans\_EH on the DAiSEE and WACV datasets underscores the model's strengths and identifies areas for improvement. While the model achieves exceptional performance on the WACV dataset, the challenges with the DAiSEE dataset offer valuable insights for further improvement.

**5. Conclusion and Future Work.** This study introduces FPN\_CSA\_Trans\_EH, a novel framework for real-time identification of student engagement in educational environments. The proposed approach integrates Channel-Spatial Attention (CSA) with Transformer-enhanced Feature Pyramid Networks (FPN), offering a robust method for automatically detecting patterns of student engagement in visual data streams. By combining attention mechanisms with hierarchical feature representation, FPN\_CSA\_Trans\_EH efficiently captures spatial and contextual information, enabling accurate determination of student engagement levels. The Transformer architecture enhances the model's ability to recognise long-range dependencies and semantic relationships within input sequences. The proposed framework outperforms baseline methods on the WACV dataset, demonstrating its potential for practical applications in educational settings. Future research will explore multiple avenues for enhancement and expansion. Initial efforts will focus on improving the model's efficiency and scalability to handle larger datasets and real-world implementation scenarios. Additionally, incorporating multimodal data sources, such as text and audio, could enhance the model's understanding of student interactions and behaviour. Finally, field tests and longitudinal studies will be conducted to evaluate the model's effectiveness in real educational settings and its impact on teaching and learning outcomes. Our goal is to advance student engagement detection techniques and contribute to the development of inclusive and effective educational technologies. Future work could extend this model by integrating audio modalities alongside visual features to predict student engagement in online settings.



## References

1. Marks H.M. Student engagement in instructional activity: Patterns in the elementary, middle, and high school years. *American Educational Research Journal*. 2000. vol. 37. pp. 153–184. DOI: 10.3102/00028312037001153.
2. Nomura K., Iwata M., Augereau O., Kise K. Estimation of student's engagement based on the posture. *Proceedings of the ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the ACM International Symposium on Wearable Computers*. ACM, 2019. pp. 164–167. DOI: 10.1145/3341162.3343767.
3. Kaur A., Mustafa A., Mehta L., Dhall A. Prediction and Localization of Student Engagement in the Wild. *Digital Image Computing: Techniques and Applications (DICTA)*. Australia, Canberra: ACT, 2018. pp. 1–8. DOI: 10.1109/DICTA.2018.8615851.
4. Hatori Y., Nakajima T., Watabe S. Body Posture Analysis for the Classification of Classroom Scenes. *Interdisciplinary Information Sciences*. 2022. vol. 28(1). pp. 55–62. DOI: 10.4036/iis.2022.a.05.
5. Liu Y., Chen J., Zhang M., Rao C. Student engagement study based on multi-cue detection and recognition in an intelligent learning environment. *Multimedia Tools Appl*. 2018. vol. 77(21). pp. 28749–28775.
6. Sharma P., Joshi S., Gautam S., Maharjan S., Khanal S.R., Reis M.C., Barroso J., de Jesus Filipe V.M. Student Engagement Detection Using Emotion Analysis, Eye Tracking and Head Movement with Machine Learning. *Technology and Innovation in Learning, Teaching and Education. TECH-EDU 2022. Communications in Computer and Information Science*. vol. 1720. pp. 52–68. DOI: 10.1007/978-3-031-22918-3\_5.
7. Bourel F., Chibelushi C. Low A. Recognition of Facial Expressions in the Presence of Occlusion. 2001. vol. 1. DOI: 10.5244/C.15.23.
8. Mao X., Xue Y., Li Z., Huang K., Lv S. Robust facial expression recognition based on RPCA and AdaBoost. *10th Workshop on Image Analysis for Multimedia Interactive Services*. 2009. pp. 113–116. DOI: 10.1109/WIAMIS.2009.5031445.
9. Jiang B., Jia Kb. Research of Robust Facial Expression Recognition under Facial Occlusion Condition. *Active Media Technology*, 2011. pp. 92–100. DOI: 10.1007/978-3-642-23620-4\_13.
10. Hammal Z., Arguin M., Gosselin F. Comparing a novel model based on the transferable belief model with humans during the recognition of partially occluded facial expressions. *Journal of vision*. 2009. vol. 9. DOI: 10.1167/9.2.22.
11. Zhang F., Zhang T., Mao Q., Xu C. Joint Pose and Expression Modeling for Facial Expression Recognition. *IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2018. pp. 3359–3368. DOI: 10.1109/CVPR.2018.00354.
12. Wang C., Wang S., Liang G. Identity- and Pose-Robust Facial Expression Recognition through Adversarial Feature Learning. 2019. pp. 238–246. DOI: 10.1145/3343031.3350872.
13. Moubayed A., Injadat M., Shami A., Lutfiyya H. Student Engagement Level in e-Learning Environment: Clustering Using K-means. *American Journal of Distance Education*. 2020. vol. 34(2). pp. 137–156. DOI: 10.1080/08923647.2020.1696140.
14. Gupta S., Kumar P., Tekchandani R. Facial emotion recognition based real-time learner engagement detection system in online learning context using deep learning models. *Multimedia Tools and Applications*. 2022. vol. 82. pp. 11365–11394. DOI: 10.1007/s11042-022-13558-9.
15. Bhardwaj P., Gupta P., Panwar H., Siddiqui M.K., Morales-Menendez R., Bhaik A. Application of Deep Learning on Student Engagement in e-learning environments. *Computers & Electrical Engineering*. 2021. vol. 93. DOI: 10.1016/j.compeleceng.2021.107277.

16. Fakhar S., Baber J., Bazai S., Marjan S., Jasiński M., Jasińska E., Chaudhry M.U., Leonowicz Z., Hussain S. Smart Classroom Monitoring Using Novel Real-Time Facial Expression Recognition System. *Applied Sciences*. 2022. vol. 12(23). DOI: 10.3390/app122312134.
17. Sümer Ö., Goldberg P., D'Mello S., Gerjets P., Trautwein U., Kasneci E., Multimodal Engagement Analysis From Facial Videos in the Classroom. *IEEE Transactions on Affective Computing*. 2023. vol. 14. no. 2. pp. 1012–1027. DOI: 10.1109/TAFFC.2021.3127692.
18. Psaltis A., Apostolakis K.C., Dimitropoulos K., Daras P. Multimodal Student Engagement Recognition in Prosocial Games. *IEEE Transactions on Games*. 2018. vol. 10. no. 3. pp. 292–303. DOI: 10.1109/TGIAIG.2017.2743341.
19. Mohamad Nezami O., Dras M., Hamey L., Richards D., Wan S., Paris C. Automatic Recognition of Student Engagement Using Deep Learning and Facial Expression. *Lecture Notes in Computer Science*. Springer, Cham. 2020. vol. 11908. pp. 273–289. DOI: 10.1007/978-3-030-46133-1\_17.
20. Yu H., Gupta A., Lee W., Arroyo I., Betke M., Allesio D., Murray T., Magee J., Woolf B.P. Measuring and integrating facial expressions and head pose as indicators of engagement and affect in tutoring systems *Adaptive Instructional Systems. Adaptation Strategies and Methods*. Cham Springer. 2021. pp. 219–233.
21. He K., Zhang X., Ren S., Sun J. Deep Residual Learning for Image Recognition. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. 2016. pp. 770–778. DOI: 10.1109/CVPR.2016.90.
22. Gao M., Song P., Wang F., Liu J., Mandelis A., Qi D. A Novel Deep Convolutional Neural Network Based on ResNet-18 and Transfer Learning for Detection of Wood Knot Defects. *Journal of Sensors*. 2021. pp. 1–16. DOI: 10.1155/2021/4428964.
23. Lin T.-Y., Dollár P., Girshick R., He K., Hariharan B., Belongie S. Feature Pyramid Networks for Object Detection. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. 2017. pp. 936–944. DOI: 10.1109/CVPR.2017.106.
24. Woo S., Park J., Lee J.Y., Kweon I.S. CBAM: Convolutional Block Attention Module. *Computer Vision – ECCV 2018. Lecture Notes in Computer Science*. Springer, Cham. 2018. vol. 11211. pp. 3–19. DOI: 10.1007/978-3-030-01234-2\_1.
25. Gupta A., DCunha A., Awasthi K., Balasubramanian V. DAiSEE: Towards User Engagement Recognition in the Wild. 2016. arXiv preprint: arXiv:1609.01885.
26. Islam M., Hossain E. Foreign Exchange Currency Rate Prediction using a GRU-LSTM Hybrid Network. *Soft Computing Letters*. 2021. vol. 3. DOI: 10.1016/j.socl.2020.100009.
27. Zhang H., Xiao X., Huang T., Liu S., Xia Y., Li J. An Novel End-to-end Network for Automatic Student Engagement Recognition. 2019 IEEE 9th International Conference on Electronics Information and Emergency Communication (ICEIEC). 2019. pp. 342–345. DOI: 10.1109/ICEIEC.2019.8784507.
28. Batra S., Wang S., Nag A., Brodeur P., Checkley M., Klinkert A., Dev S. DMCNet: Diversified Model Combination Network for Understanding Engagement from Video Screenshots. 2022. arXiv preprint: arXiv: 2204.06454. DOI: 10.48550/arXiv.2204.06454.
29. Abedi A., Khan S.S. Improving state-of-the-art in Detecting Student Engagement with Resnet and TCN Hybrid Network. 18th Conference on Robots and Vision (CRV). 2021. pp. 151–157. DOI: 10.1109/CRV52889.2021.00028.
30. Liao J., Liang Y., Pan J. Deep facial spatiotemporal network for engagement prediction in online learning. *Applied Intelligence*. 2021. vol. 51. pp. 6609–6621. DOI: 10.1007/s10489-020-02139-8.
31. Huang T., Mei Y., Zhang H., Liu S., Yang H. Fine-grained Engagement Recognition in Online Learning Environment. *IEEE 9th International Conference on Electronics*

Information and Emergency Communication (ICEIEC). 2019. pp. 338–341.  
DOI: 10.1109/ICEIEC.2019.8784559.

32. Ma X., Xu M., Dong Y., Sun Z. Automatic Student Engagement in Online Learning Environment Based on Neural Turing Machine. International Journal of Information and Education Technology. 2021. vol. 11(3). pp. 107–111.  
DOI: 10.18178/ijiet.2021.11.3.1497.

**Naveen A.** — Researcher, Department of computer science and engineering, Gitam University-Bengaluru Campus. Research interests: computer science, engineering. The number of publications — 0. a.naveen21@gmail.com; 207, Nagadenehalli Doddaballapur, Taluk, 561203, Bengaluru, Karnataka, India; office phone: +91(80)2809-8000.

**Jacob I. Jeena** — Professor, Associate professor, Department of computer science and engineering, Gitam University-Bengaluru Campus. Research interests: computer vision, deep learning. The number of publications — 53. ijacob@gitam.edu; 207, Nagadenehalli Doddaballapur, Taluk, 561203, Bengaluru, Karnataka, India; office phone: +91(80)2809-8000.

**Mandava Ajay Kumar** — Associate professor, Department of electrical, electronics and communication engineering, Gitam University-Bengaluru Campus. Research interests: computer vision, pattern recognition, signal and image processing. The number of publications — 46. amandava@gitam.edu; 207, Nagadenehalli Doddaballapur, Taluk, 561203, Bengaluru, Karnataka, India; office phone: +91(80)2809-8000.

А. НАВИН, И. ДЖЕЙКОБ, А. МАНДАВА  
**ОПРЕДЕЛЕНИЕ ВОВЛЕЧЕННОСТИ УЧАЩИХСЯ  
С ПОМОЩЬЮ СЕТЕЙ ПИРАМИДАЛЬНЫХ ПРИЗНАКОВ,  
УЛУЧШЕННЫХ ТРАНСФОРМЕРОМ, С КАНАЛЬНО-  
ПРОСТРАНСТВЕННЫМ ВНИМАНИЕМ**

*Навин А., Джейкоб И., Мандава А. Определение вовлеченности учащихся с помощью сетей пирамидальных признаков, улучшенных трансформером, с канално-пространственным вниманием.*

**Аннотация.** Одним из важнейших аспектов современных образовательных систем является определение вовлеченности учащихся, которое включает выявление того, насколько вовлечены, внимательны и активны учащиеся на занятиях в классе. Для преподавателей этот подход имеет важное значение, поскольку он дает представление об опыте обучения учащихся, позволяя адаптировать подходы в обучении и улучшать качество обучения. Традиционные методы оценки вовлеченности учащихся часто являются трудоемкими и субъективными. В этом исследовании предлагается новая система определения степени вовлеченности учащихся в реальном времени, которая использует сети пирамидальных признаков (FPN), улучшенные с помощью архитектуры Трансформера, с канално-пространственным вниманием (CSA), называемая BiusFPN\_CSA. Предлагаемый подход автоматически анализирует модели вовлеченности учащихся, такие как поза тела, зрительный контакт и положение головы, из визуальных потоков данных путем интеграции передовых методов глубокого обучения и компьютерного зрения. За счет интеграции механизма внимания CSA с возможностями иерархического представления признаков FPN, модель может точно определять уровни вовлеченности учащихся, улавливая контекстную и пространственную информацию во входных данных. Кроме того, благодаря внедрению архитектуры Трансформера, модель достигает лучшей общей производительности за счет эффективного учета долгосрочных зависимостей и семантических связей во входных последовательностях. Оценка с использованием набора данных WACV показывает, что предлагаемая модель превосходит базовые методы с точки зрения точности. В частности, вариант FPN\_CSA\_Trans\_EH предлагаемой модели превосходит FPN\_CSA на 3,28% и 4,98% соответственно. Эти результаты подчеркивают эффективность структуры BiusFPN\_CSA в определении вовлеченности учащихся в реальном времени, предлагая преподавателям ценный инструмент для повышения качества обучения, создания активной среды обучения и, в конечном итоге, улучшения результатов учащихся.

**Ключевые слова:** сеть пирамидальных признаков (FPN), канално-пространственное внимание (CSA), определение вовлеченности учащихся, трансформер.

### Литература

1. Marks H.M. Student engagement in instructional activity: Patterns in the elementary, middle, and high school years. *American Educational Research Journal*. 2000. vol. 37. pp. 153–184. DOI: 10.3102/00028312037001153.
2. Nomura K., Iwata M., Augereau O., Kise K. Estimation of student's engagement based on the posture. *Proceedings of the ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the ACM International Symposium on Wearable Computers*. ACM, 2019. pp. 164–167. DOI: 10.1145/3341162.3343767.

3. Kaur A., Mustafa A., Mehta L., Dhall A. Prediction and Localization of Student Engagement in the Wild. *Digital Image Computing: Techniques and Applications (DICTA)*. Australia, Canberra: ACT, 2018. pp. 1–8. DOI: 10.1109/DICTA.2018.8615851.
4. Hatori Y., Nakajima T., Watabe S. Body Posture Analysis for the Classification of Classroom Scenes. *Interdisciplinary Information Sciences*. 2022. vol. 28(1). pp. 55–62. DOI: 10.4036/iis.2022.a.05.
5. Liu Y., Chen J., Zhang M., Rao C. Student engagement study based on multi-cue detection and recognition in an intelligent learning environment. *Multimedia Tools Appl.* 2018. vol. 77(21). pp. 28749–28775.
6. Sharma P., Joshi S., Gautam S., Maharjan S., Khanal S.R., Reis M.C., Barroso J., de Jesus Filipe V.M. Student Engagement Detection Using Emotion Analysis, Eye Tracking and Head Movement with Machine Learning. *Technology and Innovation in Learning, Teaching and Education. TECH-EDU 2022. Communications in Computer and Information Science*. vol. 1720. pp. 52–68. DOI: 10.1007/978-3-031-22918-3\_5.
7. Bourel F., Chibelushi C., Low A. Recognition of Facial Expressions in the Presence of Occlusion. 2001. vol. 1. DOI: 10.5244/C.15.23.
8. Mao X., Xue Y., Li Z., Huang K., Lv S. Robust facial expression recognition based on RPCA and AdaBoost. *10th Workshop on Image Analysis for Multimedia Interactive Services*. 2009. pp. 113–116. DOI: 10.1109/WIAMIS.2009.5031445.
9. Jiang B., Jia K. Research of Robust Facial Expression Recognition under Facial Occlusion Condition. *Active Media Technology*, 2011. pp. 92–100. DOI: 10.1007/978-3-642-23620-4\_13.
10. Hammal Z., Arguin M., Gosselin F. Comparing a novel model based on the transferable belief model with humans during the recognition of partially occluded facial expressions. *Journal of vision*. 2009. vol. 9. DOI: 10.1167/9.2.22.
11. Zhang F., Zhang T., Mao Q., Xu C. Joint Pose and Expression Modeling for Facial Expression Recognition. *IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2018. pp. 3359–3368. DOI: 10.1109/CVPR.2018.00354.
12. Wang C., Wang S., Liang G. Identity- and Pose-Robust Facial Expression Recognition through Adversarial Feature Learning. 2019. pp. 238–246. DOI: 10.1145/3343031.3350872.
13. Moubayed A., Injadat M., Shami A., Lutfiyya H. Student Engagement Level in e-Learning Environment: Clustering Using K-means. *American Journal of Distance Education*. 2020. vol. 34(2). pp. 137–156. DOI: 10.1080/08923647.2020.1696140.
14. Gupta S., Kumar P., Tekchandani R. Facial emotion recognition based real-time learner engagement detection system in online learning context using deep learning models. *Multimedia Tools and Applications*. 2022. vol. 82. pp. 11365–11394. DOI: 10.1007/s11042-022-13558-9.
15. Bhardwaj P., Gupta P., Panwar H., Siddiqui M.K., Morales-Menendez R., Bhalk A. Application of Deep Learning on Student Engagement in e-learning environments. *Computers & Electrical Engineering*. 2021. vol. 93. DOI: 10.1016/j.compeleceng.2021.107277.
16. Fakhar S., Baber J., Bazai S., Marjan S., Jasiński M., Jasińska E., Chaudhry M.U., Leonowicz Z., Hussain S. Smart Classroom Monitoring Using Novel Real-Time Facial Expression Recognition System. *Applied Sciences*. 2022. vol. 12(23). DOI: 10.3390/app122312134.
17. Sümer Ö., Goldberg P., D’Mello S., Gerjets P., Trautwein U., Kasneci E., Multimodal Engagement Analysis From Facial Videos in the Classroom. *IEEE Transactions on Affective Computing*. 2023. vol. 14. no. 2. pp. 1012–1027. DOI: 10.1109/TAFFC.2021.3127692.

18. Psaltis A., Apostolakis K.C., Dimitropoulos K., Daras P. Multimodal Student Engagement Recognition in Prosocial Games. *IEEE Transactions on Games*. 2018. vol. 10. no. 3. pp. 292–303. DOI: 10.1109/TGIAIG.2017.2743341.
19. Mohamad Nezami O., Dras M., Hamey L., Richards D., Wan S., Paris C. Automatic Recognition of Student Engagement Using Deep Learning and Facial Expression. *Lecture Notes in Computer Science*. Springer, Cham. 2020. vol. 11908. pp. 273–289. DOI: 10.1007/978-3-030-46133-1\_17.
20. Yu H., Gupta A., Lee W., Arroyo I., Betke M., Allesio D., Murray T., Magee J., Woolf B.P. Measuring and integrating facial expressions and head pose as indicators of engagement and affect in tutoring systems *Adaptive Instructional Systems. Adaptation Strategies and Methods*. Cham Springer. 2021. pp. 219–233.
21. He K., Zhang X., Ren S., Sun J. Deep Residual Learning for Image Recognition. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. 2016. pp. 770–778. DOI: 10.1109/CVPR.2016.90.
22. Gao M., Song P., Wang F., Liu J., Mandelis A., Qi D. A Novel Deep Convolutional Neural Network Based on ResNet-18 and Transfer Learning for Detection of Wood Knot Defects. *Journal of Sensors*. 2021. pp. 1–16. DOI: 10.1155/2021/4428964.
23. Lin T.-Y., Dollár P., Girshick R., He K., Hariharan B., Belongie S. Feature Pyramid Networks for Object Detection. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. 2017. pp. 936–944. DOI: 10.1109/CVPR.2017.106.
24. Woo S., Park J., Lee J.Y., Kweon I.S. CBAM: Convolutional Block Attention Module. *Computer Vision – ECCV 2018. Lecture Notes in Computer Science*. Springer, Cham. 2018. vol. 11211. pp. 3–19. DOI: 10.1007/978-3-030-01234-2\_1.
25. Gupta A., DCunha A., Awasthi K., Balasubramanian V. DAiSEE: Towards User Engagement Recognition in the Wild. 2016. arXiv preprint: arXiv:1609.01885.
26. Islam M., Hossain E. Foreign Exchange Currency Rate Prediction using a GRU-LSTM Hybrid Network. *Soft Computing Letters*. 2021. vol. 3. DOI: 10.1016/j.socl.2020.100009.
27. Zhang H., Xiao X., Huang T., Liu S., Xia Y., Li J. An Novel End-to-end Network for Automatic Student Engagement Recognition. 2019 IEEE 9th International Conference on Electronics Information and Emergency Communication (ICEIEC). 2019. pp. 342–345. DOI: 10.1109/ICEIEC.2019.8784507.
28. Batra S., Wang S., Nag A., Brodeur P., Checkley M., Klinkert A., Dev S. DMCNet: Diversified Model Combination Network for Understanding Engagement from Video Screengrabs. 2022. arXiv preprint: arXiv: 2204.06454. DOI: 10.48550/arXiv.2204.06454.
29. Abedi A., Khan S.S. Improving state-of-the-art in Detecting Student Engagement with Resnet and TCN Hybrid Network. 18th Conference on Robots and Vision (CRV). 2021. pp. 151–157. DOI: 10.1109/CRV52889.2021.00028.
30. Liao J., Liang Y., Pan J. Deep facial spatiotemporal network for engagement prediction in online learning. *Applied Intelligence*. 2021. vol. 51. pp. 6609–6621. DOI: 10.1007/s10489-020-02139-8.
31. Huang T., Mei Y., Zhang H., Liu S., Yang H. Fine-grained Engagement Recognition in Online Learning Environment. *IEEE 9th International Conference on Electronics Information and Emergency Communication (ICEIEC)*. 2019. pp. 338–341. DOI: 10.1109/ICEIEC.2019.8784559.
32. Ma X., Xu M., Dong Y., Sun Z. Automatic Student Engagement in Online Learning Environment Based on Neural Turing Machine. *International Journal of Information and Education Technology*. 2021. vol. 11(3). pp. 107–111. DOI: 10.18178/ijiet.2021.11.3.1497.

**Навин А.** — научный сотрудник, факультет компьютерных наук и инженерии, Университет Гитам – кампус в Бангалоре. Область научных интересов: компьютерные науки, инженерия. Число научных публикаций — 0. a.naveen21@gmail.com; Нагаденехалли Доддабаллапур, Талук, 207, 561203, Бангалор, Карнатака, Индия; р.т.: +91(80)2809-8000.

**Джейкоб И. Джина** — профессор, доцент, факультет компьютерных наук и инженерии, Университет Гитам – кампус в Бангалоре. Область научных интересов: компьютерное зрение, глубокое обучение. Число научных публикаций — 53. ijacob@gitam.edu; Нагаденехалли Доддабаллапур, Талук, 207, 561203, Бангалор, Карнатака, Индия; р.т.: +91(80)2809-8000.

**Мандава Аджай Кумар** — доцент, кафедра электротехники, электроники и средств связи, Университет Гитам – кампус в Бангалоре. Область научных интересов: компьютерное зрение, распознавание образов, обработка сигналов и изображений. Число научных публикаций — 46. amandava@gitam.edu; Нагаденехалли Доддабаллапур, Талук, 207, 561203, Бангалор, Карнатака, Индия; р.т.: +91(80)2809-8000.

А.В. ЧУВАКОВ, Р.О. БОРЯЕВ  
**СИСТЕМА ПОДДЕРЖКИ ПРИНЯТИЯ ФАКТОРИНГОВЫХ  
РЕШЕНИЙ НА ОСНОВЕ ОПТИМИЗИРОВАННЫХ  
КВАНТОВЫХ АЛГОРИТМОВ QMC**

*Чуваков А.В., Боряев Р.О. Система поддержки принятия факторинговых решений на основе оптимизированных квантовых алгоритмов QMC.*

**Аннотация.** Непрерывный рост финансовых рынков диктует необходимость для его участников искать новые подходы к финансовому анализу для получения конкурентных преимуществ, в том числе за счет использования новых подходов в области вычислений. Квантовые вычисления могут быть использованы в качестве инструмента по получению данных преимуществ перед конкурентами. В частности моделирование Монте-Карло применяется широко в управлении финансовыми рисками, в то же время, требует значительных вычислительных ресурсов из-за использования большого количества сценариев, необходимых для получения более точного результата. Для оптимизации данного подхода применяются алгоритмы квантовой оценки амплитуды, которые ускоряют данный процесс, если использовать предварительно вычисленные распределения вероятностей для инициализации входных квантовых состояний. Но при отсутствии данных распределений в имеющихся подходах по данной тематике они генерируются численно с использованием классических вычислений, что полностью нивелирует преимущество квантового подхода. В данной статье предлагается решение указанной проблемы путём использования квантовых вычислений, в том числе для генерации распределений вероятностей. Была рассмотрена реализация квантовых схем для моделирования эволюции факторов риска во времени для движения капитала, процентных ставок и кредитных рисков, а также представлено объединение этих моделей с алгоритмами квантовой оценки амплитуды в качестве примера использования полученных алгоритмов для управления кредитными рисками. В завершении статьи проанализирована возможность использования полученных схем в финансовом анализе.

**Ключевые слова:** квантовые вычисления, квантовый метод Монте-Карло, квантовая оценка амплитуды, оценка финансовых рисков.

**1. Введение.** Проблематика ускорения решений численных задач в финансах неоднократно поднималась в современных исследованиях на тему практического применения квантовых вычислений [1 – 7]. В сфере управления рисками возможно применение квантовых алгоритмов для расчёта стоимости-риска и ожидаемого дефолта, ценообразования опционов и обеспеченных долговых обязательств, корректировки оценки кредита. В настоящее время квантовые вычисления не используются в производстве, но теоретические достижения в этой области необходимы финансовым институтам для подготовки к переходу на новый метод. В управлении финансовыми рисками, которое требует значительного количества вычислительных ресурсов из-за большого количества сценариев, необходимых для получения точного результата, часто используют моделирование Монте-Карло. Алгоритмы квантовой оценки



амплитуды (QAE) могут значительно ускорить этот процесс, но в то же время существует возможность оптимизировать и эти алгоритмы, если использовать предварительно вычисленные распределения вероятностей для инициализации входных квантовых состояний, так как если таких распределений нет, их нужно генерировать численно, а это нивелирует преимущество квантового подхода. В данной статье предлагается решение обозначенной проблемы оптимизации алгоритмов QAE путём использования квантовых вычислений так же и для генерации распределений вероятностей. В ходе статьи предлагается рассмотреть создание квантовых схем для моделирования эволюции факторов риска во времени для движения капитала, процентных ставок и кредитных рисков (QMC). Далее, путем объединения моделей с алгоритмами квантовой оценки амплитуды, рассматривается возможность их использования для управления кредитными рисками.

### 1.1. Моделирование на основе метода Монте-Карло.

Финансовые риски связаны с возможными потерями из-за будущих событий, таких как изменение цен на акции, процентных ставок или валютных курсов. Также к ним относится дефолт долгового инструмента или контрагента. Риск можно оценить статистически, опираясь на свойства факторов риска. Для этого генерируется распределение вероятности  $P(X_t)$  для вектора факторов риска  $X_t$  в будущем времени  $t$ , затем определить меру риска  $F(X_t)$ , которая будет зависеть от  $X_t$ , и оценить значение меры риска с помощью статистических методов.

Например, рассмотрим значение  $V_t$  портфеля, состоящего из  $q$  акций одного типа, с его ценой  $S_0$  при  $t=0$ . Здесь  $X_t$  является одномерной величиной, и единственным элементом является  $S_t$ . Цена и волатильность акции могут использоваться в качестве входных данных для моделирования эволюции  $S_t$  и генерации распределения  $P(S_t)$ . Предположим, мера риска определяется как ожидаемое значение портфеля во время  $t$ ,  $E(V_t) = qE(S_t)$  (оценка Value-at-Risk [8]). Тогда можно оценить  $E(V_t)$  либо аналитическими, либо численными методами. Вероятностная природа эволюции фактора риска требует повторной случайной выборки с последующей статистической оценкой – этот тип численного подхода называется моделированием Монте-Карло (МК). Он заключается в многократной случайной выборке с последующей статистической оценкой. Расчёты управления финансовыми рисками с помощью МК требуют от 10 000 до 1 000 000 экспериментов для достижения необходимой точности. Предположим, что величина  $X$  принимает значения 1 и 0 с вероятностями  $p$  и  $1-p$

соответственно. В моделировании МК эксперимент повторяется  $N$  раз, получая  $N_1$  наблюдений результата  $X=1$  и  $N_0$  результата  $X=0$ , при  $N_1+N_0=N$ . Это биномиальное распределение с ожидаемым значением числа появлений  $X=1$ , равным  $\bar{N}_1 = pN$ , а стандартное отклонение равно  $\delta N_1 = [p(1-p)N]^{1/2}$ . Поскольку  $N_1$  известно, но не  $p$ , можно инвертировать выражения для оценки  $p$  на основе наблюдаемых результатов:

$$p \approx \bar{p} = \frac{N_1}{N}, \quad (1)$$

$$\delta p \approx \frac{\delta N_1}{N} = \left[ \frac{\bar{p}(1-\bar{p})}{N} \right]^{1/2} \propto \left( \frac{1}{N} \right)^{1/2}. \quad (2)$$

Из второго уравнения следует, что шум выборки обратно пропорционален квадратному корню из  $N$ .

**1.2. Алгоритмы оценки квантовой амплитуды.** Уравнение (2) подразумевает, что для увеличения точности на десятичную цифру классические алгоритмы требуют в 100 раз больше экспериментов, поэтому типичные значения  $N$  находятся в диапазоне 10000–1000000. В то же время QAE могут достигать квадратичного ускорения по сравнению с классическими алгоритмами [9]. Для примера из раздела 1.1 можно оценить вероятность  $p$  с помощью квантового компьютера, кодируя распределение вероятностей случайной величины в квантовом состоянии кубита:

$$|\psi\rangle = \sqrt{1-p}|0\rangle + \sqrt{p}|1\rangle = \cos(\theta/2)|0\rangle + \sin(\theta/2)|1\rangle \quad (3)$$

так, что состояние  $|1\rangle$  измеряется с вероятностью  $p$ . Повторение квантового эксперимента и измерение значения каждый раз не обходят классическое ограничение точности, которая пропорциональна квадратному корню из числа измерений. Вместо этого алгоритмы QAE используют квантовую интерференцию для ускорения вычислений. Входной кубит инициализируется в суперпозиции состояний, а затем применяется алгоритм QAE, основанный на фазовой отдаче и квантовой оценке амплитуды. Вероятность  $p$  (угол  $\theta$ ) запечатлется как фаза  $\phi = \pm k\theta$  на выходных кубитах. Далее следует обратное квантовое преобразование Фурье, которое преобразует фазы

в двоичное число. Это позволяет считать угол  $\theta$  и значение  $p$  напрямую. При таком подходе оценка  $p$  и его шум равны:

$$p \approx \bar{p} = \sin^2\left(\frac{\theta}{2}\right), \quad (4)$$

$$\delta p \approx \sin\theta \frac{\pi}{2N} \propto \frac{1}{N}, \quad (5)$$

где  $N=2^n$  – число возможных результатов, и  $\delta p$  обратно пропорционален  $N$ . Для квантовой оценки амплитуды, в области финансового риска, рассматривают схемы с заранее известными распределениями вероятностей случайных величин  $P(X_i)$  [10]. Однако, если эти распределения нельзя точно закодировать в квантовом состоянии, то алгоритмы, включающие интеграцию классического варианта МК для подготовки состояния, не обеспечивают квадратичного ускорения. Если распределение  $P(X_i)$  доступно, то существуют методы эффективной загрузки его в квантовый регистр и оптимизации схем. Такой подход помогает упростить подготовку состояния и уменьшить сложность схемы [11, 12]. Включение генерации сценариев в квантовую схему позволяет обойти проблему медленной сходимости классических алгоритмов при вычислении  $P(X_i)$ . В этом случае точное распределение случайной величины можно представить как суперпозицию квантовых состояний. Таким образом, уменьшается зависимость от передачи данных между классическими и квантовыми системами. В данной статье рассматривается квантовая реализация стохастических моделей риска для генерации сценариев факторов риска – симуляция QMC. Далее полученные распределения объединяются с вентилями QAE. В зависимости от сложности стохастических дифференциальных уравнений, описывающих эволюцию факторов риска, симуляции МК дополнительно оптимизируются с помощью квантово-ускоренных многоуровневых методов МК.

**1.3. Обзор схем QAE.** Для оценки статистической меры  $F(X_i) \in [F_{min}, F_{max}]$  фактора риска  $X_i$  схема QAE обычно состоит из  $m$  кубитов «фактора риска» ( $rf$ ) для моделирования распределения  $X_i$  (начальное состояние обозначается как  $|0\rangle_{rf}^{\otimes m}$ ), одного кубита «меры риска» ( $rm$ ) для кодирования нормализованного значения меры риска  $f(X_i) \in [0, 1]$  в угле  $\theta \in [0, \pi]$  (начальное состояние обозначается как  $|0\rangle_{rm}$ ,

п выходных кубитов для впечатывания кратных  $\theta$  в их фазы (начальное состояние обозначается как  $|0\rangle_{out}^{\otimes n}$ ). Без потери общности, в следующем примере выбирается  $f(X_i)=p$ , т.е. мера риска будет вероятностью результата, зависящего от  $X_i$ . Общая структура квантовой схемы QMC/QAE на рисунке 1.

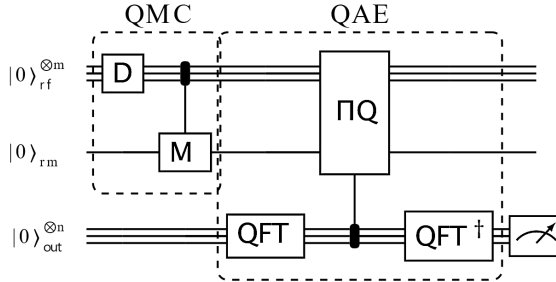


Рис. 1. Общая структура квантовой схемы QMC/QAE

$D$  – вентиль, который генерирует входное распределение с использованием  $m$  кубитов «фактора риска»,  $M$  – управляемый вентиль, который кодирует меру риска в угол  $\theta$  кубита меры риска,  $PQ$  – повторное применение управляемого вентиля  $Q$  для запечатления  $\theta$  в фазах  $n$  выходных кубитов, а  $QFT/QFT^\dagger$  квантовое преобразование Фурье и его обратная величина для измерения фазы выходных кубитов с помехами. Эти компоненты схемы описаны ниже более подробно (таблица 1).

**1.3.1. Вентиль  $D$ : подготовка распределения  $P(X_j)$ .** С помощью  $m$  кубитов,  $|0\rangle_{rf}^{\otimes m}$ , можно смоделировать дискретное распределение вероятностей  $2^m$  возможных результатов, каждый с вероятностью  $|a_j|^2$ , с  $j \in \{0, 1, \dots, 2^m - 1\}$ . Вентиль  $D$  используется для загрузки распределения  $X_i$  в кубиты «фактора риска» (рисунок 2) и выражается математически как:

$$|\psi\rangle_{rf} = D|0\rangle_{rf}^{\otimes m} = \sum_{j=0}^{2^m-1} a_j |b_{jm-1} \dots b_{j1} b_{j0}\rangle_{rf} = \sum_{j=0}^{2^m-1} a_j |j\rangle_{rf}, \quad (6)$$

где  $j$  – целое число в десятичной системе счисления, представляющее двоичное число  $b_{jm-1} \dots b_{j1} b_{j0}$ , а  $b_{jl} \in \{0, 1\}$  –  $l$ -я цифра  $j$ -го состояния.

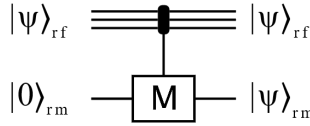

 Рис. 2. Вентиль  $D$ 

Если распределение предварительно вычислено, его можно загрузить, используя шаблонную схему квантового преобразования [12], квантовые генеративные состязательные сети [13], разложение Фурье [14] или квантовую обработку сигналов QSP [15]. В данной работе предлагается генерировать распределения с помощью квантовых вентилей, которые реализуют стохастические модели для эволюции фактора риска.

Таблица 1. Обозначение состояний

Состояние	Описание
$ b_{m-1} \dots b_1 b_0\rangle$	$m$ кубитов в двоичном виде, например $ 101\rangle$
$ j\rangle,  x\rangle,  z\rangle$	кубиты как целое число, например $ 5\rangle$
$ \psi\rangle$	суперпозиция состояний
$ \psi\rangle_{in}$	входной кубит(ы)
$ \psi\rangle_{out}$	выходной кубит(ы)
$ \psi\rangle_{rf}$	кубит(ы) «фактора риска» (вход)
$ \psi\rangle_{rm}$	кубит «меры риска» (вход)
$ \psi\rangle_{st}$	кубит(ы) «состояния»
$ \psi\rangle_c$	кубит(ы) «количества»
$ \psi\rangle_{anc}$	кубит(ы) «вспомогательный»

**1.3.2. Вентиль  $M$ : расчет меры риска  $f(X_i)=p$ .** С распределением  $P(X_i)$ , закодированным в состоянии кубитов «фактора риска»,  $|\psi\rangle_{rf}$ , кодируются меры риска  $f(X_i)$  – например, вероятность,  $p=P(X_i < K)$ , где  $K \in [X_{min}, X_{max}]$  – в состоянии кубита «меры риска»,  $|0\rangle_{rm}$ . Значение  $p$  может быть зафиксировано в угле  $\theta$  кубита с помощью управляемого вентиля  $M$ , который считывает распределение  $|\psi\rangle_{rf}$  и кодирует  $p \in [0, 1] \rightarrow \theta \in [0, \pi]$  в  $|0\rangle_{rm}$  (рисунок 3).


 Рис. 3. Вентиль  $M$ 

$$\begin{aligned}
 |\psi\rangle_{in} = M|\psi\rangle_{rf}|0\rangle_{rm} = \sqrt{1-p}|\psi_0\rangle_{rf}|0\rangle_{rm} + \sqrt{p}|\psi_1\rangle_{rf}|1\rangle_{rm} = \\
 \sqrt{1-p}|\psi_0\rangle_{in} + \sqrt{p}|\psi_1\rangle_{in}, \quad (7)
 \end{aligned}$$

где  $\sqrt{1-p} = \cos \theta / 2$ ,  $\sqrt{p} = \sin \theta / 2$ , и, упростив обозначения,  $|\psi_0\rangle_{in} = |\psi_0\rangle_{rf}|0\rangle_{rm}$  и  $|\psi_1\rangle_{in} = |\psi_1\rangle_{rf}|1\rangle_{rm}$

### 1.3.3. Вентили $PQ$ и $QFT$ : оценка $p$

**1.3.3.1. Усиление амплитуды.** Следующая часть схемы – это управляемый  $Q$ -вентиль, оператор, основанный на алгоритме поиска Гровера [16].  $Q$  состоит из двух отражений,  $|\psi_0\rangle_{in} \rightarrow -|\psi_0\rangle_{in}$  и  $|\psi_1\rangle_{in} \rightarrow -|\psi_1\rangle_{in}$ :

$$Q = Q_\psi Q_{\psi_0} = (1 - 2|\psi\rangle_{in}\langle\psi|_{in})(1 - 2|\psi_0\rangle_{in}\langle\psi_0|_{in}), \quad (8)$$

которые усиливают/уменьшают амплитуды  $|\psi_0\rangle_{in}$  и  $|\psi_1\rangle_{in}$  в зависимости от значения  $p$ :

$$Q|\psi\rangle_{in} = (1 - 4p)\sqrt{1-p}|\psi_0\rangle_{in} + (3 - 4p)\sqrt{1-p}|\psi_1\rangle_{in}. \quad (9)$$

На рисунке 4 показаны множители  $(1 - 4p)^2$  и  $(3 - 4p)^2$ , которые увеличивают/уменьшают вероятности измерения  $|0\rangle_{rm}$  и  $|1\rangle_{rm}$  соответственно.

Ключевым свойством вентиля  $Q$  является то, что сохраняет состояние

$$|\psi^\pm\rangle_{in} = \frac{1}{\sqrt{2}}|\psi_1\rangle_{in} \pm i|\psi_0\rangle_{in} \quad (10)$$

не измененным после применения  $k$  раз, но вводит глобальную фазу:

$$Q^k |\psi_{\pm}\rangle_{in} = e^{\pm ik\theta} |\psi_{\pm}\rangle_{in}. \quad (11)$$

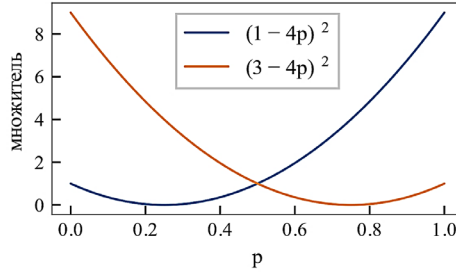


Рис. 4. Увеличение вероятностей состояний  $|\psi_0\rangle_{in}$  (синий) и  $|\psi_1\rangle_{in}$  (оранжевый), когда  $Q$  действует на  $\Psi$  (9)

**1.3.3.2. Фазовый откат.** Фазовый откат может использоваться при реализации управляемого вентиля  $Q$ , действующего на  $|\psi\rangle_{in}$ , который выражается как:

$$|\psi\rangle_{in} = -i \frac{1}{\sqrt{2}} \left( e^{i\theta/2} |\psi_+\rangle_{in} - e^{-i\theta/2} |\psi_-\rangle_{in} \right). \quad (12)$$

Целью этого управляющего вентиля являются выходные кубиты, после первого вращения их начального состояния,  $|0\dots 00\rangle_{out}$ , оператором «H» таким образом, что оно становится  $|+\dots+\rangle_{out}$ . Это эквивалентно квантовому преобразованию Фурье (рисунок 5) и математически выражается как:

$$QFT |0\rangle_{out}^{\otimes n} = \bigotimes_{l=0}^{n-1} H |0\rangle_{out}^l = |+\rangle_{out}^{\otimes n}. \quad (13)$$



Рис. 5 Квантовое преобразование Фурье

Далее используется фазовый откат для отпечатывания кратных углов  $\theta$  на их фазы. На основе уравнения 11 фаза выходного кубита  $|+\rangle'_{out} = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$  станет  $\frac{1}{\sqrt{2}}(|0\rangle + e^{ik\theta})$  для  $|\psi_+\rangle_{in}$ , и  $\frac{1}{\sqrt{2}}(|0\rangle + e^{-ik\theta})$  для  $|\psi_-\rangle_{in}$ , где устанавливается  $k=2^l$ . А именно,  $k=2^0$  для первого выходного кубита,  $k=2^1$  для второго, ..., и  $k=2^{n-1}$  для последнего ( $n$ -го). Повторное применение управляемого вентиля  $Q$  на рисунке 6.

Состояние выходных кубитов можно записать как:

$$\prod_{l=0}^{n-1} Q^{2^l} |\psi_+\rangle_{in} |+\rangle_{out}^{\otimes n} = |\psi\rangle_{in} \left\{ \bigotimes_{l=0}^{n-1} \left[ \frac{1}{\sqrt{2}} (|0\rangle + e^{\pm i 2^l \theta} |1\rangle) \right] \right\} = |\psi\rangle_{in} \left\{ \frac{1}{2^{n/2}} \sum_{x=0}^{2^n-1} e^{\pm i x \theta} |x\rangle \right\}, \quad (14)$$

где  $x = \sum_{l=0}^{n-1} 2^l b^l$  – двоичное число  $b_{n-1} \dots b_1 b_0$ , выраженное в десятичной системе счисления.

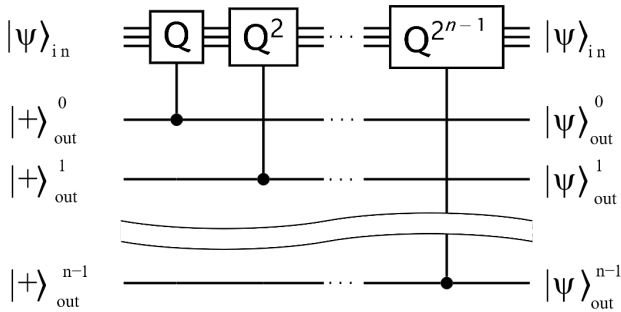


Рис. 6. Повторное применение управляемого вентиля  $Q$

**1.3.3.3. Интерференция.** Состояние  $\sum_{x=0}^{2^n-1} a_x |x\rangle$ , где  $a_x = \left(1/2^{n/2}\right) e^{\pm i x \theta}$  может быть преобразовано в стандартный базис



$\sum_{z=0}^{2^n-1} a_z |z\rangle$  с помощью обратного квантового преобразования Фурье, которое изображается вентилем  $QFT^\dagger$  (рисунок 7) и математически выражается:

$$|\psi\rangle_{out} = QFT^\dagger \frac{1}{2^{n/2}} \sum_{x=0}^{2^n-1} e^{\pm ix\theta} |x\rangle = \frac{1}{2^n} \sum_{z=0}^{2^n-1} \sum_{x=0}^{2^n-1} e^{ix(\pm\theta-2\pi z/2^n)} |z\rangle, \quad (15)$$

включая суперпозиции состояний  $|z\rangle$ , где  $z$  – целое число в десятичной системе счисления.

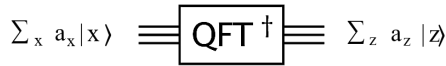


Рис. 7. Обратное квантовое преобразование Фурье

**1.3.3.4. Измерение.** Измерение состояния  $\sum_z a_z |z\rangle$  приведет к коллапсу суперпозиции и даст одно из возможных значений  $z$ ,  $z_0$ , (рисунок 8).

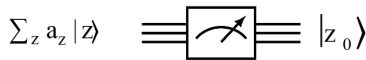


Рис. 8. Измерение состояния  $\sum_z a_z |z\rangle$

Если  $\theta$  имеет такое значение, что существует целое число  $z_0 \in [0, 2^{n-1}]$ , которое делает  $\pm\theta - 2\pi z/2^n$  кратным  $2\pi$ , а именно,  $z_0 = 2^n \theta / 2\pi$  или  $z_0 = 2^n (2\pi - \theta) / 2\pi$ , то уравнение (15) подразумевает, что измеренное состояние является одним из:

$$|z_0\rangle = |2^n \theta / 2\pi\rangle, \quad (16)$$

$$|z_0\rangle = |2^n (2\pi - \theta) / 2\pi\rangle. \quad (17)$$

Затем вычисляется вероятность  $p$ , заменив  $\theta$  в уравнении (3) на измеренное значение  $z_0$ :

$$p = \sin^2 \left( \frac{\theta}{2} \right) = \sin^2 \left( \frac{z_0}{2^n} \pi \right). \quad (18)$$

Если нет такого  $z_0$ , что  $\pm\theta - 2\pi z_0/2^n$  кратного  $2\pi$ , ближайшее целое число  $z_0$  измеряется с вероятностью  $|a_{z_0}|^2 \sim 20\%$  для  $\theta$  или  $2\pi - \theta$ , что дает общую вероятность приблизительно 40% для получения ближайшего значения  $p$ .

Точность, с которой оценивается  $\theta$ , составляет  $\delta\theta \approx 2\pi/2^n$ , и,  $p$  равна:

$$\delta p \approx \sin\theta \frac{\pi}{N} \propto \frac{1}{N}, \quad (19)$$

где  $N=2^n$  – общее количество возможных исходов.

Важно, что шум в QAE уменьшается пропорционально  $1/N$  (19), намного быстрее, чем  $1/N^{1/2}$  классических алгоритмов (2).

## 2. Квантовые вентили и схемы для генерации сценариев

**2.1. Факторы риска процентной ставки.** Эволюцию процентных ставок можно смоделировать с помощью краткосрочных моделей возврата к среднему значению, основная из которых – модель Васичека [17]:

$$dr_t = a(b - r_t)dt + \sigma dW_t, \quad (20)$$

где  $r_t$  – мгновенная процентная ставка в момент времени  $t$ ,  $b$  – долгосрочное среднее значение,  $a$  – скорость возврата,  $\sigma$  – волатильность, а  $W_t$  – винеровский процесс. Ожидаемое значение и дисперсия асимптотически постоянны при  $\delta t \rightarrow \infty$ , тогда как при конечном  $\delta t$  ожидаемое значение зависит от  $r_t$ :

$$E(r_{t+\delta t}) = r_t e^{-a\delta t} + b(1 - e^{-a\delta t}), \quad (21)$$

$$Var(r_{t+\delta t}) = \frac{\sigma^2}{2a} (1 - e^{-2a\delta t}). \quad (22)$$

Такие модели часто дискретизируются с помощью триномиальных деревьев (рисунок 9).

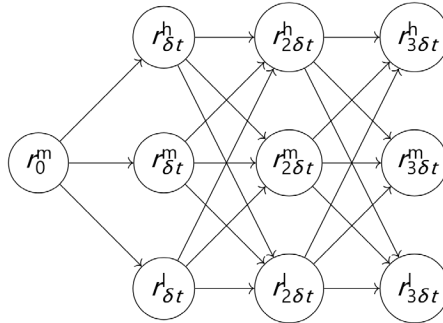


Рис. 9. Тринomialное дерево

Например, рассмотрим простое дерево, которое ограничено как низкими, так и высокими процентными ставками, и на каждом временном шаге  $rt$  может принимать одно из трех возможных значений:  $r_r^h = b + \delta r$  (высокое),  $r_r^m = b$  (среднее) и  $r_r^l = b - \delta r$  (низкое), где  $\delta r$  является константой.

Вероятности перехода  $q_{t \rightarrow t+\delta t}$  из узла  $r_t$  в узел  $r_{t+\delta t}$  зависят от узлов  $t$  и  $t+\delta t$ , и, таким образом,  $q_{t \rightarrow t+\delta t}$  представляет собой массив из  $3 \times 3 = 9$  значений. Эти вероятности можно вывести, приравняв ожидаемое значение и дисперсию непрерывной и дискретной моделей соответственно.

**2.1.1. Вентиль  $D_{r_i}$ : распределение  $P(r_i)$ .** Учитывая, что на каждом временном шаге  $t$  существует 3 возможных результата, нам нужны два кубита,  $|0\rangle_{rf_1}^t |0\rangle_{rf_0}^t$ , для кодирования вероятностей  $q$ ; следовательно,  $m$  временных шагов требуют  $2m$  кубитов «фактора риска». На каждом временном шаге используется первый кубит ( $rf_0$ ) из пары для представления перехода процентной ставки на средний уровень:  $r_t \rightarrow r_{t+\delta t}^m$ . В частности, состояние  $|0\rangle_{rf_0}^t$  описывает переход на средний узел, а состояние  $|1\rangle_{rf_0}^t$  – переход на другой узел. В последнем случае второй кубит ( $rf_1$ ) используется для моделирования перехода  $r_t$  на высокий ( $r_t \rightarrow r_{t+\delta t}^h$ , описанный как  $|1\rangle_{rf_1}^t$ ) или низкий ( $r_t \rightarrow r_{t+\delta t}^l$ , описанный как  $|0\rangle_{rf_1}^t$ ) уровни соответственно. Поскольку вероятности перехода зависят от уровня процентной ставки, состояние кубита «фактора риска» необходимо считывать на каждом временном шаге перед кодированием

вероятностей следующего перехода. Это можно облегчить, включив 3 дополнительных кубита «состояния» – наряду с  $2m$  кубитами «фактора риска» – для хранения значения  $r_t$  на каждом временном шаге; т.е.  $|001\rangle_{st}^t$ ,  $|010\rangle_{st}^t$  и  $|100\rangle_{st}^t$ , которые представляют высокий, средний и низкий уровни процентной ставки соответственно. Квантовое состояние на первом временном шаге:

$$\begin{aligned}
 &|\psi\rangle_{rf}^{\delta t} |\psi\rangle_{st}^{\delta t} = \\
 &\sqrt{q_{mh}} |11\rangle_{rf}^{\delta t} |001\rangle_{st}^{\delta t} + \sqrt{q_{mm}} |00\rangle_{rf}^{\delta t} |010\rangle_{st}^{\delta t} + \sqrt{q_{ml}} |01\rangle_{rf}^{\delta t} |100\rangle_{st}^{\delta t}, \quad (23)
 \end{aligned}$$

где первый и второй индексы вероятностей перехода обозначают начальный и конечный узлы. Чтобы собрать вентиль  $D_{tr}$ , собирается оператор «чтения»  $R_t = R_t^l R_t^m R_t^h$ , который считывает состояние  $|\psi\rangle_{st}^t$  и кодирует вероятности перехода в  $|\psi\rangle_{rf}^t$  (рисунок 10), где

$$\theta_0^{mm} = 2 \arcsin \sqrt{q_{mm}}, \quad (24)$$

$$\theta_1^{mh} = 2 \arcsin \sqrt{q_{mh}/(1-q_{mm})}. \quad (25)$$

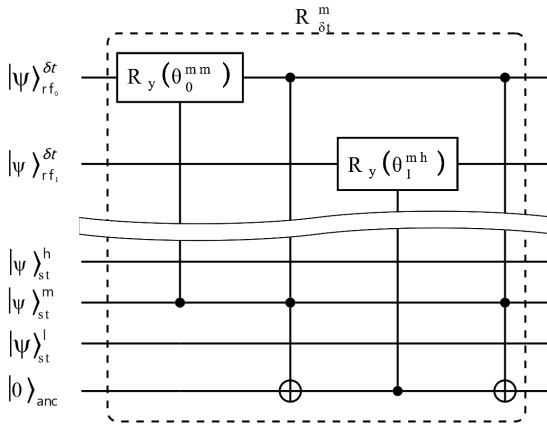


Рис. 10. Оператор  $R_{\delta t}^m$

Здесь управляемый вентиль  $R_y(\theta_0^{hm})$  кодирует вероятность  $q_{mm}$  – а именно вероятность того, что ставка останется на среднем уровне  $r_{\delta t} = r_0 = b$  – в  $|\psi\rangle_{rf_0}^{\delta t}$ . Второй вентиль, «И», проверяет, являются ли оба  $r_0 = b$  и  $r_{\delta t} \neq b$  истинными – а именно, изменилась ли процентная ставка – и записывает результат во «вспомогательный» кубит.

Третий вентиль, управляемый  $R_y(\theta_1^{hm})$ , проверяет результат, хранящийся во «вспомогательном» кубите, и если он равен  $|1\rangle$ , кодирует условную вероятность перехода  $r_{\delta t}$  к высокому значению процентной ставки, учитывая, что он не остался на среднем уровне:  $P(r_{\delta t} = b + \delta r | r_{\delta t} \neq b) = P(r_{\delta t} = b + \delta r) / P(r_{\delta t} \neq b) = q_{mh} / (1 - q_{mm})$ .

Четвертый вентиль, «И», гарантирует, что «вспомогательный» кубит находится в исходном состоянии. Логика для вентилях  $R_{\delta t}^h$  и  $R_{\delta t}^l$  аналогична (их разложение на рисунке 11), где показан пример всего вентиля «Чтение» на первом временном шаге,  $R_{\delta t}$ .

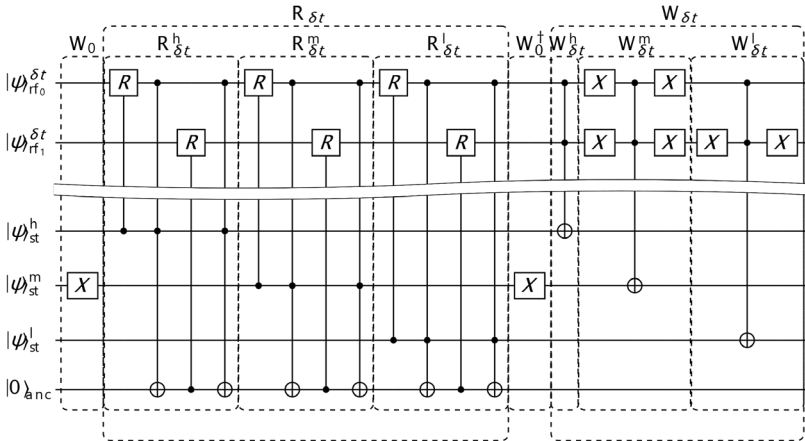


Рис. 11. Разложение операторов «чтения»  $R_t$  и «записи»  $W_t$  для первого временного шага, предполагая, что  $r_0 = b$  (т.е.  $W_0 |000\rangle_{st} = |100\rangle_{st}^0$ ). Слева направо вентили  $R$ :  $R_y(\theta_0^{hm})$ ,  $R_y(\theta_1^{hh})$ ,  $R_y(\theta_0^{mm})$ ,  $R_y(\theta_1^{mh})$ ,  $R_y(\theta_0^{lm})$ ,  $R_y(\theta_1^{lh})$

Вентили «Записи»  $W_t$  считывают пару кубитов «фактора риска»  $|\psi\rangle_{rf}^t$  и записывают результат в кубиты «состояния»,  $|\psi\rangle_{st}^t$ :

$$W_t |11\rangle_{rf}^t |000\rangle_{st} = |11\rangle_{rf}^t |001\rangle_{st}, \quad (26)$$

$$W_t |00\rangle_{rf}^t |000\rangle_{st} = |00\rangle_{rf}^t |010\rangle_{st}, \quad (27)$$

$$W_t |01\rangle_{rf}^t |000\rangle_{st} = |01\rangle_{rf}^t |100\rangle_{st}, \quad (28)$$

Предыдущее состояние должно быть стерто; поэтому применению  $W_t$  всегда предшествует вентиль  $W_{t-\delta t}^\dagger$ , который сбрасывает кубиты «состояния» обратно в  $|000\rangle_{st}$ . Обратные вентили «чтения» и «записи»,  $R_t^\dagger$  и  $W_t^\dagger$ , состоят из компонентов  $R_t$  и  $W_t$ , расположенных в обратном порядке. Наконец, оператор  $D_{ir}$  может быть собран последовательностью операторов «чтения» и «записи» для каждого временного шага, как показано на рисунке 12.

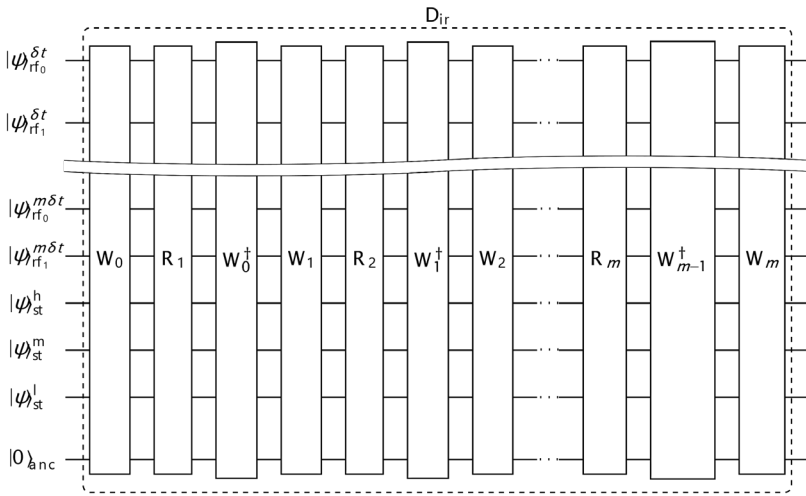


Рис. 12. Вентиль  $D_{ir}$  состоит из последовательности вентилях «чтения» и «записи»

**2.1.2. Вентиль  $M_m$ : меры риска  $F(r_t)$ .** Для реализации триномиального дерева мы рассматриваем в качестве меры риска вероятность того, что процентная ставка будет равна своему долгосрочному среднему уровню после  $m$  временных шагов. В этом случае вентиль  $M_m$  будет выглядеть как на рисунке 13 с  $M_m^\dagger = M_m$ .

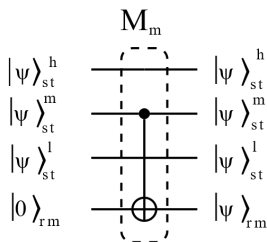


Рис. 13 Вентиль  $M_m$

**2.1.3. Результаты.** В таблице 2 перечислены параметры квантовой схемы, реализующей триномиальные деревья для эволюции процентной ставки (из-за использования параметризации не выбираются значения для  $a$  и  $b$ ). Для каждого из  $m = 3$  временных шагов нужно 2 кубита «фактора риска», всего 6:  $|\psi\rangle_{rf}$ . Так же нужно еще 3 кубита для хранения состояния процентной ставки  $|\psi\rangle_{st}$ , 1 кубит для меры риска  $|\psi\rangle_{rm}$ , и 9 «вспомогательных» кубитов  $|\psi\rangle_{anc}$  для вентиля  $A$ . Свободно выбирается  $\delta r$ , который устанавливается кратно стандартному отклонению. В тоже время определение  $0 \leq q_{t \rightarrow t+\delta t} \leq 1$  ограничивает выбор  $\delta t$ . Определив  $\delta r$  и  $\delta t$ , вычисляются вероятности перехода (таблица 2).

Таблица 2. Список параметров триномиального дерева для эволюции фактора риска процентной ставки

Параметр	Значение
$m$	3
$n$	1-9
$\delta r$	$\sqrt{3Var}$
$\delta t$	$3a/12$

Если выбрать  $r_0=b$ , вероятность измерения значения  $r_{3\delta t}=b$  после 3 временных шагов вычисляется, сложив вероятности всех возможных путей:

$$P(r_{3\delta t} = b) = \sum_{s_1} \left[ q_{ms_1} \left( \sum_{s_2} q_{s_1s_2} q_{s_2m} \right) \right], \quad (29)$$

где  $s_1, s_2 \in \{h, m, l\}$  – узлы первого и второго временных шагов соответственно.

На рисунке 14 показано распределение вероятностей кубитов «состояния» (левая панель) и распределение кубита «меры риска» (правая панель) при  $t=3\delta t$  при измерении этих кубитов напрямую с 10000 запусками.

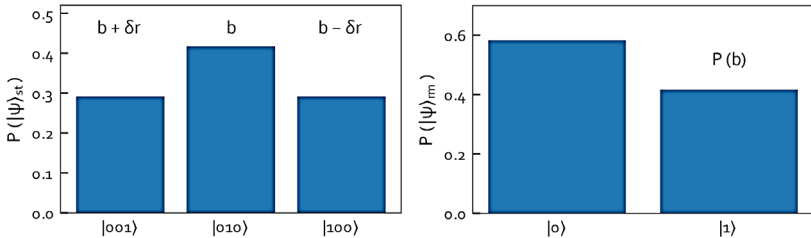


Рис. 14. Распределение  $P(r_{3\delta t})$  (левая панель) и суперпозиция кубита «фактора риска», который кодирует  $P(r_{3\delta t}=b)$  (правая панель)

## 2.2. Факторы кредитного риска

**2.2.1. Модели кредитного риска в сокращенной форме.** Один из подходов к оценке вероятности дефолта – это модели кредитного риска в сокращенной форме, в которых дефолт моделируется как статистический процесс. Вероятность выживания от времени  $t_0$  до времени  $t$  определяется по формуле:

$$P(t_0 + t) = e^{-t/T_{def}}, \quad (30)$$

где  $T_{def}$  – характерная временная шкала, обратная мере опасности,  $\lambda=1/T_{def}$ . Чтобы вычислить вероятность «выживания» в момент времени  $T$ , можно дискретизировать временной интервал  $t \in [0, T]$  с  $m$  временными шагами,  $\delta t=T/m$ , так что вероятность «выживания» равна:



$$P(T) = e^{-\sum \delta t / T} = \prod_{j=1}^m P(\delta t). \quad (31)$$

**2.2.1.1. Вентиль  $D_{surv}$ : распределение  $P(T)$ .** На каждом из  $m$  временных шагов можно представить «выживание» одним кубитом «фактора риска», таким образом, что  $|0\rangle$  представляет «выживание», а  $|1\rangle$  дефолт. Таким образом, если компания выжила до момента времени  $t=(l-1)\delta t$ , то  $|\psi\rangle_{rf}^{\delta t} = \sqrt{1-q_{def}}|0\rangle + \sqrt{q_{def}}|1\rangle$ , где  $q_{def}=1-P(\delta t)$ . Если компания допустила дефолт в предыдущий временной шаг, то:  $|\psi\rangle_{rf}^{\delta t} = |1\rangle$ .

Вентиль  $D_{surv}$  можно собрать с помощью вентилях вращения  $R_y^{\theta_{def}}$  ( $\theta_d$  для краткости), полученных из  $q_{def}=\sin(\theta_{def}/2)$ . Кроме того, чтобы гарантировать, что в случае дефолта в момент времени  $t$ , компания также будет в дефолте в момент  $t+\delta t$ , включается управляемый вентиль  $R_y^{\theta_{def}^c}$ , где  $\theta_{def}^c + \theta_{def} = \pi$ : это нарушает суперпозицию кубита, представляющего  $t+\delta t$ , и устанавливает его в  $|1\rangle$ . Таким образом, вентиль  $D_{surv}$  имеет вид как на рисунке 15.

Обратный вентиль,  $D_{surv}^\dagger$ , содержит вентили вращения в обратном порядке и знаки углов вращения перевернуты.

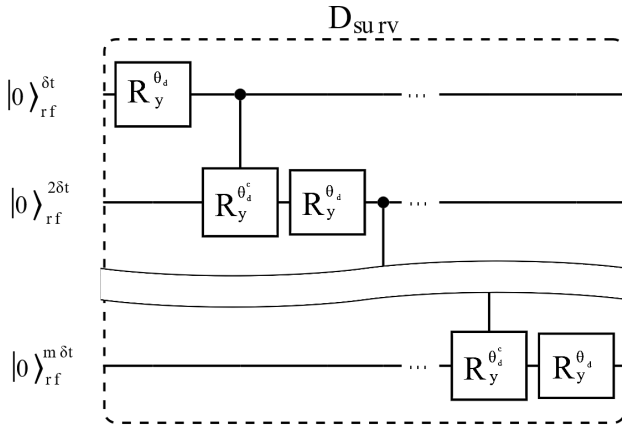


Рис. 15. Вентиль  $D_{surv}$

**2.2.1.2. Вентиль  $M_{surv}$ : вероятность выживания.** Вероятность выживания при  $t=T$  представлена состоянием  $|0\dots 00\rangle_{rf}^T$ . Поэтому можно использовать вентиль  $M_{surv}=M_{min}$ , который переворачивает кубит «меры риска», если все кубиты «фактора риска» находятся в состоянии  $|0\rangle$ . Ожидаемое значение вероятности выживания в момент времени  $T$  равно  $(1-q_{def})^m$ .

**2.2.1.3. Результаты.** Таблица 3 содержит параметры квантовой схемы, а рисунок 16 – выживание ( $|1\rangle_{rm}$ ) и вероятности по умолчанию ( $|0\rangle_{rm}$ ), закодированные в кубите «меры риска». На рисунке 17 показано измерение выходного состояния и его сходимость аналогично предыдущим вариантам использования.

Таблица 3. Список параметров для модели кредитного риска в сокращенной форме

Параметер	Значение
m	6
n	1-9
T	1
$q_{def}$	2%
$\theta_{def} 180^\circ/\pi$	$\sim 16.3^\circ$

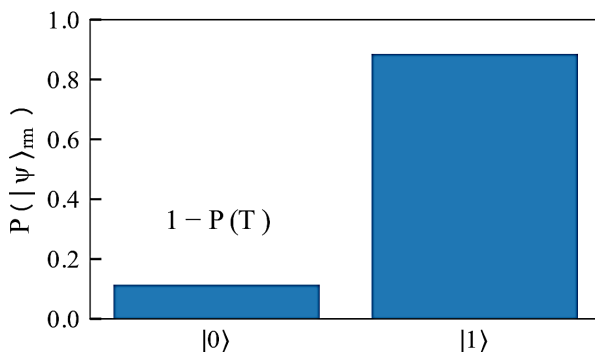


Рис. 16. Вероятности состояний кубита «меры риска»:  $|1\rangle$  обозначает выживание до  $t = T$ , а  $|0\rangle$  обозначает дефолт на любом временном шаге до достижения  $T$

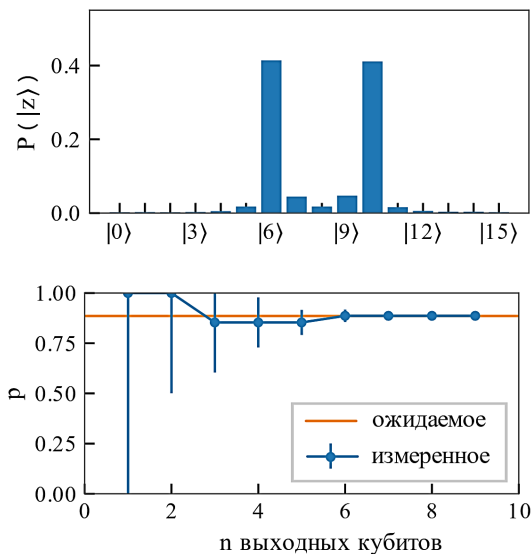


Рис. 17. Верхняя панель: распределение вероятностей выходных кубитов для  $m=4$  и 10000 запусков. Нижняя панель: расчетная вероятность (один запуск) и ее сходимости при увеличении числа выходных кубитов

**2.3. Миграция кредитного рейтинга.** Более сложный подход к моделированию вероятностей дефолта – это матрицы миграции кредитного рейтинга. Например, рассмотрим следующие три рейтинга:  $A$  (инвестиционный уровень),  $B$  (высокая доходность) и  $D$  (дефолт), и предположим, что при  $t=0$  эмитент имеет рейтинг  $A_0$ .

Изменение кредитного рейтинга и вероятность дефолта в момент времени  $t$  можно смоделировать с помощью вентиля  $D_{migr}=D_{ir}$  (с небольшой корректировкой для установки начального рейтинга на  $A$ , что будет соответствовать высокому уровню процентной ставки), а вероятность дефолта может быть закодирована с помощью  $M_{def}=M_l$  (соответствующей измерению низкого уровня процентной ставки).

**3. Оценка результатов.** Рассмотрим, сколько кубитов и какой глубины схема потребуется для типичных случаев использования в управлении финансовыми рисками. Предлагается проанализировать только генерацию сценариев и не учитывать дополнительные кубиты и вентили, которые могут потребоваться для других задач. Количество кубитов «фактора риска» ( $m_{rf}$ ) определяет количество временных шагов, а значит, и точность модели риска. Количество выходных кубитов ( $n$ ) определяет погрешность оценки, то есть точность

измерения. Чем больше входных и выходных кубитов, тем выше достоверность и точность результатов.

Модели финансового риска не учитывают внутрисуточные изменения, поэтому минимальный временной шаг – один рабочий день. Для долгосрочных сценариев, например охватывающих десятилетия, может использовать временные шаги вплоть до одного месяца. Модели кредитного риска могут использоваться для оценки вероятности выживания организации в течение нескольких лет. Такой подход может быть полезен при свопах по кредитному дефолту или анализе кредитного риска контрагента. Таким образом мы можем ограничить количество кубитов «фактора риска» вентиля  $D_{surv}$  несколькими сотнями (например, 600 для 50 лет). Поскольку  $M_{surv}$  требуется ряд «вспомогательных» кубитов, общее число входных кубитов для этого типа схем не будет превышать  $\sim 1200$ .

Для сценариев процентных ставок количество кубитов «фактора риска» зависит от числа временных шагов. Поскольку процентные ставки меняются не так сильно, как акции, временного шага в один месяц часто достаточно для моделирования их эволюции. Каждый новый дополнительный кубит может удвоить количество моделируемых значений процентной ставки. Однако масштабирование является сублинейным. Кроме того, для вентиля  $M_m$  требуется только один дополнительный кубит – «мера риска» и моделирование 50-летнего периода потребует не более тысячи входных кубитов.

Тот же анализ применим и к многочленным деревьям для моделирования миграции кредитного рейтинга. Для достижения точности в 1 базисный пункт (0,01%) потребуется около 14 выходных кубитов. Таким образом, для высокоточных результатов в целом потребуется чуть больше тысячи кубитов с учётом как входных, так и выходных.

Оценим глубину схемы: чем больше число вентиляй используется, тем больше ошибок из-за шума. Глубина определяется численно, разложением вентиляй на их составляющие. Этот процесс повторяется до тех пор, пока дальнейшее разложение невозможно. Левые панели рис. 18 показывают глубину вентиляй  $D$  и  $M$  как функцию количества входных кубитов  $m$  (точность модели). Можно сделать вывод, что глубина линейно масштабируется с количеством входных кубитов. Для  $m \approx 10$  глубина составляет  $\leq 500$ , а для  $m \approx 1000$  – порядка  $\leq 50000$ .

Правые панели рис. 18 показывают общую глубину схемы (она включает в себя вентиляй  $D$ ,  $M$ ,  $PQ$ ,  $QFT$  и  $QFT^t$ ) в логарифмическом

масштабе как функцию количества выходных кубитов  $n$  (точность модели).

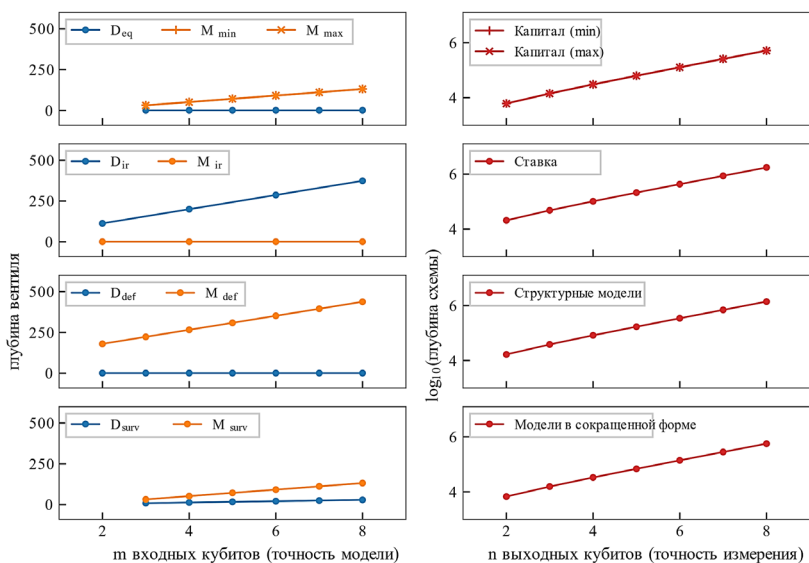


Рис. 18. Левые панели: глубина вентилей  $D$  и  $M$  как функция числа входных кубитов,  $t$ , что представляет точность модели. Правые панели: глубина (в логарифмической шкале) всей квантовой схемы как функция числа выходных кубитов,  $n$ , что представляет точность измерения

Глубина увеличивается экспоненциально по отношению к  $n$ , поскольку каждый дополнительный выходной кубит удваивает количество раз, когда применяется вентиль  $Q$  (14). Для  $n=14$  общая глубина будет порядка  $\leq 10^8$ . Оценки ценообразования различных деривативов говорят о числе кубитов и глубин схем, которые находятся на том же порядке величины,  $\sim 10^3-10^4$  и  $\sim 10^8$  соответственно [18, 15].

В ближайшие несколько лет [19, 20] появятся квантовые компьютеры с тысячей кубитов, для таких вариантов использования потребуется значительный прогресс в отказоустойчивости. Для портфеля, который зависит от нескольких факторов риска, понадобится кратное количество входных кубитов при параллельной обработке или обработка будет происходить последовательно. Из-за ограничений оборудования более вероятен последовательный запуск генерации сценариев.

**4. Заключение.** Алгоритмы QAE обеспечивают квадратичное ускорение по сравнению с их классическими аналогами.

В работах [21, 22] предварительно вычисляют распределение факторов риска (с помощью классических вычислений) и фокусируются на расчете общих мер риска с нагрузкой распределения и оптимизации QAE. В результате проведенного исследования за счет последовательной интеграции генерации сценариев в квантовые вычисления QAE устраняются ограничения классических алгоритмов, нивелировавшие квадратичное ускорение, получаемое за счет применения квантовых алгоритмов. В частности, были собраны квантовые вентили, которые реализуют стохастические модели риска для процентной ставки и факторов кредитного риска. Для процентных ставок сценарии генерируются путем дискретизации стохастических дифференциальных уравнений возврата к среднему с помощью ограниченных тринomialных деревьев. Для факторов кредитного риска вычисляется вероятность выживания из моделей в сокращенной форме на основе процесса Пуассона. Также описывается, как мультиномиальные деревья могут быть использованы для реализации матриц миграции кредитных рейтингов. Для каждого варианта использования построена сквозная схема QMC/QAE, которая включает: генерацию сценариев факторов риска (QMC), кодирование меры риска и оценку значения меры риска (QAE). Данные схемы были проверены с использованием эмулятора квантового компьютера [23] и продемонстрировано, что измеренное значение сходится к ожидаемому значению, а ошибка стремится к нулю по мере увеличения числа выходных кубитов. Таким образом, показана возможность практического применения QAE с целью оценки рисков в финансовых операциях при сохранении преимуществ квантового подхода над классическими вычислениями.

Для типичной точности модели и точности измерения реалистичных финансовых ситуаций оценивается, что эволюция фактора риска требует  $\lesssim 1200$  кубитов, что находится в пределах возможностей квантовых компьютеров, ожидаемых в ближайшие несколько лет. Однако так же стоит уточнить, что типичная глубина схемы составляет порядка  $\lesssim 10^8$ , что представляет самую большую проблему, пока не появятся более отказоустойчивые квантовые устройства.

Квантовые вентили моделируют выходные состояния случайной величины, а суперпозиция позволяет квантовой схеме одновременно моделировать все возможные значения зависимой от времени случайной величины, что устраняет ограничения классических вычислений при итерации между наборами данных. Квантовые приложения финансового риска выигрывают от последовательного

включения генерации сценариев в моделирование QMC/QAE, что снижает их зависимость от классических компьютеров.

### Литература

1. Orús R., Mugel S., Lizaso E. Quantum computing for finance: Overview and prospects // *Reviews in Physics*. 2019. vol. 4. DOI: 10.1016/j.revip.2019.100028.
2. Egger D.J., Gambella C., Marecek J., McFaddin S., Mevissen M., Raymond R., Simonetto A., Woerner S., Yndurain E. Quantum computing for finance: State-of-the-art and future prospects // *IEEE Transactions on Quantum Engineering*. 2020. vol. 1. pp. 1–24. DOI: 10.1109/TQE.2020.3030314.
3. Gómez A., Rodriguez A.L., Manzano A., Nogueiras M., Ordóñez G., Vázquez C. A Survey on quantum computational finance for derivatives pricing and VaR // *Archives of Computational Methods in Engineering*. 2022. vol. 29. pp. 4137–4163. DOI: 10.1007/s11831-022-09732-9.
4. Herman D., Googin C., Liu X., Galda A., Safro I., Sun Y., Pistoia M., Alexeev Y.A. Survey of Quantum Computing for Finance // 2022. arxiv preprint arxiv: 2201.02773.
5. Wilkens S., Moorhouse J. Quantum computing for financial risk measurement // *Quantum Information Processing*. 2023. vol. 22. DOI: 10.1007/s11128-022-03777-2.
6. Intallura P., Korpas G., Chakraborty S., Kungurtsev V., Marecek J. A Survey of Quantum Alternatives to Randomized Algorithms: Monte Carlo Integration and Beyond // 2023. arxiv preprint arxiv: 2303.04945.
7. Dalzell A.M., McArdle S., Berta M., Bienias P., Chen C., Gilyén A., Hann C.T., Kastoryano M.J., Khabiboulline E.T., Kubica A., Salton G., Wang S., Brandão F.G.S.L. Quantum algorithms: A survey of applications and end-to-end complexities // 2023. arxiv preprint arxiv:2310.03011.
8. Romano C., Di Clichemnte A. Measuring Portfolio Value-at-Risk by a Copula-EVT Based Approach // *Studi economici*. 2005. vol. 2005(85). pp. 29–65.
9. Brassard G., Hoyer P., Mosca M., Tapp A. Quantum amplitude amplification and estimation // *Quantum Computation and Information*. 2002. vol. 305. pp. 53–74. DOI: 10.1090/conm/305/05215.
10. Herbert S. No quantum speedup with grover-rudolph state preparation for quantum monte carlo integration // *Physical Review E*. 2021. vol. 103(6-1). DOI: 10.1103/PhysRevE.103.063302.
11. Vazquez A.C., Woerner S. Efficient state preparation for quantum amplitude estimation // *Physical Review Applied*. 2021. vol. 15(3). DOI: 10.1103/PhysRevApplied.15.034027.
12. McArdle S., Gilyén A., Berta M. Quantum state preparation without coherent arithmetic // 2022. arxiv preprint arxiv:2210.14892.
13. Zoufal C., Lucchi A., Woerner S. Quantum generative adversarial networks for learning and loading random distributions // *Quantum Information*. 2019. vol. 5(103). DOI: 10.1038/s41534-019-0223-2.
14. Li J., Kais S. A universal quantum circuit design for periodical functions // *New Journal of Physics*. 2021. vol. 23. DOI: 10.1088/1367-2630/ac2cb4.
15. Stamatopoulos N., Zeng W.J. Derivative Pricing using Quantum Signal Processing // 2023. arxiv preprint arxiv:2307.14310.
16. Grover L.K. A fast quantum mechanical algorithm for database search // *Proceedings of the Twenty Eighth Annual ACM Symposium on the Theory of Computing*. 1996. pp. 212–219. DOI: 10.1145/237814.237866.
17. Vasicek O. An equilibrium characterization of the term structure // *Journal of Financial Economics*. 1977. vol. 5(2). pp. 177–188.

18. Chakrabarti S., Krishnakumar R., Mazzola G., Stamatopoulos N., Woerner S., Zeng W.J. A Threshold for Quantum Advantage in Derivative Pricing // 2021. Quantum 5. vol. 463. DOI: 10.22331/q-2021-06-01-463.
19. Castelveccchi D. IBM releases first-ever 1,000-qubit quantum chip, 2023. URL: <https://www.nature.com/articles/d41586-023-03854-1> (accessed 18.09.2024).
20. IBM. Charting the course to 100,000 qubits. 2023. URL: <https://research.ibm.com/blog/100k-qubit-supercomputer>. (accessed 18.09.2024).
21. Боряев Р.О., Чуваков А.В. Квантовые вычисления в автоматизированных системах управления факторинговыми операциями // Вестник СамГТУ, Технические науки. 2023. № 2(78). С. 6–19.
22. Боряев Р.О., Чуваков А.В. Необходимость использования квантовых вычислений в автоматизированных системах управления факторинговыми операциями // Сборник материалов IX Международной научно-практической конференции (школы-семинара) молодых ученых (г. Тольятти, 18-20 апреля). 2023. С. 407–412.
23. AWS Cloud Quantum Computing Service. IQM, 2024. URL: <https://aws.amazon.com/ru/braket/quantum-computers/iqm/>. (accessed 18.09.2024).

**Чуваков Александр Владимирович** — канд. хим. наук, доцент, заведующий кафедрой, кафедра «информатика и вычислительная техника», Самарский Государственный Технический Университет. Область научных интересов: искусственный интеллект, методы проектирования мультиагентных систем контроля и управления, квантовые вычисления. Число научных публикаций — 120. [avch2105@gmail.com](mailto:avch2105@gmail.com); улица Молодогвардейская, 244, 443100, Самара, Россия; р.т.: +7(846)337-1286.

**Боряев Родион Олегович** — аспирант, кафедра «информатика и вычислительная техника», Самарский Государственный Технический Университет. Область научных интересов: экспертная поддержка принятия решений, квантовые вычисления, анализ кредитных рисков. Число научных публикаций — 5. [r.boryaev@gmail.com](mailto:r.boryaev@gmail.com); улица Молодогвардейская, 244, 443100, Самара, Россия; р.т.: +7(846)337-1286.



A. CHUVAKOV, R. BORYAEV  
**FACTORING DECISION SUPPORT SYSTEM BASED ON  
OPTIMIZED QUANTUM ALGORITHMS QMC**

*Chuvakov A., Boryaev R. Factoring Decision Support System Based on Optimized Quantum Algorithms QMC.*

**Abstract.** The continuous growth of financial markets dictates the need for its participants to seek new approaches to financial analysis to gain competitive advantages, including through the use of new approaches in the field of computing. Quantum computing can be used as a tool for obtaining these advantages over competitors. In particular, Monte Carlo modeling, although widely used in financial risk management, requires significant computing resources due to the large number of scenarios required to obtain an accurate result. To optimize this approach, quantum amplitude estimation algorithms are used, which accelerate this process if pre-calculated probability distributions are used to initialize input quantum states. However, in the absence of these distributions in existing approaches to this topic, they are generated numerically using classical computing, which completely negates the advantage of the quantum approach. This article proposes a solution to this problem by using quantum computing, including for the generation of probability distributions. The article discusses the creation of quantum circuits for modeling the evolution of risk factors over time for capital flows, interest rates, and credit risks, and presents the combination of these models with quantum amplitude estimation algorithms as an example of using the obtained algorithms for credit risk management. In conclusion, the article analyzes the possibility of using the obtained circuits in financial analysis.

**Keywords:** quantum computing, quantum Monte Carlo method, quantum amplitude estimation, financial risk assessment.

## References

1. Orús R., Revil S., Lizaso E. Quantum computing for finance: Overview and prospects. *Reviews in Physics*. 2019. vol. 4. DOI: 10.1016/j.revip.2019.100028.
2. Egger D.J., Gambella C., Marecek J., McFaddin S., Mevissen M., Raymond R., Simonetto A., Woerner S., Yndurain E. Quantum computing for finance: State-of-the-art and future prospects. *IEEE Transactions on Quantum Engineering*. 2020. vol. 1. pp. 1–24. DOI: 10.1109/TQE.2020.3030314.
3. Gómez A., Rodríguez A.L., Manzano A., Nogueiras M., Ordóñez G., Vázquez C. A Survey on quantum computational finance for derivatives pricing and VaR. *Archives of Computational Methods in Engineering*. 2022. vol. 29. pp. 4137–4163. DOI: 10.1007/s11831-022-09732-9.
4. Herman D., Googin C., Liu X., Galda A., Safro I., Sun Y., Pistoia M., Alexeev Y.A. Survey of Quantum Computing for Finance. 2022. arxiv preprint arxiv: 2201.02773.
5. Wilkens S., Moorhouse J. Quantum computing for financial risk measurement. *Quantum Information Processing*. 2023. vol. 22. DOI: 10.1007/s1128-022-03777-2.
6. Intallura P., Korpas G., Chakraborty S., Kungurtsev V., Marecek J. A Survey of Quantum Alternatives to Randomized Algorithms: Monte Carlo Integration and Beyond. 2023. arxiv preprint arxiv: 2303.04945.
7. Dalzell A.M., McArdle S., Berta M., Bienias P., Chen C., Gilyén A., Hann C.T., Kastoryano M.J., Khabiboulline E.T., Kubica A., Salton G., Wang S., Brandão F.G.S.L. Quantum algorithms: A survey of applications and end-to-end complexities. 2023. arxiv preprint arxiv:2310.03011.

8. Romano C., Di Climemnte A. Measuring Portfolio Value-at-Risk by a Copula-EVT Based Approach. *Studi economici*. 2005. vol. 2005(85). pp. 29–65.
9. Brassard G., Hoyer P., Mosca M., Tapp A. Quantum amplitude amplification and estimation. *Quantum Computation and Information*. 2002. vol. 305. pp. 53–74. DOI: 10.1090/conm/305/05215.
10. Herbert S. No quantum speedup with grover-rudolph state preparation for quantum monte carlo integration. *Physical Review E*. 2021. vol. 103(6-1). DOI: 10.1103/PhysRevE.103.063302.
11. Vazquez A.C., Woerner S. Efficient state preparation for quantum amplitude estimation. *Physical Review Applied*. 2021. vol. 15(3). DOI: 10.1103/PhysRevApplied.15.034027.
12. McArdle S., Gilyén A., Berta M. Quantum state preparation without coherent arithmetic. 2022. arxiv preprint arxiv:2210.14892.
13. Zoufal C., Lucchi A., Woerner S. Quantum generative adversarial networks for learning and loading random distributions. *Quantum Information*. 2019. vol. 5(103). DOI: 10.1038/s41534-019-0223-2.
14. Li J., Kais S. A universal quantum circuit design for periodical functions. *New Journal of Physics*. 2021. vol. 23. DOI: 10.1088/1367-2630/ac2cb4.
15. Stamatopoulos N., Zeng W.J. Derivative Pricing using Quantum Signal Processing. 2023. arxiv preprint arxiv:2307.14310.
16. Grover L.K. A fast quantum mechanical algorithm for database search. *Proceedings of the Twenty Eighth Annual ACM Symposium on the Theory of Computing*. 1996. pp. 212–219. DOI: 10.1145/237814.237866.
17. Vasicek O. An equilibrium characterization of the term structure. *Journal of Financial Economics*. 1977. vol. 5(2). pp. 177–188.
18. Chakrabarti S., Krishnakumar R., Mazzola G., Stamatopoulos N., Woerner S., Zeng W.J. A Threshold for Quantum Advantage in Derivative Pricing. 2021. *Quantum* 5. vol. 463. DOI: 10.22331/q-2021-06-01-463.
19. Castelvecchi D. IBM releases first-ever 1,000-qubit quantum chip, 2023. Available at: <https://www.nature.com/articles/d41586-023-03854-1> (accessed 18.09.2024).
20. IBM. Charting the course to 100,000 qubits., 2023. Available at: <https://research.ibm.com/blog/100k-qubit-supercomputer>. (accessed 18.09.2024).
21. Боряев Р.О., Чуваков А.В. Квантовые вычисления в автоматизированных системах управления факторинговыми операциями. *Вестник СамГТУ, Технические науки*. 2023. № 2(78). С. 6–19.
22. Boryaev R.O., Chuvakov A.V. [The need to use quantum computing in automated factoring management systems] *Sbornik materialov IX Mezhdunarodnoj nauchno-prakticheskoy konferencii* [Proc. 9th Int. scientific and practical conference] (г. Тольятти, 18-20 апреля). 2023. pp. 407–412. (In Russ.).
23. AWS Cloud Quantum Computing Service. IQM, 2024. Available at: <https://aws.amazon.com/ru/braket/quantum-computers/iqm/>. (accessed 18.09.2024).

**Chuvakov Aleksandr** — Ph.D., Associate Professor, Head of the department, Department of computer engineering, Samara State Technical University. Research interests: artificial intelligence, methods for designing multi-agent control and management systems, quantum computing. The number of publications — 120. [avch2105@gmail.com](mailto:avch2105@gmail.com); 244, Molodogvardeyskaya St., 443100, Samara, Russia; office phone: +7(846)337-1286.

**Boryaev Rodion** — Postgraduate student, Department of computer engineering, Samara State Technical University. Research interests: expert support for decision-making, quantum computing, credit risk analysis. The number of publications — 5. [r.boryaev@gmail.com](mailto:r.boryaev@gmail.com); 244, Molodogvardeyskaya St., 443100, Samara, Russia; office phone: +7(846)337-1286.

А.В. ИВАЩЕНКО, О.К. ГОЛОВНИН, А.А. ГОЛОВНИНА, Е.А. ДОДОНОВА  
**КОМБИНИРОВАННАЯ ГЕОИНФОРМАЦИОННАЯ  
МНОГОСЛОЙНАЯ ПРОСТРАНСТВЕННО-ВРЕМЕННАЯ  
МОДЕЛЬ**

*Иващенко А.В., Головнин О.К., Головнина А.А., ДодONOва Е.А.* **Комбинированная геоинформационная многослойная пространственно-временная модель.**

**Аннотация.** Современные технологии организационного управления предусматривают сбор и обработку большого объема данных для расчета параметров функционирования исследуемых объектов и процессов. Поскольку основная особенность собираемых параметров состоит в их привязке к территориям, с одной стороны, и отнесении к периодам времени, с другой стороны, требуется применение геоинформационных систем и технологий. Несмотря на развитие современных геоинформационных технологий, вопросы их практического применения для поддержки принятия решений с учетом комбинированного влияния пространственного и временного факторов, в полной мере не решены. В статье предложена комбинированная геоинформационная многослойная пространственно-временная модель, которая представляет собой граф, вершинами которого являются значения параметров, упорядоченные по слоям с размещением временных меток в слой времени, а дугами – отношения между ними, разделенные на три типа: топологические, семантические и хронологические. Сопряжение и упорядочивание параметров, согласно предложенной модели, позволяет корректно поставить и решить задачу оптимизации, а, следовательно, устранить проблему практического использования накапливаемой аналитики в процессах поддержки принятия управленческих решений. Предложенная модель использована в цифровой платформе интегрального мониторинга для цифровой трансформации процессов сбора, анализа и визуализации данных коммунальных ресурсов. Рассмотрена общая задача управления, а также приведен конкретный пример для одной из актуальных задач регионального управления – социальной газификации, в котором производится оптимизация процесса обработки заявок по подключению жилых домов к системе газоснабжения в границах выбранного региона. Комбинированная геоинформационная многослойная пространственно-временная модель позволяет формулировать универсальные постановки задач поддержки принятия решений для различных приложений геоинформатики в логистике, при управлении транспортными ресурсами, а также в ситуационных центрах управления предприятиями и регионами, в системах бизнес-аналитики и управления организационными системами.

**Ключевые слова:** геоинформационная система, геоинформационная модель, пространственная модель, пространственно-временная модель, пространственная оптимизация, интегральный мониторинг, ГИС.

**1. Введение.** Современные технологии организационного управления предусматривают сбор и обработку большого объема данных. Например, планирование и реализация мероприятий социально-экономического развития регионов в настоящее время основывается на результатах мониторинга и бизнес-аналитики, для чего активно внедряются информационные системы поддержки принятия управленческих решений. Повсеместное внедрение

информационных технологий и цифровая трансформация процессов позволяет обеспечить руководство регионов данными практически по всем аспектам общественной жизни населения.

Однако большой объем информации и множество взаимозависимостей затрудняет ее обработку в рамках поддержки принятия управленческих решений. Обычно лицо, принимающее решения, сосредотачивается на наиболее критичных, важных для конкретного региона показателях. Сформирована система показателей, по которым осуществляется государственный контроль. Однако задачи прогнозирования и оптимизации показателей социально-экономического развития на системном уровне в настоящее время до конца не решены.

Для того чтобы устранить этот недостаток, предлагается в рамках цифровых систем интегрального мониторинга и управления предусмотреть включение контура поддержки принятия решений. Поскольку основная особенность собираемых показателей состоит в их привязке к территориям субъектов Российской Федерации, регионов, районов, населенных пунктов, конкретных домовладений и т.п., с одной стороны, и отнесении к периодам времени (годам, кварталам, месяцам, суткам и т.п.), с другой стороны, решение видится в области внедрения геоинформационных систем и технологий.

Расширение возможностей современных геоинформационных систем в направлении реализации функциональности моделирования и прогнозирования может быть произведено за счет добавления в модель параметра времени как нового измерения с разбиением геоинформационных данных на слои, соответствующие разным этапам развития обстановки во времени, разным вариантам развития обстановки (ситуациям), а также разным масштабам представления обстановки лицам, принимающим решения.

В данной статье представлена такая модель, позволяющая описывать динамику развития обстановки в пространстве и времени, а также результаты ее реализации для планирования ресурсного снабжения жителей региона на примере их подключения к газораспределительной сети.

**2. Существующие подходы к решению проблемы.** В рамках теории геоинформационных систем (ГИС) [1, 2] обычно рассматривают понятие геоинформационного пространства – среды, в которой функционируют цифровая геоинформация и геоизображения разных видов и назначения. Геоинформационное пространство содержит совокупность геоинформационных территориальных моделей обстановки, охватывающих наземные,

подземные и надземные пространственные объекты, связанные между собой единой координатной основой.

В ГИС геоинформационное пространство может быть представлено в двумерном и трехмерном видах, а существующие объекты в качестве своих характеристик обязательно имеют геометрический размер и координату, по которой можно однозначно определить их расположение в пространстве. В современных ГИС существует возможность задания полигональных пространственных объектов, линий, обозначения регионов и территорий. Возможно также обобщение нескольких объектов в один и, наоборот, их декомпозиция в случае изменения масштаба.

Модель «Пространство-время» расширяет геоинформационное пространство путем введения дополнительного равноправного временного измерения [3]. Нерелятивистская классическая механика рассматривает время как универсальную величину измерения, которая является однородной во всем пространстве и которая отделена от пространства. Классическая механика предполагает, что время имеет постоянную скорость течения, которая не зависит от состояния движения наблюдателя [4].

В философии концепция связи пространственно-временных координат описывается с помощью модели хронотопа [5, 6]. Функциональная определенность пространственно-распределенных групп образуется за счет увязки во времени, скоростях, ритмах и сроках действия [7]. Взаимосвязь временных и пространственных отношений играет существенную роль в литературе [8]. В рамках имитационного моделирования можно реализовать своеобразное перемещение в пространстве-времени, что соответствует изменению объектов и их расположению в пространстве.

В системах геоинформационного мониторинга и аналитики [9, 10] геоинформационная пространственно-временная модель расширяется путем добавления информационных слоев, описывающих изменение заданного набора параметров в пространстве и времени. Эти слои используются как для визуализации многомерных данных, так и для их аналитической обработки средствами статистического анализа. Каждый слой может описывать изменения, как отдельного параметра, так и агрегированного показателя, обобщающего разную информацию, поступающую из нескольких источников. В результате формируется пространственно-временная модель.

Проблемы моделирования и анализа данных активно изучаются в современной теории геоинформационных систем [11 – 13]. С точки

зрения компьютерных наук понятие «пространство» занимает особое место [14]. Параметр «время» часто вводится как дополнительный в модели, учитывающей отраслевые особенности предметной области исследования [15, 16]. Этот параметр играет существенную роль при сборе географических данных в режиме реального времени, например, в дистанционном зондировании Земли [17, 18].

Также отмечается высокая актуальность решения задачи внедрения ГИС в нефтяной и газовой отраслях [19], где динамика изменения параметров ниже, но сложность их представления возрастает благодаря появлению новых зависимостей между объектами.

В зарубежной литературе вопросам развития современных геоинформационных систем и их приложений уделяется особое внимание [20, 21]. Логическому отношению пространства и времени в геоинформационном пространстве посвящена статья [22]. Отмечается, что время и пространство интегрированы и равноправны в формировании потоков, событий и процессов. Такой подход позволяет по-новому рассматривать проблемы моделирования и системного анализа в рамках вычислительной геоинформатики [23].

В статье [24] предложен новый «временно-географический» аналитический подход к решению задачи моделирования процессов взаимодействия людей и животных в пространстве и времени. В работе [25] представлена модель данных, интегрированных по времени, для геологического картирования. Особое внимание уделено концептуализации взаимоотношений между пространством и временем при моделировании геологических данных.

Статья [26] описывает результаты моделирования времени, изменений и темпоральности в ГИС. Моделирование движения в пространстве открывает новые возможности восприятия геоинформационной среды и построения дополнительных измерений в цифровых представлениях пространственно-временных встреч. Особенности интеграции временных параметров в ГИС посвящена работа [27].

О роли времени в географических исследованиях, способах его представления в предметно-ориентированных моделях, детализации измерений времени и тесной интеграции пространства и времени сообщается в обзоре [28]. Моделирование времени связано с представлением событий [29, 30]. Отмечается, что интеграция времени в ГИС требует всестороннего анализа пространственно-временных явлений. Для решения этой задачи использована концептуальная событийно-ориентированная модель времени и пространства, подходящая для географических приложений.

Выделение событий как отдельных сущностей позволяет упорядочить моменты изменения характеристик объектов и территорий во времени и реализовать их хранение в реляционной базе данных. Темпоральная логика на основе событий как отдельных концептов обеспечивает инструментарий для адаптивной дискретизации процессов, при которой поток событий имеет неравномерную дискретизацию, определяемую темпо-ритмом изменения обстановки, например, в режиме реального времени.

В системах бизнес-аналитики [31, 32] геоинформационное пространство отображается в виде информационных срезов, посвященных описанию отдельных аспектов рассматриваемой проблемы. Обычно такое представление осуществляется в виде двумерных графиков на специализированных виджетах, отображающих динамику изменения отдельных параметров, либо взаимные корреляционные зависимости между несколькими параметрами, позволяющими выявить причинно-следственные связи. Информационные виджеты обычно привязываются визуально к пространственным объектам, сгруппированным по слоям в ГИС.

Такое представление обстановки схематично выглядит в виде графа (сети) заданного вида, наложенного поверх геоинформационной подложки (карты) с дополнительными справочными контекстными графиками, описывающими динамику изменения ситуации во времени. Оно позволяет применять современные алгоритмы оптимизации и алгоритмы теории графов [33, 34].

Описанный подход широко используется в ГИС транспортной и авиационной логистики и на железнодорожном транспорте, в системах дистанционного зондирования Земли, в ситуационных центрах и мониторинговых платформах регионального и государственного управления и многих других практических приложениях [35 – 37]. Однако применение задач бизнес-аналитики для поддержки принятия управленческих решений затруднено в силу большого объема отображаемых данных и сложности настройки такой системы для каждого отдельного случая. На практике такие системы широко применяются для мониторинга обстановки, сбора информации и подготовки аналитических отчетов, однако решать на их основе оптимизационные задачи затруднительно в силу отсутствия единой методики выполняемых операций.

В [38] отмечается растущая тенденция к использованию ГИС в качестве одного из компонентов системы поддержки принятия решений для семантической интеграции данных и использования экспертного интеллекта. Необходимость организации доступа

в системах поддержки принятия решений к различным наборам пространственных данных, а также к функциональным возможностям ГИС и средствам анализа многокритериальных решений отмечается и в [39]. В [40] показано, что включение платформы ГИС в систему планирования распределительной сети может значительно повысить производительность труда планировщиков.

Таким образом, геоинформационные многослойные пространственно-временные модели предназначены для решения задач визуализации данных не только в системах распределенного мониторинга, бизнес-аналитики, экспертной оценки и информационной поддержки принятия решений, но и в системах управления организационными системами. Однако, для этой области необходимо решение задач многокритериальной оптимизации, что затруднено в силу сложности и неоднородности данных, полученных с различными временными характеристиками и сгруппированных по множеству пространственных слоёв. Для решения этой задачи в практических приложениях можно предложить следующую модель.

**3. Комбинированная геоинформационная многослойная пространственно-временная модель.** Опишем каждую точку в геоинформационном пространстве в локальной системе координат в виде переменной

$$\gamma_k = \{x_k, y_k, z_k\}, \quad (1)$$

где  $x_k, y_k, z_k$  – пространственные координаты точки в трехмерном пространстве в локальной (местной) системе координат.

Отметим, что точка  $\gamma_k$  также может быть задана в глобальной системе координат в виде тройки географических координат: широта  $\phi_k$ , долгота  $\lambda_k$  и высота  $h_k$ :

$$\gamma_k = (\phi_k, \lambda_k, h_k). \quad (2)$$

Рассмотрим каждый информационный слой, отображаемый средствами геоинформационной системы, в виде множества значений параметров для показателя  $w_i$ :

$$v_{i,j,k} = v_{i,j,k} (w_i, t_{i,j,k}, g_{i,j,k}), \quad (3)$$



где  $v_{i,j,k}$  – значение параметра, принадлежащее одному из множеств: пустое множество, множество целых чисел, множество вещественных чисел, множество символов, множество строк;

$t_{i,j,k}$  – временная метка контроля (измерения);

$g_{i,j,k}$  – отметка о наличии значения параметра в геоинформационном пространстве  $\{\gamma_k\}$ .

Будем считать, что для существующих, то есть известных значений параметра  $v_{i,j,k}$  отмечается  $g_{i,j,k} = true$ . Это означает существование узла координатной сетки в точке  $(\varphi_k, \lambda_k, h_k)$ . Отсутствующие значения параметра соответствуют  $g_{i,j,k} = false$ .

Комбинированная геоинформационная многослойная пространственно-временная модель может быть задана множеством точек  $v_{i,j,k}$ . Однако, в геоинформационных системах некоторые параметры обстановки, т.е. значений параметров в определенный момент времени с учетом пространственного расположения, удобно задавать относительно заданной области пространства, или региона  $G_m = \{\gamma_k, \tau_{k,m}\}$ ,  $m = 1..N_G$  в виде обобщенного параметра, формирующего свой геоинформационный слой для каждого показателя  $w_i$ :

$$V_{i,m,n} = \sum_{j,k} v_{i,j,k} \cdot \delta(g_{i,j,k} \in G_m) \cdot \delta(t_{i,j,k} \in \Delta T_n), \quad (4)$$

где  $\delta(x) = \begin{cases} 1, & \text{если } x = true, \\ 0, & \text{если } x = false. \end{cases}$ ;

$\Delta T_n$  – временной интервал контроля (актуальности) обобщенного показателя.

Таким образом, комбинированная геоинформационная многослойная пространственно-временная модель представляет собой граф, вершинами которого являются значения параметров  $v_{i,j,k}$ , упорядоченные по слоям, сформированным для каждого показателя  $w_i$ , с размещением временных меток в слое времени.

Дуги графа представляют логические связи значений параметров  $v_{i,j,k}$ , которые можно разделить в свою очередь на три типа:

- топологические – отношение территориальной или транспортной доступности, определяемые взаимным расположением отметок  $g_{i,j,k}$ ;
- семантические – отношение семантической близости параметров, определяющих их включение в один слой;
- хронологические – отношение следования событий контроля (измерения) параметров по времени  $t_{i,j,k}$ .

В качестве значений контролируемого показателя  $w_i$  может выступать, например, количество оказанных услуг, их стоимость, доходность, своевременность и т.п.

В предложенной модели геоинформационное пространство является дискретным и представляется в виде череды событий изменения значений параметров обстановки, каждое событие привязано к географической координате и моменту времени возникновения. Можно считать, что события изменения параметров упорядочены по времени в интервальной темпоральной логике.

Интервалы между событиями, описывающими изменение разных параметров или последовательное изменение одного параметра, могут быть разными, что позволяет описывать их с помощью моделей неэквидистантных временных рядов [41]. Эти модели могут описывать потерю информации о событии с заданной вероятностью, «дрожание» момента события по времени, что связано в основном с задержками фиксации изменения параметра, или задавать адаптивную дискретизацию временного ряда, описывающего изменение параметра, что свойственно процессам с изменяющейся динамикой развития обстановки.

Использование интервальной темпоральной логики, то есть учет временного аспекта при описании параметров, с одной стороны, позволяет достаточно гибко описывать изменения обстановки, уточняя последовательность событий в случае более высокой динамики и наоборот, увеличивая масштаб в случае малых изменений обстановки или их отсутствия. С другой стороны, такой подход требует реализации интервальной логики сравнения потоков событий между собой. В этом случае невозможно точно задать одновременность возникновения событий и их взаимосвязь, в связи с чем, предполагается устанавливать данные факты в случае, если

разница по времени между событиями не превышает определенной величины, установленной пропорционально масштабу наблюдений.

Предложенная модель позволяет описывать также несколько вариантов развития обстановки, что предоставляет достаточно удобный инструмент для имитационного моделирования. События, характеризующие разные варианты развития обстановки, могут выстраиваться в независимые или взаимосвязанные цепочки, что влияет на изменение выбранного набора параметров, описывающих ситуацию.

В этом случае, например, можно выделить отдельно сценарии негативного, позитивного или нейтрального варианта развития обстановки. Отметим, что неравномерная дискретизация событий по времени также здесь обеспечивает возможность описания этих сценариев с разной плотностью событий, соответствующей разной динамике.

Высокая динамика развития обстановки, фиксируемой в реальности или построенной в результате имитационного моделирования, отображается оператору в виде последовательно упорядоченных ситуаций. Это позволяет ему судить о скорости и важности принимаемых решений. Таким образом, дополнительно к основным параметрам, характеризующим текущую обстановку, могут быть заданы косвенные параметры, свидетельствующие о скорости ее изменения, адаптивной реакции на внешние события и взаимосвязи параметров. Например, небольшое событие может требовать всякий раз перестройки всей наблюдаемой картины параметров, что характеризует его важность и значимость для принятия решений.

Также различия динамики событий, описывающих разные сценарии развития обстановки при ее имитационном моделировании, могут характеризовать объективность их описания и достаточность данных для моделирования и оценки. Например, низкая плотность событий изменения обстановки в случае позитивного сценария может свидетельствовать о недостатке исходных данных, в то время как детальная проработка негативного варианта может быть слабо адаптирована к воздействию случайных факторов.

Реализация данной логики в рамках системы поддержки принятия решений или ситуационного центра управления обстановкой позволяет вырабатывать решения не только на основе изменяющихся значений параметров, но и корректировать расчеты значений этих параметров. В этом случае решения должны включать задания по сбору дополнительной информации, уточнению отдельных

фрагментов сценария мониторинга обстановки или ее моделирования, сравнительного сопоставления позитивных и негативных прогнозов.

**4. Постановка оптимизационной задачи с использованием комбинированной геоинформационной многослойной пространственно-временной модели.** Как было указано выше, предложенная модель позволяет формулировать универсальные постановки задач поддержки принятия решений для различных приложений геоинформатики в логистике, при управлении транспортными ресурсами, а также в ситуационных центрах управления предприятиями и регионами и системах бизнес-аналитики.

Например, представление системы взаимосвязанных показателей в виде графа востребовано при построении системы мониторинга развития региона. Сбор показателей по различным направлениям социально-экономического развития обеспечивает построение достаточно сложной сети взаимосвязанных индикаторов (показателей), использование которых для автоматизированного формирования организационно-управленческих решений традиционно представляет сложность.

Сопряжение и упорядочивание этих индикаторов, согласно предложенной модели, позволяет корректно поставить задачу оптимизации, а следовательно, решить проблему практического использования накапливаемой аналитики в организационном управлении.

В данном разделе рассмотрим общую задачу управления, а в следующем – приведем конкретный пример для одной из актуальных задач регионального управления.

Введем дополнительные обозначения. Пусть задан обобщенный параметр  $V_{[service],m_1,n_1} = V_{1,m_1,n_1}$ , с областью начальных значений  $V_{1,m_1,n_1}(0)$ , где  $t'_0 = t'_{1,m_1,n_1}(0)$  – начальная отметка времени. Определим целевой параметр как  $V_{[goal],m_2,N_2} = V_{2,m_2,N_2}$ , достижимый в момент времени  $t'_{2,m_2,N_2}$ , а также обобщенный параметр ограничения как  $V_{[constr],m_3,n_3} = V_{3,m_3,n_3}$ .

Сформулируем задачу оптимизации обобщенного параметра в виде:

$$F(w_1, t'_0) = \sum_{m_1, m_2=1}^{NG} \left| V_{[goal], m_2, N_2} - V_{[service], m_1, N_1} \right| \rightarrow \min, \quad (5)$$

$$\forall (m_1 = m_3, n_1 = n_3) : V_{1, m_1, n_1} \leq V_{[constr], m_3, n_3}.$$

Допустим, существует такая функциональная зависимость между  $V_{[service], m_1, n_1}$  и другим параметром,  $V_{[costs], m_4, n_4} = V_{4, m_4, n_4}$ , при которой величина  $V_{[costs], m_4, n_4}$  влияет на изменение  $V_{[service], m_1, n_1}$ . Пример такой зависимости – выделение финансирования на социальное обеспечение отдельных государственных услуг.

В этом случае решение задачи (5) может быть выполнено градиентным методом:

$$V_{[costs], m_4, n_4}^{(s+1)} = V_{[costs], m_4, n_4}^{(s)} - k_{m_4} \cdot \nabla \sum_{G_m} \left| V_{[service], m_1, n_1} \right|^2;$$

$$\sum_{m_1, m_2=1}^{NG} V_{[costs], m_4, n_4}^{(s+1)} \leq \Delta C_{n_4}^{(s+1)}(w_4); \quad (6)$$

$$\forall (m_1 = m_3, n_1 = n_3) : V_{1, m_1, n_1} \leq V_{3, m_3, n_3}.$$

Отметим, что в данной модели не обязательна синхронизация изменений  $V_{[costs], m_4, n_4}$  и  $V_{[service], m_1, n_1}$  по времени, так как справедливо только, что

$$\forall (i, m) : t'_{i, m, n}^{(s+1)} > t'_{i, m, n}^{(s)};$$

$$t'_{[service], m, n}^{(s)} \geq t'_{[costs], m, n}^{(s)}. \quad (7)$$

Решение также сохраняется при наличии множественных зависимостей  $V_{[service], m_1, n_1}$  от нескольких  $V_{[costs], m, n}$ .

Для решения данной задачи на практике упростим ее следующим образом. Обычно в системах организационного управления требуется на основе данных мониторинга по различным индикаторам или показателям эффективности выбрать, какие мероприятия наиболее важны для реализации. Например, в региональном управлении в условиях ограниченного финансирования речь идет о периодической расстановке приоритетов

между разными программами и проектами. При этом внешние условия на каждой итерации принятия решений также изменяются.

Пусть пространство целей также меняется во времени и задается на текущую дату. Тогда задачу можно преобразовать к виду:

$$F(w_1, t') = \sum_{m_1, m_2=1}^{N_G} \left| V_{[goal], m_2, n_2} - V_{[service], m_1, n_1} \right| \rightarrow \min, \quad (8)$$

$$\forall (m_1 = m_3, n_1 = n_3) : V_{1, m_1, n_1} \leq V_{[constr], m_3, n_3}.$$

Для такой постановки практически полезное решение может быть представлено в виде матрицы приоритетов:

$$P_{[service], m_1, n_1}^{(s+1)} = \frac{\left| V_{[goal], m_2, n_2}^{(s)} - V_{[service], m_1, n_1}^{(s-1)} \right| - \left| V_{[constr], m_3, n_3}^{(s)} - V_{[service], m_1, n_1}^{(s-1)} \right|}{\left| V_{[costs], m_4, n_4}^{(s)} - V_{[costs], m_4, n_4}^{(s-1)} \right|}. \quad (9)$$

**5. Задача планирования заявок с использованием комбинированной геоинформационной многослойной пространственно-временной модели.** Рассмотрим пример использования временной модели для оптимизации процесса обработки заявок по подключению жилых домов к системе газоснабжения в границах выбранного региона. Этот процесс называют газификацией, отдельно выделяется социальная догазификация, призванная подвести газ к домам в газифицированных населенных пунктах в соответствии с постановлением Правительства РФ от 13.09.2021 N 1547.

Процесс социальной догазификации может быть автоматизирован с использованием цифровой платформы интегрального мониторинга [42 – 44]. Цифровая платформа позволяет реализовать мониторинг и контроль заявок на подключение газа и выполняемых работ по их исполнению. При этом оптимизация работ состоит в консолидации заявок и планировании ресурсов.

На рисунке 1 представлен пример района подключений. Точками отмечены домовладения, требующие подключения к системе газоснабжения. Кластеризация точек подключения, например, по абрису преграды ландшафта, позволяет группировать работы по подключению и планировать их с учетом имеющихся ресурсов. Другим критерием для планирования работ является срок подключения: законодательством

установлены сроки подключения частных домов по программе догазификации, которые составляют до 9 месяцев.

Географическая обстановка в данном случае выступает в качестве системы ограничений, поскольку она влияет на стоимость и трудоемкость выполнения работ. Логично предположить, что последовательное освоение близких локаций более выигрышно, чем случайный проход распределенных в пространстве точек.

Фактор времени также ограничивает область допустимых решений. Регламент подключения определяет интервалы допустимых смещений в случае опережения или опоздания и задает интервальную темпоральную логику возможных объединений нескольких подключений между собой.

Планирование мероприятий по догазификации частных домовладений представляет собой многокритериальную задачу, которая может быть поставлена и решена в терминах предложенной модели.

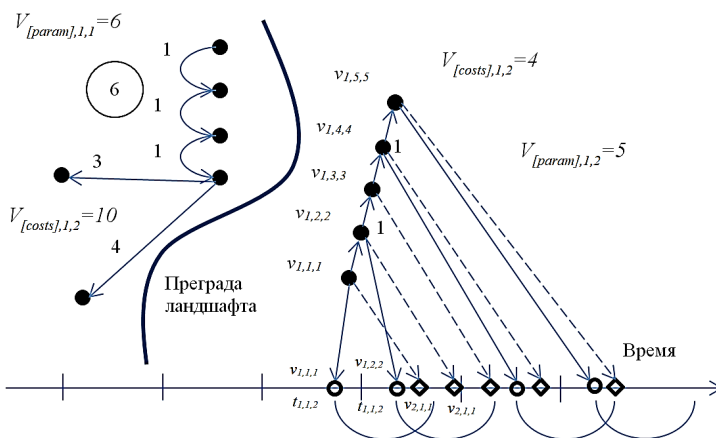


Рис. 1. Пример заявок на догазификацию

Для оптимизации стоимости подключения в рамках решения данной задачи предлагается организационная новация – планировать подключение газа домовладениям при отсутствии заявок в случае технологической оптимальности таких действий. Данный подход не лишен недостатков, в частности он требует проактивной позиции газовых служб по предложению услуг населению, что может привести к дополнительным расходам. Однако, с учетом общей цели полномасштабной газификации России, эти затраты компенсируются сокращением временных и материальных расходов за счет

комплексного подхода по социально-экономическому развитию территорий.

Обозначим пространство заявок на догазификацию  $v_{i,j,k}$ ,  $t_{i,j,k}$  – время поступления заявки,  $t'_{i,j,k}$  – плановое время исполнения заявки. Фактическое время исполнения может отличаться от планового, однако на рассматриваемом этапе принятия решений это можно не учитывать.

Тогда задачу планирования мероприятий по догазификации можно сформулировать в виде:

$$\begin{aligned} \sum_{m,n} V_{[service],m,n} &\rightarrow \max; \\ \sum_{m,n} V_{[costs],m,n} &\rightarrow \min; \end{aligned} \quad (10)$$

где  $V_{[service],m,n} = \sum_{j,k} e_{j,k} \cdot \delta(g_{i,j,k} \in G_m)$  – консолидация заявок;

$V_{[costs],m,n} = \sum_{j,k} c_{j,k} \cdot \delta(g_{i,j,k} \in G_m)$  – общая стоимость;

$e_{j,k} = v_{1,j,k}$  – заявки на подключение, единиц;

$c_{j,k} = v_{2,j,k}$  – стоимость подключения (перехода от одной заявки к другой), рублей;

$t'_{i,j,k} - t_{i,j,k} \leq \Delta\tau$  – ограничение на срок подключения;

$\Delta\tau$  – норматив на выполнение заявки, мес.

Таким образом, для решения задачи планирования заявок предлагается представить сеть подключений в виде модели, содержащей два пространственных слоя и один временной.

**6. Алгоритм планирования заявок с использованием комбинированной геоинформационной многослойной пространственно-временной модели.** Данная задача в указанной постановке может быть решена путем поиска групп вершин на графе и состоит из 2 частей.

Первая часть – группировка вершин по технологической доступности. Данная часть алгоритма является частным случаем задачи поиска компонент связности – следует разбить вершины графа на несвязные группы так, чтобы внутри группы можно было пройти от одной вершины до любой другой, а между группами пути



отсутствуют. Идея алгоритма стандартна – находим для каждой непройденной вершины все смежные непройденные. Шаги алгоритма:

Шаг 1. Выбираем любую непройденную вершину  $e_{j,k}$ ;

Шаг 2. Помечаем ее как пройденную;

Шаг 3. Каждую смежную вершину добавляем в компонент связности  $V_{1,m,n}$ , переходим на нее;

Шаг 4. Повторяем Шаги 2 и 3 пока остаются переходы;

Шаг 5. Переходим к Шагу 1 пока не пометим все вершины.

Вторая часть – топологическая сортировка по времени и стоимости, выполняется для каждой группы (компоненты связности)  $V_{1,m,n}$ .

Вариант 1 (ASAP, срочное распределение, при котором критерием выступает наименьшее время реализации заявки):

Шаг 1. Устанавливаем для всех вершин  $t'_{i,j,k} = t_{i,j,k}$ ;

Шаг 2. Переходим на самую раннюю вершину группы;

Шаг 3. Переходим в следующую смежную вершину по ребру наименьшей стоимости  $c_{j,k}$ ;

Шаг 4. Корректируем  $t'_{i,j,k} \rightarrow \min$ , при условии  $t'_{i,j,k} - t_{i,j,k} \leq \Delta\tau$  если для вершины существует заявка  $e_{j,k} = 1$ ;

Шаг 5. Повторяем Шаги 3 и 4 пока остаются переходы;

Шаг 6. Переходим к Шагу 2 пока не пометим все вершины;

Шаг 7. Фиксируем для группы  $V_{[costs],m,n}$ .

Вариант 2 (JIT, точное распределение, при котором критерием выступает полное исполнение заявок):

Шаг 1. Устанавливаем для всех вершин  $t'_{i,j,k} = t_{i,j,k} + \Delta\tau$ ;

Шаг 2. Переходим на самую позднюю вершину группы;

Шаг 3. Переходим в следующую смежную вершину по ребру наименьшей стоимости  $c_{j,k}$ ;

Шаг 4. Корректируем  $t'_{i,j,k} \rightarrow \max$ , при условии  $t'_{i,j,k} - t_{i,j,k} \leq \Delta\tau$  если для вершины существует заявка  $e_{j,k} = 1$ ;

Шаг 5. Повторяем Шаги 3 и 4 пока остаются переходы;

Шаг 6. Переходим к Шагу 2 пока не пометим все вершины;

Шаг 7. Фиксируем для группы  $V_{[costs],m,n}$ .

Выбор между вариантами реализации второго этапа зависит от плотности заявок на участке (в выбранном регионе). При высокой

плотности следует минимизировать риски, в этом случае необходимо выбрать первый вариант срочного распределения. Однако, в этом случае возможны отказы от выполнения заявок и опоздания. При низкой плотности заявок лучше подходит второй вариант, так как он позволяет выполнить подключение для собственников при отсутствии заявок с целью оптимизации общей стоимости догазификации региона. На рисунке 2 показано отличие при реализации данных алгоритмов в количестве запланированных заявок в формируемом плане на газификацию.

Аналогичная логика будет полезна при имитационном моделировании изменения обстановки под влиянием внешних событий. Задержки и опоздания, нехватка человеческих и материальных ресурсов могут привести к смещению сроков влево для чего можно либо предусмотреть дополнительные интервалы времени, либо разработать графики опережения или компенсации возникающих отставаний.

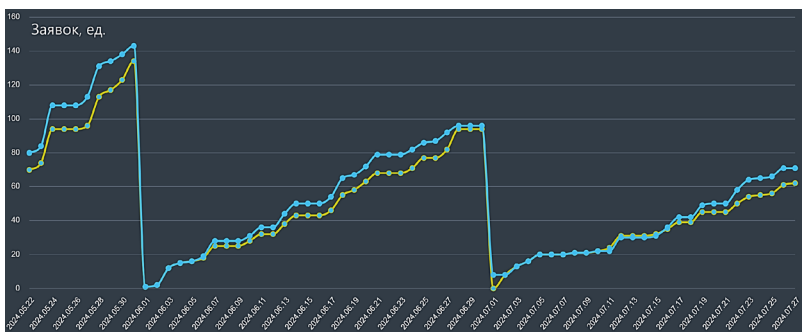


Рис. 2. План-график количества заявок на газификацию: желтый – вариант 1, бирюзовый – вариант 2

В предложенной модели это будет соответствовать появлению новых дополнительных ветвей обстановки, которые, с одной стороны, потребуют поиска дополнительной информации для их построения, а, с другой стороны, обеспечат заблаговременную проработку вариантов управленческих решений. Каждая ветвь будет иметь собственную частоту дискретизации, адаптивную к темпо-ритму событий.

Алгоритм планирования заявок в пространстве-времени может быть рекомендован для реализации в различных ситуационных центрах и системах мониторинга как в роли инструмента решения задач оптимизации, так и в качестве средства имитационного

моделирования возможных решений по оптимизации по времени использования географических распределенных ресурсов.

**7. Практическая реализация.** Предложенные решения были использованы в рамках внедрения цифровой платформы интегрального мониторинга (разработчик – ООО «Открытый код», г. Самара) для цифровой трансформации процессов сбора, анализа и визуализации данных коммунальных ресурсов [44]. Реализация предложенной модели потребовала проведения интеграции с источниками данных и функциональными пользователями цифровой платформы: региональным оператором газификации, эксплуатирующим собственную информационную систему с поддержкой механизма API, газораспределительными организациями региона, ведущими реестр газифицируемых объектов с использованием электронных таблиц, а также с информационными системами многофункциональных центров через Единую систему межведомственного электронного взаимодействия. Кроме того, для представления в модели данных от жителей потребовалось подключение к portalу Госуслуги с аутентификацией через Единую систему идентификации и аутентификации. Поскольку первичные данные не содержали в явном виде координат газифицируемых объектов, выполнялось прямое геокодирование текстовой адресной информации. Подключение к источникам данных осуществлялось с помощью встроенных в цифровую платформу интегрального мониторинга средств интеграции.

Интерфейс программного обеспечения цифровой платформы приведен на рисунке 3.

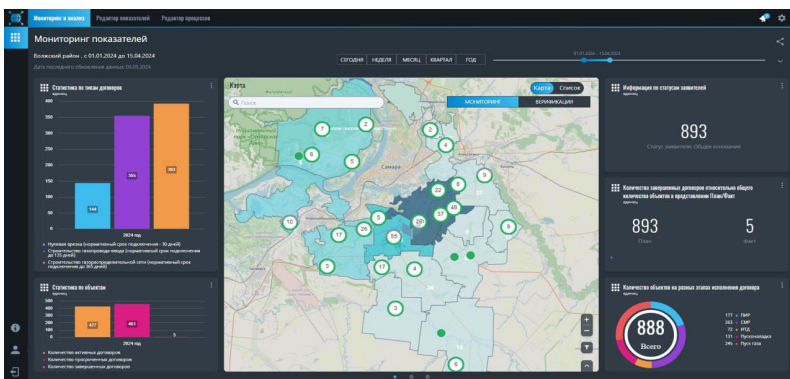


Рис. 3. Фрагмент графического интерфейса системы поддержки принятия решений по ресурсоснабжению населения на основе цифровой платформы интегрального мониторинга

В программе реализованы функциональные возможности геокодирования объектов на интерактивной карте, сбора, обработки и анализа данных по мероприятиям ресурсоснабжения; контроля исполнения полного жизненного цикла процессов по обеспечению населения коммунальными ресурсами; визуального отображения данных в режиме дашбордов и решения оптимизационных задач на основании агрегации данных.

Пользовательский интерфейс решения организован в форме интерактивной аналитической панели (дашборда), содержащей взаимно сопряженные графические приложения (виджеты), отражающие разные параметры развития обстановки. На электронной карте в центре экрана отображается геоинформационное пространство принятия решений с разбиением данных по информационным слоям. Отдельно выделены регионы обеспечения и точки спроса.

Применение предложенной модели позволяет не только пересмотреть способ формализации геоинформационных данных, но и по-новому организовать пользовательский интерфейс с учетом современных тенденций научной визуализации. С одной стороны, новые возможности включают конфигурируемость, интерактивность и адаптивность пользовательского интерфейса, с другой стороны, – они обеспечивают задание достаточно строгого и понятного пользователю регламента работы с данными, обеспечивающими результативность информационной поддержки принятия решений.

Последний тезис представляет собой важное преимущество для систем обработки информации и управления, построенных на базе геоинформационных систем. Проблема их использования на практике состоит в том, что они предоставляют пользователю избыточно большой объем данных и набор функциональных возможностей по их обработке, что затрудняет использование систем лицами, принимающим решения. Несколько решить проблему позволяет организационная мера по созданию вокруг геоинформационной системы аналитического центра, обеспечивающего лицо, принимающее решения, информационными справками и докладами, однако это серьезно усложняет и повышает стоимость внедрения платформы.

Для решения этой проблемы необходимо, как ни странно, сократить перечень возможностей системы и редуцировать набор отображаемых высшему руководству данных, обеспечив их интегральность и интерпретируемость для принятия управленческих решений. Такую возможность дает сценарный подход, когда для текущей ситуации вырабатывается несколько вариантов развития

обстановки, отличающихся разными внешними условиями. Например, может быть реализовано моделирование развития ситуации для позитивного, негативного и реалистичного сценария.

Предлагаемый подход позволяет реализовать краткосрочный прогноз изменения выбранных показателей во времени в соответствии с меняющейся обстановкой. При этом от руководителя – лица, принимающего решение, требуется выбор одного из нескольких вариантов, что реально реализовать в рамках ситуационного центра.

Общая концепция реализации цифровой платформы близка к современным системам класса Business Intelligence (BI). Как и BI системы, цифровая платформа интегрального мониторинга обеспечивает перевод транзакционной деловой информации в визуальную форму, удобную для восприятия человеком. Основные интерфейсы платформы по типу использования включают цифровую интерактивную карту (построенную на базе геоинформационной системы), масштабируемые графики и диаграммы, отображающие основные зависимости параметров между собой, а также прогнозы их изменения во времени.

Панели управления интерфейсом содержат наборы виджетов, описывающий по-отдельности каждую область мониторинга, а также сборные дашборды, соответствующие решаемым задачам или контролируемым ситуациям. Ситуационные наборы виджетов формируются динамически для конкретных решаемых задач и отображаются в верхней части пользовательского интерфейса. Таким образом, обеспечивается привлечение внимания пользователя к ключевым изменениям параметров и влияющим факторам, прогнозируя исход различных вариантов действий.

Комбинированная геоинформационная многослойная пространственно-временная модель позволяет подобрать масштаб отображаемой обстановки, как в территориальном, так и во временном аспекте, в соответствии с поставленной задачей и темпом изменения событий. Отметим, что в случае многопользовательского режима работы с цифровой платформой интегрального мониторинга, масштаб обстановки для каждого пользователя может быть разным, что обеспечивает согласование мнений в условиях различной осведомленности от отличий точек зрения.

Методика мониторинга обстановки в условиях использования предложенной модели состоит, таким образом, в последовательном выполнении семи действий или этапов: 1) отслеживание текущей ситуации; 2) корректировка оптимизационной задачи или постановка новой; 3) моделирование реалистичного решения задачи;

4) построение позитивного сценария и моделирование обстановки в его случае; 5) построение негативного сценария и моделирование обстановки в его случае; 6) формирование справки-отчета; 7) задание интервала контроля, по истечении которого все действия повторяются.

Предложенная методика позволила организовать сбор и обработку данных по обеспеченности населения коммунальными ресурсами на основе процессного подхода. Реализация возможностей бизнес-аналитики в предложенной модели позволила согласовать разные масштабы сбора и обработки информации в пространственном и временном аспектах на уровне подразделений коммунальных служб, аналитиков ситуационного центра и руководства региона в условиях высокой динамики изменения ситуации. В результате появилась возможность оптимизировать работы по обеспечению ресурсами в заданных временных ограничениях и в соответствии с отраслевыми регламентами.

В системе была предусмотрена возможность консолидации запросов на оказание услуг по ресурсоснабжению планирования работ по их комплексному исполнению. Применение предложенной модели в данной системе позволяет не только решить задачу оптимального планирования порядка исполнения заявок с учетом территориальной доступности и технологических возможностей сетей ресурсоснабжения, но и визуализировать информацию в рамках поддержки принятия решений по социально-экономическому развитию региона.

**8. Заключение.** Современные возможности геоинформационных технологий позволяют по-новому взглянуть на задачу организационного управления. Моделирование обстановки на карте с учетом динамики изменения событий становится мощным инструментом мониторинга, поскольку обеспечивает упорядочивание и согласование разнородной информации, поступающей из различных источников, и отражающей разные аспекты наблюдений. Реализация предложенной модели позволяет при этом перейти от мониторинга обстановки к поддержке принятия решений за счет реализации возможности моделирования различных сценариев развития обстановки и возможности решения оптимизационных задач.

Предложенная модель может быть рекомендована в качестве базиса для автоматизации поддержки принятия управленческих решений в системах бизнес-аналитики больших данных и ситуационных центрах. Опыт успешного ее применения в цифровой платформе интегрального мониторинга для аналитики и оптимизации ресурсоснабжения подтверждает правильность выбранного подхода.

Следует отметить, что непосредственное внедрение такого рода систем требует дополнительных усилий по их встраиванию в общий процесс сбора и обработки информации и выработки решений.

Дальнейшие исследования связаны с сопряжением предложенной модели с компонентами искусственного интеллекта для генерации слоев на базе нейросетевой кластеризации и классификации. Применение предложенной модели в этом случае позволяет решить важную задачу расширения объема обучающей выборки, накапливая данные не только по реальным обстановкам, но и по случаям негативного и позитивного развития событий, что обеспечивает необходимый и достаточный объем данных для решения задач классификации. Перспективным представляется также расширение области применения модели при решении широкого круга современных задач организационного управления.

### Литература

1. Gospodinov S.G. Geoinformatics as a science of space // *European Journal of Technology and Design*. 2022. vol. 10(1). pp. 3–8.
2. Бескид П.П., Куракина Н.И., Орлова Н.В. Геоинформационные системы и технологии // СПб.: Российский государственный гидрометеорологический университет, 2010. 173 с.
3. Массер Д. Что есть пространство-время? // *В мире науки*. 2018. № 8-9. С. 78–82.
4. Rynasiewicz R. Newton's Views on Space, Time, and Motion / Ed.: Edward N. Zalta // *The Stanford Encyclopedia of Philosophy*. 2012.
5. Ухтомский А.А. Доминанта // М.: АСТ, 2022. 320 с.
6. Флоренский П.А. Анализ пространственности и времени в художественно-изобразительных произведениях // М.: Прогресс, 1993. 324 с.
7. Мещеряков Б.Г., Зинченко В.П. Большой психологический словарь. Изд. 4-е, расш. // М.: АСТ, 2009. 811 с.
8. Бахтин М.М. Формы времени и хронотопа в романе. Очерки по исторической поэтике // *Литературно-критические статьи*. М.: Худ. лит., 1986. С. 121–290.
9. Корнилков С.В., Рыбникова Л.С., Рыбников П.А., Смирнов А.Ю. Геоинформационный мониторинг для решения экологических задач горнопромышленных территорий Среднего Урала // *Горная промышленность*. 2022. № S1. С. 127–133.
10. Имамов Ф.Н. Вахтеров А.Р., Решетников А.Г. Геоинформационная система для мониторинга и аналитического контроля локальных магнитных аномалий // *Молодежный Вестник УГАТУ*. 2021. № 1 (24). С. 17–24.
11. Цветков М.В., Смирнова О.В., Гальяно Ф.Р. Система мониторинга ледовой обстановки и обеспечения безопасного судоходства по северному морскому пути на базе интеллектуальной ГИС // *Труды СПИИРАН*. 2014. № 5(36). С. 28–43. DOI: 10.15622/sp.36.2.
12. Косяков С.В., Гадалов А.Б., Садыков А.М. Моделирование пространственных данных при решении задач дискретной оптимизации в среде ГИС // *Информационные технологии*. 2012. № 7. С. 27–31.
13. Воробьева Г.Р., Воробьев А.В., Орлов Г.О. Концепция обработки, анализа и визуализации геофизических данных на основе элементов тензорного исчисления // *Информатика и автоматизация*. 2024. Т. 23(2). С. 572–604.

14. Попович В.В., Ведешин Л.А. Понятие "пространство" в географических информационных системах // Системы высокой доступности. 2019. Т. 15. № 2. С. 47–56.
15. Кузьмин В.А., Шаныгин С.И., Чунин С.А., Никитин Г.С., Мкртчян М.Э., Каурова З.Г., Орехов Д.А., Цыганов А.В., Айдиев А.Б., Мищенко Н.В., Ачилов В.В. Моделирование пространственно-временных данных об окружающей среде в ГИС // Нормативно-правовое регулирование в ветеринарии. 2022. № 3. С. 43–50.
16. Рыжиков А.И. Картография, ГИС и время // Геодезия и картография. 2006. № 7. С. 43–47.
17. Маризе-Томокала Г.Де.Г., Сисей Ф.К. Данные дистанционного зондирования Земли // Научный Лидер. 2023. № 19(117). С. 21–33.
18. Куфтинова Н.Г. Проблемы интеллектуального анализа данных при моделировании транспортных потоков мегаполиса // Мир транспорта. 2020. Т. 18. № 5. С. 24–40.
19. Опоку Ф.К. Роль цифровых технологий в повышении безопасности и эффективности нефте- и газопроводов в России // Аллея науки. 2024. Т. 1. № 1(88).
20. de Lange N. Geoinformation Systems // Geoinformatics in Theory and Practice. 2023. pp. 375–433. DOI: 10.1007/978-3-662-65758-4\_9.
21. Merry K., Bettinger P., Crosby M., Boston K. Geographic information systems // Geographic Information System Skills for Foresters and Natural Resource Managers. 2023. pp. 1–23. DOI: 10.1016/B978-0-323-90519-0.00007-8.
22. Yuan M. Relationships between Space and Time // Geographic Information Science & Technology Body of Knowledge. 2020. DOI: 10.22224/gistbok/2020.3.7.
23. Yuan M. From representation to geocomputation: some theoretical accounts of geographic information science // New Thinking in GIScience. 2022. pp. 1–8.
24. Dodge S., Su R., Johnson J., Simcharoen A., Goulias K., Smith J.L.D., Ahearn S.C. ORTEGA: An object-oriented time-geographic analytical approach to trace space-time contact patterns in movement data // Computers, Environment and Urban Systems. 2021. vol. 88. DOI: 10.1016/j.compenvurbsys.2021.101630.
25. van Gasselst S., Nass A. Time and geology in geographic information systems: some formalism for describing temporal events // Journal for Geographic Information Science. 2015. vol. 3. pp. 199–208. DOI: 10.1553/giscience2015s199.
26. Shields R. Bergson's GIS: experience, time and memory in geographical information systems // Media Theory. Geospatial Memory. 2018. vol. 2(1). pp. 316–332.
27. Ott T., Swiaczny F. Implementation of time in GIS // Time-Integrative Geographic Information Systems. 2001. pp. 77–126. DOI: 10.1007/978-3-642-56747-6\_4.
28. Dai J., An L. 1.21 –Time geography // Reference Module in Earth Systems and Environmental Sciences. Comprehensive Geographic Information Systems. 2017. pp. 303–312. DOI: 10.1016/B978-0-12-409548-9.09625-1.
29. Claramunt C., Thériault M. Managing time in GIS: an event-oriented approach // Recent Advances in Temporal Databases, Proceedings of the International Workshop on Temporal Databases. 1995. pp. 23–42. DOI: 10.1007/978-1-4471-3033-8\_2.
30. Wang M., Zhang J., Cao Y., Li S., Chen M. A study on a spatiotemporal entity-based event data model // ISPRS International Journal of Geo-Information. 2024. vol. 13(10). DOI: 10.3390/ijgi13100360.
31. Домрачева А.А., Сайбель Н.Ю. Применение геоинформационных систем в бизнес-анализе // Взаимодействие науки и общества: проблемы и перспективы. 2016. С. 73–75.



32. Chaudhuri S. Application of web-based geographical information system (GIS) in e-business // *Digital Marketing and Consumer Engagement: Concepts, Methodologies, Tools, and Applications*. IGI Global, 2018. pp. 649–665.
33. Кормен Т., Лейзерсон Ч., Ривест Р., Штайн К. Алгоритмы: построение и анализ // пер. с англ. М.: Вильямс, 2019. 1296 с.
34. Кристофидес Н. Теория графов. Алгоритмический подход // М.: Мир, 1978. 434 с.
35. Воробьев А.В., Воробьева Г.Р. Подход к динамической визуализации разнородных геопространственных векторных изображений // *Компьютерная оптика*. 2024. Т. 48. № 1. С. 123–138.
36. Плякин А.В., Орехова Е.А., Огарков Д.И. Геоинформационные системы для обеспечения транспортно-логистической деятельности // *Интеллектуальная логистика*. 2021. С. 149–153.
37. Оюунханд Б., Касьянова Е.Л. Использование ДЗЗ и ГИС при создании географических основ для тематических карт // *Вестник СГУГиТ*. 2021. Т. 26. № 5. С. 119–125.
38. Sowmiya Narayanan K.J., Manimaran A. Recent developments in geographic information systems across different application domains: a review // *Knowledge and Information Systems*. 2024. vol. 66. pp. 1523–1547.
39. Jelokhani-Niaraki M., Moradi-Pour S., Samany N.N., Mohammadkhan S. A multiple models-multiple users group GIS-based decision support system for land use problems // *Land Use Policy*, 2023. vol. 134. DOI: 10.1016/j.landusepol.2023.106916.
40. Wang D. Research and development of intelligent decision support system for distribution network planning based on GIS // *IEEE 3rd International Conference on Data Science and Computer Application (ICDSCA)*. 2023. pp. 786–791.
41. Прохоров С.А. Прикладной анализ неэквидистантных временных рядов. Самара: СГАУ, 2001. 375 с.
42. Сурнин О.Л., Ситников П.В., Иващенко А.В., Головнин О.К., Дубинина И.Н., Додонова Е.А. Применение цифровой платформы интегрального мониторинга как средства бизнес-аналитики социально-экономического развития региона // Конференция «Информационные технологии в управлении» (ИТУ-2022). Сборник материалов. СПб.: СПбГЭТУ «ЛЭТИ», 2022. С. 158–161.
43. Ivaschenko A., Dubinina I., Golovnin O., Golovnina A., Sitnikov P. Digital integrated monitoring platform for intelligent social analysis // *Communications in Computer and Information Science*. Springer, Cham. 2023. vol. 1909. pp. 365–376. DOI: 10.1007/978-3-031-44615-3\_25.
44. Свидетельство о государственной регистрации программы для ЭВМ № 2023616797 РФ. Цифровая платформа интегрального мониторинга ресурсоснабжения: № 2023614655: заявл. 14.03.2023: опубл. 03.04.2023 / заявитель ООО «Открытый код».

**Иващенко Антон Владимирович** — д-р техн. наук, профессор, директор, передовая медицинская инженерная школа, Самарский государственный медицинский университет Министерства здравоохранения Российской Федерации. Область научных интересов: управление в организационных системах, мультиагентные технологии, медицинская инженерия. Число научных публикаций — 400. anton.ivashenko@gmail.com; улица Чапаевская, 89, 443099, Самара, Россия; р.т.: +7(846)374-1004.

**Головнин Олег Константинович** — д-р техн. наук, доцент, заведующий кафедрой, кафедра медицинской физики, математики и информатики, Самарский государственный медицинский университет Министерства здравоохранения Российской Федерации. Область научных интересов: системный анализ, геоинформационные системы,

интеллектуальные транспортные системы. Число научных публикаций — 250. golovnin@bk.ru; улица Чапаевская, 89, 443099, Самара, Россия; р.т.: +7(846)374-1004.

**Головнина Анастасия Александровна** — канд. техн. наук, доцент кафедры, кафедра информационных систем и технологий, Самарский национальный исследовательский университет имени академика С.П. Королева. Область научных интересов: системный анализ, геоинформационные системы, искусственный интеллект, обработка больших данных. Число научных публикаций — 85. anastasiya.stolbova@bk.ru; Московское шоссе, 34, 443086, Самара, Россия; р.т.: +7(846)267-4370.

**Додонова Евгения Александровна** — аспирант, передовая медицинская инженерная школа, Самарский государственный медицинский университет Министерства здравоохранения Российской Федерации. Область научных интересов: управление в организационных системах, мультиагентные технологии, медицинская инженерия. Число научных публикаций — 20. dodonova.evg@gmail.com; улица Чапаевская, 89, 443099, Самара, Россия; р.т.: +7(846)374-1004.

A. IVASCHENKO, O. GOLOVNIN, A. GOLOVNINA, E. DODONOVA  
**AN ASSEMBLED MODEL OF MULTILAYER GEOINFORMATION  
SPACE-TIME**

*Ivaschenko A., Golovnin O., Golovnina A., Dodonova E. An Assembled Model of Multilayer Geoinformation Space-Time.*

**Abstract.** Modern organizational management technologies involve collecting and processing large amounts of data to calculate the parameters of the functioning of the objects and processes under study. Since the main feature of the collected parameters is their binding to territories, on the one hand, and attribution to time periods, on the other hand, the use of geographic information systems and technologies is required. Despite the development of modern geographic information technologies, the issues of their practical application to support decision-making, taking into account the combined influence of spatial and temporal factors, have not been fully resolved. The article proposes an assembled model of geoinformation multilayer space-time, which is a graph whose vertices are the parameter values ordered by layers with the placement of time marks in the time layer, and the arcs describe the relations between them that are divided into three types: topological, semantic and chronological. Conjugation and ordering of parameters, according to the proposed model, allows you to correctly pose and solve the optimization problem, and, consequently, eliminate the problem of the practical use of accumulated analytics in the processes of supporting management decision-making. The proposed model is used in the digital platform of integral monitoring for the digital transformation of the processes of collecting, analyzing and visualizing utility resource data. The general management task is considered, and a specific example is given for one of the urgent tasks of regional management, i.e. social gasification, in which the optimization of the process of processing applications for connecting residential buildings to the gas supply system within the boundaries of the selected region is carried out. The assembled model of geoinformation multilayer space-time allows formulating universal statements of decision support problems for various geoinformatics applications in logistics, transport resource management, as well as in situational centers for enterprise and regional management, business analytics systems and organizational systems management.

**Keywords:** geographic information system, geographic information model, spatial model, spatio-temporal model, spatial optimization, integral monitoring, GIS.

## References

1. Gospodinov S.G. Geoinformatics as a science of space. *European Journal of Technology and Design*. 2022. vol. 10(1). pp. 3–8.
2. Beskid P.P., Kurakina N.I., Orlova N.V. *Geoinformacionnye sistemy i tehnologii [Geographic information systems and technologies]*. Saint Peterburg: Rossijskij gosudarstvennyj gidrometeorologičeskij universitet, 2013. 173 p. (In Russ.).
3. Masser D. [What is space-time?]. *V mire nauki – In the world of science*. 2018. no. 8-9. pp. 78–82. (In Russ.).
4. Rynasiewicz R. *Newton's Views on Space, Time, and Motion / Ed.: Edward N. Zalta. The Stanford Encyclopedia of Philosophy*. 2012.
5. Uhtomskij A.A. *Dominanta [Dominant]*. Moscow: AST, 2022. 320 p. (In Russ.).
6. Florenskij P.A. *Analiz prostranstvennosti i vremeni v hudozhestvenno-izobrazitel'nyh proizvedenijah [Analysis of spatiality and time in artistic and visual works]*. Moscow, Progress, 1993. 324 p (In Russ.).

7. Meshherjakova B.G., Zinchenko V.P. Bol'shoj psihologicheskij slovar' [Large psychological dictionary]. Moscow: AST, 2009. 811 p. (In Russ.).
8. Bahtin M.M. [Forms of Time and Chronotope in the Novel. Essays on Historical Poetics]. Literaturno-kriticheskie stat'i – Literary and critical articles. 1986. pp. 121–290. (In Russ.).
9. [Geoinformation monitoring for solving environmental problems of mining areas of the Middle Urals]. Gornaja promyshlennost' – Mining industry. 2022. no. S1. pp. 127–133. (In Russ.).
10. Imamov F.N., Vahterov A.R., Reshetnikov A.G. [Geoinformation system for monitoring and analytical control of local magnetic anomalies]. Molodezhnyj Vestnik UGATU – UGATU Youth Bulletin. 2021. no. 1(24). pp. 17–24. (In Russ.).
11. Tsvetkov M., Smirnova O., Galiano P. [Ice Situation and Safe Navigation Monitoring System on the Northern Sea Route Based on Intelligent GIS]. Trudy SPIIRAN – SPIIRAS Proceedings. 2014. no. 5(36). pp. 28–43. DOI: 10.15622/sp.36.2. (In Russ.).
12. Kosyakov S.V., Gadalov A.B., Sadykov A.M. [Modeling spatial data in solving discrete optimization problems in a GIS environment]. Informacionnye tekhnologii – Information technology. 2012. no. 7. pp. 27–31. (In Russ.).
13. Vorobeva G., Vorobev A., Orlov G. [The Concept of Processing, Analysis and Visualization of Geophysical Data Based on Elements of Tensor Calculus]. Informatika i avtomatizaciya – Informatics and Automation. 2024. vol. 23(2). pp. 572–604. (In Russ.).
14. Popovich V.V., Vedeshin L.A. [The concept of "space" in geographic information systems]. Sistemy vysokoj dostupnosti – High availability systems. 2019. vol. 15. no. 2. pp. 47–56. (In Russ.).
15. Kuz'min V.A., SHanygin S.I., Chunin S.A., Nikitin G.S., Mkrtychyan M.E., Kaurova Z.G., Orekhov D.A., Cyganov A.V., Ajdiev A.B., Mishchenko N.V., Achilov V.V. [Modeling of spatio-temporal environmental data in GIS]. Normativno-pravovoe regulirovanie v veterinarii – Regulatory and legal regulation in veterinary medicine. 2022. no. 3. pp. 43–50. (In Russ.).
16. Ryzhikov A.I. [Cartography, GIS and time]. Geodeziya i kartografiya. 2006. no. 7. pp. 43–47. (In Russ.).
17. Marize-Tomokala G.De.G., Sisej F.K. [Earth remote sensing data]. Nauchnyj Lider – Scientific Leader. 2023. no. 19(117). pp. 21–33. (In Russ.).
18. Kuftinova N.G. [Problems of data mining in modeling metropolitan transport flows]. Mir transporta – The world of transport. 2020. vol. 18. no. 5. pp. 24–40. (In Russ.).
19. Opoku F.K. [The role of digital technologies in improving the safety and efficiency of oil and gas pipelines in Russia]. Alleya nauki – Alley of Science. 2024. vol. 1. no. 1(88). (In Russ.).
20. de Lange N. Geoinformation Systems. Geoinformatics in Theory and Practice. 2023. pp. 375–433. DOI: 10.1007/978-3-662-65758-4\_9.
21. Merry K., Bettinger P., Crosby M., Boston K. Geographic information systems. Geographic Information System Skills for Foresters and Natural Resource Managers. 2023. pp. 1–23. DOI: 10.1016/B978-0-323-90519-0.00007-8.
22. Yuan M. Relationships between Space and Time. Geographic Information Science & Technology Body of Knowledge. 2020. DOI: 10.22224/gistbok/2020.3.7.
23. Yuan M. From representation to geocomputation: some theoretical accounts of geographic information science. New Thinking in GIScience. 2022. pp. 1–8.
24. Dodge S., Su R., Johnson J., Simcharoen A., Goulias K., Smith J.L.D., Ahearn S.C. ORTEGA: An object-oriented time-geographic analytical approach to trace space-time contact patterns in movement data. Computers, Environment and Urban Systems. 2021. vol. 88. DOI: 10.1016/j.compenvurbsys.2021.101630.

25. van Gasselt S., Nass A. Time and geology in geographic information systems: some formalism for describing temporal events. *Journal for Geographic Information Science*. 2015. vol. 3. pp. 199–208. DOI: 10.1553/giscience2015s199.
26. Shields R. Bergson's GIS: experience, time and memory in geographical information systems. *Media Theory. Geospatial Memory*. 2018. vol. 2(1). pp. 316–332.
27. Ott T., Swiaczyn F. Implementation of time in GIS. *Time-Integrative Geographic Information Systems*. 2001. pp. 77–126. DOI: 10.1007/978-3-642-56747-6\_4.
28. Dai J., An L. 1.21 –Time geography. *Reference Module in Earth Systems and Environmental Sciences. Comprehensive Geographic Information Systems*. 2017. pp. 303–312. DOI: 10.1016/B978-0-12-409548-9.09625-1.
29. Claramunt C., Thériault M. Managing time in GIS: an event-oriented approach. *Recent Advances in Temporal Databases, Proceedings of the International Workshop on Temporal Databases*. 1995. pp. 23–42. DOI: 10.1007/978-1-4471-3033-8\_2.
30. Wang M., Zhang J., Cao Y., Li S., Chen M. A study on a spatiotemporal entity-based event data model. *ISPRS International Journal of Geo-Information*. 2024. vol. 13(10). DOI: 10.3390/ijgi13100360.
31. Domracheva A.A., Sajbel' N.Ju. [Application of Geographic Information Systems in Business Analysis]. *Vzaimodejstvie nauki i obshchestva: problemy i perspektivy – Interaction between science and society: problems and prospects*. 2016. pp. 73–75. (In Russ.).
32. Chaudhuri S. Application of web-based geographical information system (GIS) in e-business. *Digital Marketing and Consumer Engagement: Concepts, Methodologies, Tools, and Applications*. IGI Global, 2018. pp. 649–665.
33. Kormen T., Lejzerson Ch., Rivest R., Shtajn K. *Algoritmy: postroenie i analiz [Algorithms: construction and analysis]*. Moscow, Vil'jams, 2019, 1296 p. (In Russ.).
34. Kristofides N. *Teoriya grafov. Algoritmicheskij podhod [Graph Theory. Algorithmic Approach]*. Moscow, Mir, 1978, 434 p. (In Russ.).
35. Vorob'ev A.V., Vorob'eva G.R. [An approach to dynamic visualization of heterogeneous geospatial vector images]. *Komp'yuternaja optika – Computer Optics*. 2024. vol. 48. no. 1. pp. 123–138. (In Russ.).
36. Pljakin A.V., Orehova E.A., Ogarkov D.I. [Geographic information systems for transport and logistics activities]. *Intellektual'naja logistika – Intelligent logistics*. 2021. pp. 149–153. (In Russ.).
37. Ojuunhand B., Kas'janova E.L. [Using Remote Sensing and GIS to Create Geographical Bases for Thematic Maps]. *Bulletin of the SGUGiT Vestnik SGUGiT*. 2021. vol. 26. no. 5. pp. 119–125. (In Russ.).
38. Sowmiya Narayanan K.J., Manimaran A. Recent developments in geographic information systems across different application domains: a review. *Knowledge and Information Systems*. 2024. vol. 66. pp. 1523–1547.
39. Jelokhani-Niaraki M., Moradi-Pour S., Samany N.N., Mohammadkhan S. A multiple models-multiple users group GIS-based decision support system for land use problems. *Land Use Policy*, 2023. vol. 134. DOI: 10.1016/j.landusepol.2023.106916.
40. Wang D. Research and development of intelligent decision support system for distribution network planning based on GIS. *IEEE 3rd International Conference on Data Science and Computer Application (ICDSCA)*. 2023. pp. 786–791.
41. Prokhorov S.A. *Prikladnoj analiz neekvidistantny'x vremenny'x ryadov [Applied analysis of non-equidistant time series]*. Samara, Samara State Aerospace University, 2001. 375 p. (In Russ.).
42. Surmin O.L., Sitnikov P.V., Ivashhenko A.V., Golovnin O.K., Dubinina I.N., Dodonova E.A. *Primenenie cifrovoj platformy integral'nogo monitoringa kak sredstva biznes-analitiki social'no-jekonomicheskogo razvitiya regiona [Application of a digital platform for integrated monitoring as a means of business analytics of the socio-*

- economic development of a region] Konferencija «Informacionnye tehnologii v upravlenii» (ITU-2022). Sbornik materialov [Conference "Information Technologies in Management" (ITU-2022). Collection of materials]. SPb.: SPbGJeTU «LJeTb», 2022. pp. 158–161. (In Russ.).
43. Ivaschenko A., Dubinina I., Golovnin O., Golovnina A., Sitnikov P. Digital integrated monitoring platform for intelligent social analysis. Communications in Computer and Information Science. Springer, Cham. 2023. vol. 1909. pp. 365–376. DOI: 10.1007/978-3-031-44615-3\_25.
44. Cifrovaja platforma integral'nogo monitoringa resursosnabzhenija [Digital platform for integrated resource supply monitoring]. Russian patent no. 2023616797, 2023. (In Russ.).

**Ivaschenko Anton** — Ph.D., Dr.Sci., Professor, Director, Higher school of medical engineering, Samara State Medical University. Research interests: management in organizational systems, multi-agent technologies, medical engineering. The number of publications — 400. anton.ivashenko@gmail.com; 89, Chapaevskaya St., 443099, Samara, Russia; office phone: +7(846)374-1004.

**Golovnin Oleg** — Ph.D., Dr.Sci., Associate Professor, Head of the department, Head of the department of medical physics, mathematics and informatics, Samara State Medical University. Research interests: systems analysis, geographic information systems, intelligent transport systems. The number of publications — 250. golovnin@bk.ru; 89, Chapaevskaya St., 443099, Samara, Russia; office phone: +7(846)374-1004.

**Golovnina Anastasia** — Ph.D., Associate professor of the department, Department of information systems and technologies, Samara National Research University. Research interests: systems analysis, geographic information systems, artificial intelligence, big data processing. The number of publications — 85. anastasiya.stolbova@bk.ru; 34, Moskovskoe Hwy., 443086, Samara, Russia; office phone: +7(846)267-4370.

**Dodonova Evgeniya** — Postgraduate student, Higher school of medical engineering, Samara State Medical University. Research interests: management in organizational systems, multi-agent technologies, medical engineering. The number of publications — 20. dodonova.evg@gmail.com; 89, Chapaevskaya St., 443099, Samara, Russia; office phone: +7(846)374-1004.

## Руководство для авторов

Взаимодействие автора с редакцией осуществляется через личный кабинет на сайте журнала «Информатика и автоматизация» <http://ia.spcras.ru/>. При регистрации авторам рекомендуется заполнить все предложенные поля данных. Подготовка статьи ведется с помощью текстовых редакторов MS Word 2007 и выше или LaTeX. Объем основного текста (до раздела Литература) - от 20 до 30 страниц включительно. Переносы запрещены. Номера страниц не проставляются. Основная часть текста статьи разбивается на разделы, среди которых являются обязательными: введение, хотя бы один «содержательный» раздел и заключение. Допускается также мотивированное содержанием и структурой материал а выделение подразделов. В основную часть опускается помещать рисунки, таблицы, листинги и формулы. Правила их оформления подробно рассмотрены на нашем сайте в разделе «Руководство для авторов».

## Author guidelines

Interaction between each potential author and the Editorial board is realized through the personal account on the website of the journal "Informatics and Automation" <http://ia.spcras.ru/>. At the registration the authors are requested to fill out all data fields in the proposed form. The submissions should be prepared using MS Word 2007, LaTeX. The text of the paper in the main part should not exceed 30 pages. Pages are not numbered; hyphenations are not allowed. Certain figures, tables, listings and formulas are allowed in the main section, and their typography is considered in more detail at the journal web.

---

Signed to print 28.03.2025. Passed for print 01.04.2025.

Printed in Publishing center GUAP.

Address: 67 litera A, B. Morskaya, St. Petersburg, 190000, Russia

---

Founder and Publisher: SPC RAS.

Address: 39 litera A, 14th Line V.O., St. Peterburg, 199178, Russia.

The journal is registered in the Federal Service for Supervision of Communications, Information Technology, and Mass Media, Registration Certificate (registration number) ПИ № ФС77-79228 dated September 25, 2020 Subscription Index П5513, Russian Post Catalog

---

Подписано к печати 28.03.2025. Дата выхода в свет 01.04.2025.

Формат 60×90 1/16. Усл. печ. л. 20,6. Заказ № 91. Тираж 300 экз., цена свободная.

Отпечатано в Редакционно-издательском центре ГУАП.

Адрес типографии: Б. Морская, д. 67, лит. А, г. Санкт-Петербург, 190000, Россия

---

Учредитель и издатель: СПб ФИЦ РАН.

Адрес учредителя и издателя: 14-я линия В.О., д. 39, лит. А, г. Санкт-Петербург, 199178, Россия

Журнал зарегистрирован Федеральной службой по надзору в сфере связи, информационных технологий и массовых коммуникаций, свидетельство о регистрации (регистрационный номер) ПИ № ФС77-79228 от 25 сентября 2020 г.

Подписной индекс П5513 по каталогу «Почта России»