

ISSN 2663-4228 (Print)
ISSN 2663-4236 (Online)

Вестник Донецкого национального университета



НАУЧНЫЙ
ЖУРНАЛ

*Основан в 1997 году,
«Серия Г: Технические науки»
основана в 2018 году*

Серия Г:
Технические
науки

4/2024

Редакционная коллегия журнала
«Вестник Донецкого национального университета. Серия Г: Технические науки»

Главный редактор – д-р техн. наук, проф. **В.В. Данилов**.

Зам. главного редактора – канд. техн. наук, доц. **И.А. Третьяков**.

Ответственный секретарь – канд. техн. наук **Е.Д. Пометун**.

Члены редколлегии: д-р техн. наук, проф. **Г.В. Аверин**; д-р техн. наук, проф. **Л.В. Алексейчик**; канд. физ.-мат. наук, с.н.с. **П.В. Асланов**; канд. техн. наук, доц. **Т.С. Башева**; д-р техн. наук, доц. **И.А. Бершадский**; д-р техн. наук, проф. **А.Б. Бирюков**; д-р техн. наук, проф. **С.В. Борщевский**; д-р техн. наук, доц. **К.А. Вытовтов**; д-р техн. наук, проф. **С.А. Горожанкин**; канд. техн. наук, доц. **С.В. Гридин**; канд. техн. наук, доц. **Т.В. Ермоленко**; канд. техн. наук, доц. **Т.В. Завадская**; д-р техн. наук, доц. **А.В. Звягинцева**; д-р техн. наук, доц. **С.А. Зори**; д-р техн. наук, доц. **В.В. Карнаух**; канд. техн. наук, с.н.с. **А.А. Курушин**; канд. техн. наук, доц. **И.А. Молоковский**; д-р техн. наук, проф. **Н.Г. Насонкина**; канд. техн. наук, доц. **В.В. Паслён**; д-р техн. наук, проф. **А.С. Петров**; д-р хим. наук, проф. **А.И. Сердюк**; академик РАН, д-р техн. наук, проф. **И.А. Соколов**; канд. физ.-мат. наук, доц. **В.И. Тимченко**; д-р техн. наук, проф. **В.К. Толстых**; д-р техн. наук, проф. **М.Н. Чепцов**; д-р техн. наук, доц. **Д.В. Шевцов**; канд. техн. наук, доц. **О.Г. Шелехова**.

The Editorial Board of the journal
«Bulletin of Donetsk National University. Series G: Technical Sciences»

The Editor-in-Chief – Dr. of tech., prof. **V.V. Danilov**.

The Deputy of the Editor-in-Chief – Cand. of tech., doc. **I.A. Tretiakov**.

Assistant Editor – Cand. of tech. **E.D. Pometun**.

The Members of the Editorial Board: Dr. of tech., prof. **G.V. Averin**; Dr. of tech., prof. **L.V. Alekseichik**; Cand. of phys. and math., s. researcher **P.V. Aslanov**; Cand. of tech., doc. **T.S. Bashevaia**; Dr. of tech., doc. **I.A. Bershadskii**; Dr. of tech., prof. **A.B. Biriukov**; Dr. of tech., prof. **S.V. Borshevskii**; Dr. of tech., doc. **K.A. Vytovtov**; Dr. of tech., prof. **S.A. Gorozhankin**; Cand. of tech., doc. **S.V. Gridin**; Cand. of tech., doc. **T.V. Ermolenko**; Cand. of tech., doc. **T.V. Zavadskaia**; Dr. of tech., doc. **A.V. Zviagintseva**; Dr. of tech., doc. **S.A. Zori**; Dr. of tech., doc. **V.V. Karnaukh**; Cand. of tech., s. researcher **A.A. Kurushin**; Cand. of tech., doc. **I.A. Molokovskii**; Dr. of tech., prof. **N.G. Nasonkina**; Cand. of tech., doc. **V.V. Paslen**; Dr. of tech., prof. **A.S. Petrov**; Dr. of chem., prof. **A.I. Serdiuk**; RAS Academician, Dr. of tech., prof. **I.A. Sokolov**; Cand. of phys. and math., doc. **V.I. Timchenko**; Dr. of tech., prof. **V.K. Tolstykh**; Dr. of tech., prof. **M.N. Sheptsov**; Dr. of tech., doc. **D.V. Shevtsov**; Cand. of tech., doc. **O.G. Shelekhova**.

Адрес редакции: 283001, Российская Федерация, Донецкая Народная Республика, г. Донецк, пр. Театральный, д. 13.

Тел.: +7 (856) 302 07 81

E-mail: vestnik_g@mail.ru

URL: <http://donnu.ru/vestnikG>

С 12.02.2024 г. журнал включен в Перечень рецензируемых научных изданий, в которых должны быть опубликованы основные научные результаты диссертаций на соискание ученой степени кандидата наук, на соискание ученой степени доктора наук (**Перечень ВАК РФ**) по следующим научным специальностям и соответствующей им отрасли (технические науки):

- 2.2.9. Проектирование и технология приборостроения и радиоэлектронной аппаратуры,
- 2.2.11. Информационно-измерительные и управляющие системы,
- 2.2.13. Радиотехника, в том числе системы и устройства телевидения,
- 2.2.15. Системы, сети и устройства телекоммуникаций,
- 2.3.1. Системный анализ, управление и обработка информации, статистика,
- 2.3.2. Вычислительные системы и их элементы,
- 2.3.3. Автоматизация и управление технологическими процессами и производствами,
- 2.3.5. Математическое и программное обеспечение вычислительных систем, комплексов и компьютерных сетей,
- 2.3.8. Информатика и информационные процессы,
- 2.4.1. Теоретическая и прикладная электротехника,
- 2.4.2. Электротехнические комплексы и системы,
- 2.4.6. Теоретическая и прикладная теплотехника.

Журнал включен в национальную информационно-аналитическую систему **РИНЦ** (Лицензионный договор № 378-06/2016 от 24.06.2016 г. в ред. от 27.03.2019 г.): http://elibrary.ru/title_about_new.asp?id=69984

Журнал включен в международную индексационную наукометрическую базу **Index Copernicus International Journals Master List**: <http://journals.indexcopernicus.com/search/details?id=63327>

*Печатается по решению Ученого совета ФГБОУ ВО «Донецкий государственный университет»:
Протокол №13 от 03.12.2024 г.*

**Вестник
Донецкого
национального
университета**

Серия Г: Технические науки

НАУЧНЫЙ ЖУРНАЛ

ОСНОВАН В 1997 ГОДУ,
СЕРИЯ Г ОСНОВАНА В 2018 ГОДУ

Номер посвящён памяти члена
редколлегии журнала, кандидата
физико-математических наук,
доцента Владимира Ивановича
Тимченко

№ 4/2024

СОДЕРЖАНИЕ

ЭЛЕКТРОНИКА, ФОТОНИКА, ПРИБОРОСТРОЕНИЕ И СВЯЗЬ

- Наумов М.А., Карповский А.Ю., Курносков В.Г.* Исследование влияния длины посылки на расстояние, время и скорость передачи данных по беспроводной связи LORA 5
- Алексейчик Л.В., Курушин А.А.* Антенны на диэлектрических резонаторах 19
- Акапьев В.Л., Савотченко С.Е.* Совершенствование структурной модели и параметров компактного Фурье-спектрометра для анализа газового состава атмосферы 27
- Казаков Е.А., Остапенков П.С., Соловьев Е.Д., Щукин А.В.* Оценка моделей полевых транзисторов для широкополосных усилителей класса D 35
- Пометун Е.Д., Лебедев В.Н.* Измерение интенсивности турбулентности по упрощенной методике градуировки проволочного датчика термоанемометра 47

ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ И ТЕЛЕКОММУНИКАЦИИ

- Подлесный В.Ю., Данилов В.В., Третьяков И.А., Колесник Е.В., Тимченко В.И.* Оптимальное дискретное представление непрерывных измерительных процессов в АСУ И АСНИ 52
- Нижникова О.Л., Дорохина Г.В., Большакова С.А.* Цифровые платформы для онлайн-консультирования: сравнительный анализ и перспективы 63
- Глушанков Е.С., Гольцев А.С., Мельник А.-В.В., Мироненко А.Б.* Построение компьютерной модели решения задач теории потенциала с помощью МГЭ 77
- Елисеев В.О., Бондаренко В.И.* Методы проектирования архитектуры модели интеллектуального ассистента для образовательной организации 87
- Криворучко К.А., Максименко И.И.* Распознавание недостоверной информации в СМИ с помощью нейронных сетей 101
- Секерин А.В., Кудинов В.А.* Алгоритм классификации идеологем на основе наборов цепочек религиозно-политических ассоциаций 110
- Нестругина Е.С., Бельский Р.С.* Использование искусственного интеллекта в строительстве 121
- Булыгин В.В.* Кибернетический подход к пониманию языка 132

<i>Дмитрюк Т.Г., Зори С.А.</i> Системный анализ функциональных особенностей предприятия как объекта управления	137
<i>Евдачѳва Е.Д., Максименко И.И.</i> Детектирование трендов модной одежды с обучением нейросети	154
<i>Ермоленко Т.В., Шалун Н.А.</i> Фильтрация спама методами глубокого обучения	165
<i>Чепцов М.Н., Лозинская В.Н., Климов В.В.</i> Обоснование комплексного показателя качества работы транспортной сети мобильного оператора связи	175
<i>Киреев В.А., Бобынцев Д.О.</i> Исследование работы системы помощи водителю при перестроении на основе радарного датчика	183

ЭНЕРГЕТИКА И ЭЛЕКТРОТЕХНИКА

<i>Малеваный М.В., Коновалов Д.А., Кожухов Н.Н., Черниченко В.В.</i> Однофазный конвективный теплоперенос в компактном теплообменнике, заполненном анизотропной пористой средой	193
<i>Батухтин А.Г., Рудой В.И., Савиных Н.К., Кузютин К.А.</i> Микроскопическое изучение форм мелкодисперсных частиц веществ, применяемых в технических агрегатах	201
<i>Бершадский И.А., Мых А.Д., Згарбул А.В., Якимшина В.В.</i> Методика анализа аварийных режимов работы контактных соединений электросети 0,4/0,22 кВ	207
<i>Бирюков А.Б., Гаращенко В.Н.</i> Анализ различных схем использования тепловой энергии: микроуровень теплотребления	218
<i>Давыдов А.И., Каштанов А.Л., Подгорная С.О.</i> Методология анализа небаланса электрической энергии на тягу поездов	228
<i>Начкебия Н.С., Бирюков А.Б., Асламова Я.Ю.</i> Моделирование температурного поля насадок регенеративных горелок	234
<i>Городько В.В., Городько О.Ю., Гридин С.В.</i> Перспективы применения современных силовых электронных компонентов в модернизации станочного парка Российской Федерации	242

ИНФОРМАЦИОННАЯ РУБРИКА

Памяти Владимира Ивановича Тимченко	251
Правила для авторов	253

Bulletin of Donetsk National University

SCIENTIFIC JOURNAL

FOUNDED IN 1997,
SERIES G FOUNDED IN 2018

Number dedicated to the memory
of a member of editorial board
of the journal, Candidate of
Physico-mathematical Sciences,
Associate Professor Vladimir
Ivanovich Timchenko

Series G: Technical sciences

№ 4/2024

CONTENTS

ELECTRONICS, PHOTONICS, INSTRUMENT ENGINEERING AND COMMUNICATION

- Naumov M.A., Karpovskii A.IU., Kurnosov V.G.* Studying the effect of the length of a parcel on the distance, time and speed of data transmission via LORA wireless communication 5
- Alekseichik L.V., Kurushin A.A.* Antennas on dielectric resonator 19
- Akapev V.L., Savotchenko S.E.* Improvement of the structural model and parameters of a compact Fourier spectrometer for analysis of the gas composition of the atmosphere 27
- Kazakov E.A., Ostapenkov P.S., Solovev E.D., SHCHukin A.V.* Models analysis of switching mosfets for wideband D class amplifiers 35
- Pometun E.D., Lebedev V.N.* Measurement of turbulence intensity using a simplified calibration method for a wire sensor of a thermoanemometer 47

INFORMATION TECHNOLOGY AND TELECOMMUNICATIONS

- Podlesnyi V.IU., Danilov V.V., Tretiakov I.A., Kolesnik E.V., Timchenko V.I.* Optimal discrete representation of continuous measurement processes in ACS and ASRS 52
- Nizhnikova O.L., Dorokhina G.V., Bolshakova S.A.* Digital platforms for online consulting: comparative analysis and prospects 63
- Glushankov E.S., Goltsev A.S., Melnik A.-V.V., Mironenko A.B.* Construction of a computer model for solving problems of potential theory using BEM 77
- Eliseev V.O., Bondarenko V.I.* Methods of designing the architecture of an intelligent assistant model for an educational organization 87
- Krivoruchko K.A., Maksimenko I.I.* Recognition of incorrect information in the media using neural networks 101
- Sekerin A.V., Kudinov V.A.* Classification algorithm of ideologemes based on sets of chains of religious and political associations 110
- Nestrugina E.S., Belskii R.S.* Use of artificial intelligence in construction 121
- Bulygin V.V.* Cyberenic approach to understanding language 132

<i>Dmitriuk T.G., Zori S.A.</i> System analysis of the enterprise functional features as a control object	137
<i>Evdacheva E.D., Maksimenko I.I.</i> Detecting fashion trends with neural network training	154
<i>Ermolenko T.V., SHalun N.A.</i> Spam filtering using deep learning methods	165
<i>CHeptsov M.N., Lozinskaia V.N., Klimov V.V.</i> Mobile communication operator's transport network quality comprehensive indicator justification	175
<i>Kireev V.A., Bobyntsev D.O.</i> Experimental study of the method of emergency detection by radar sensor during realignment	183

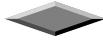
ENERGETICS AND ELECTRICAL ENGINEERING

<i>Malevanyi M.V., Konovalov D.A., Kozhukhov N.N., CHernichenko V.V.</i> Single-phase convective heat transfer in a compact heat exchanger filled with anisotropic porous medium	193
<i>Batukhtin A.G., Rudoi V.I., Savinykh N.K., Kuziutin K.A.</i> Microscopic study of the shapes of fine particles of substances used in technical aggregates	201
<i>Bershadskii I.A., Mykh A.D., Zgarbul A.V., IAKimishina V.V.</i> The method of analyzing emergency modes of operation of contact connections of the 0.4/0.22 kV power grid	207
<i>Biriukov A.B., Garashchenko V.N.</i> Analysis of various schemes for the use of thermal energy: micro-level of heat consumption	218
<i>Davydov A.I., Kashtanov A.L., Podgornaia S.O.</i> Methodology for analyzing the imbalance of electrical energy for train traction	228
<i>Nachkebiia N.S., Biriukov A.B., Aslamova IA.IU.</i> Modeling the temperature field of nozzles of regenerative burners	234
<i>Gorodko V.V., Gorodko O.IU., Gridin S.V.</i> Prospects for the use of modern power electronic components in the modernization of the machine tool fleet of the Russian Federation	242

INFORMATION SECTION

In memory of Vladimir Ivanovich Timchenko	251
Instructions for authors	253

Э Л Е К Т Р О Н И К А , Ф О Т О Н И К А ,
П Р И Б О Р О С Т Р О Е Н И Е И С В Я З Ъ



EDN SEDIYM
УДК 654.937

DOI 10.5281/zenodo.14514504

ИССЛЕДОВАНИЕ ВЛИЯНИЯ ДЛИНЫ ПОСЫЛКИ НА РАССТОЯНИЕ, ВРЕМЯ
И СКОРОСТЬ ПЕРЕДАЧИ ДАННЫХ ПО БЕСПРОВОДНОЙ СВЯЗИ LORA

© 2024 *Наумов М.А., Карповский А.Ю., Курносков В.Г.*

В этой статье анализируется влияние длины посылки на расстояние, время и скорость передачи данных по беспроводной связи LoRa. В статье исследуется влияние длины пакета - регулируемого параметра при передаче данных LoRa - на ключевые показатели производительности. Показано, что длина передаваемого пакета влияет на расстояние, пройденное сигналом, время, затрачиваемое на передачу, и общую скорость передачи данных. Приведены данные о нескольких экспериментах, проведенных в различных условиях. Представлен всесторонний анализ экспериментальных данных, в котором обращается внимание на четкие закономерности и корреляции, наблюдаемые в ходе эксперимента. Проведена аппроксимация экспериментальных данных и построены соответствующие графики.

Ключевые слова: беспроводная связь, длина посылки, расстояние передачи, время передачи, скорость передачи, LoRa, аппроксимация, ошибка.

Введение. Организация радиопередачи является классической задачей при проектировании систем управления подвижными машинами и механизмами. В горно-шахтном оборудовании, в том числе и при автоматизации очистных комбайнов (машин для выемки пластовых полезных ископаемых) подсистема радиопередачи должна обеспечивать выполнение своих функций как в зоне видимости комбайна, а также за её пределами. Увеличение дальности радиопередачи является особо важным, и позволяет повысить безопасность ведения работ по выемке угольных масс путем минимизации присутствия человека в опасной зоне. Новым решением на рынке радиосвязи выступает технология LoRa.

Цель исследований: получить функциональные зависимости, описывающие влияние длины посылки на расстояние, скорость и время передачи данных по беспроводному каналу связи LoRa.

Для полного изучения всех интересующих нас параметров беспроводной радиосвязи необходимо провести полный факторный эксперимент (ПФЭ), при котором выполняется перебор всех возможных сочетаний факторов, а именно мощность приемника, коэффициент распространения, скорость кодирования, скорость передачи данных, полоса пропускания, длина посылки. При этом количество опытов при полном факторном эксперименте определяется по формуле:

$$n = 2^k = 2^6 = 64 \text{ опыта,}$$

где k - количество факторов.

В ходе проведения исследований был осуществлен эксперимент, по результатам которого проведем необходимый анализ для изучения влияния длины посылки на качество беспроводной связи LoRa.

В таблице 1 приведены основные характеристики модуля LoRa SX1278 используемые при эксперименте [1]. В таблице 2 приведены результаты работы модуля LoRa SX1278 в связке с микроконтроллером при различных отношениях сигнал/шум и коэффициентах расширения спектра [2-4].

Таблица 1. Характеристики модуля LoRa SX1278 используемые при эксперименте

Характеристика	Значение
Напряжение питания	3.3 В
Частота	433 МГц
Интерфейс связи с микроконтроллером	SPI
Мощность передатчика	100 мВт
Коэффициент распространения	7
Полоса пропускания	91 кГц

Таблица 2. Результаты проведения экспериментов

Длина посылки, байт	Расстояние, м	Время передачи, мс	Скорость, байт/с	RSSI, dBm	Отношение сигнал/шум, dB
1	430	54	18,52	-116,4	2,82
4	423	54	18,52	-120	3,38
7	360	64	15,63	-116,38	2,86
11	356	75	13,33	-116,761	2,63
14	354	85	11,77	-114,02	1,74
16	343	95	10,53	-114,49	3,38

На рисунке 1 приведен график зависимости длины посылки от расстояния передачи данных (а), график зависимости длины посылки от времени (в мс) передачи данных (б) и график зависимости длины посылки от скорости (в байт/с) передачи данных (в).

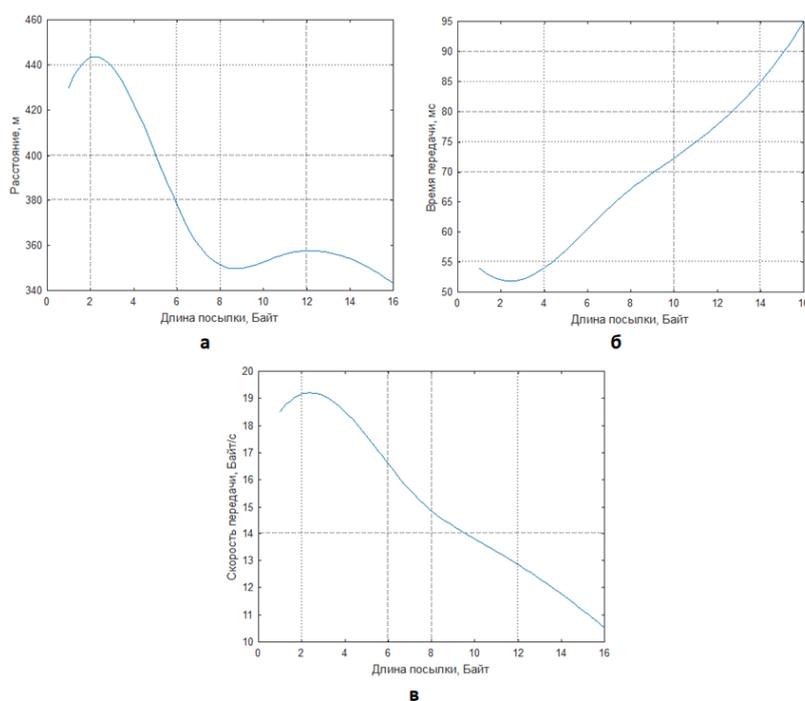


Рис. 1. График зависимости длины посылки от расстояния передачи данных (а), от времени передачи данных (б) и от скорости (в байт/с) передачи данных (в)

Графики, представленные на рисунке 1 построены по табличным экспериментальным значениям. В ходе исследования в открытых источниках информации не было найдено функциональных зависимостей, описывающих закон, по которым изменяются переменные расстояния, времени и скорости передачи данных в зависимости от длины передаваемой посылки по беспроводной связи LoRa. Для выведения функциональных зависимостей из табличных значений воспользуемся аппроксимацией [5].

Аппроксимация зависимости расстояния передачи данных от длины посылки. Аппроксимируем график, изображенный на рисунке 1 (а) с помощью линейной регрессии:

$$\hat{y} = a * x + b, \quad (1)$$

где x – длина посылки;

y – расстояние передачи данных.

Найдем коэффициент «а» по формуле (2):

$$a = \frac{\sum x_i * \sum y_i - n * \sum x_i y_i}{(\sum x_i)^2 - n * \sum x_i^2} = \frac{53 * 2266 - 6 * 19002}{53^2 - 6 * 639} = -5,94, \quad (2)$$

где x – длина посылки;

y – расстояние передачи данных;

n – количество проведенных опытов.

Найдем коэффициент «b» по формуле (3):

$$b = \frac{\sum x_i * \sum x_i y_i - \sum x_i^2 * \sum y_i}{(\sum x_i)^2 - n * \sum x_i^2} = \frac{53 * 19002 - 639 * 2266}{53^2 - 6 * 639} = 430,12, \quad (3)$$

где x – длина посылки;

y – расстояние передачи данных;

n – количество проведенных опытов.

Тогда уравнение (1) линейной регрессии, описывающее зависимость длины посылки от расстояния передачи данных, будет выглядеть следующим образом:

$$\hat{y} = -5,94 * x + 430,12.$$

Среднюю смещенную процентную ошибку аппроксимаций будем находить по формуле (4):

$$\bar{A} = \frac{1}{n} \sum \left| \frac{y_i - \hat{y}_i}{y_i} \right| * 100\%, \quad (4)$$

где \hat{y} – аппроксимированное значение расстояния передачи данных;

y – расстояние передачи данных;

n – количество проведенных опытов.

Аппроксимируем график, изображенный на рисунке 1 (а) с помощью квадратичной регрессии.

Уравнение квадратичной регрессии показано на формуле (5):

$$\hat{y} = a * x^2 + b * x + c, \quad (5)$$

где x – длина посылки;

y – расстояние передачи данных.

Найдем коэффициенты «а», «b» и «с» с помощью системы уравнений:

$$\begin{cases} a * \sum x_i^2 + b * \sum x_i + n * c = \sum y_i \\ a * \sum x_i^3 + b * \sum x_i^2 + \sum x_i = \sum x_i y_i \\ a * \sum x_i^4 + b * \sum x_i^3 + \sum x_i^2 = \sum x_i^2 y_i \end{cases} = \begin{cases} a * 639 + b * 53 + 6 * c = 2266 \\ a * 8579 + b * 639 + c * 53 = 19002 \\ a * 121251 + b * 8579 + c * 639 = 225106 \end{cases} = \begin{cases} a = 0,43 \\ b = -13,33 \\ c = 449,58 \end{cases}$$

где x – длина посылки;

y – расстояние передачи данных;

n – количество проведенных опытов.

Тогда уравнение (5) квадратичной регрессии, описывающее зависимость длины посылки от расстояния передачи данных, будет выглядеть следующим образом:

$$\hat{y} = 0,43 * x^2 - 13,33 * x + 449,58.$$

Аппроксимируем график, изображенный на рисунке 1 (а) с помощью кубической регрессии. Уравнение кубической регрессии показано на формуле (6).

$$\hat{y} = a * x^3 + b * x^2 + c * x + d, \quad (6)$$

где x – длина посылки;

y – расстояние передачи данных.

Найдем коэффициенты «а», «b», «с» и «d» с помощью системы уравнений:

$$\begin{cases} a * \sum x_i^3 + b * \sum x_i^2 + c * \sum x_i + n * d = \sum y_i \\ a * \sum x_i^4 + b * \sum x_i^3 + \sum x_i^2 + \sum x_i = \sum x_i y_i \\ a * \sum x_i^5 + b * \sum x_i^4 + \sum x_i^3 + \sum x_i^2 = \sum x_i^2 y_i \\ a * \sum x_i^6 + b * \sum x_i^5 + \sum x_i^4 + \sum x_i^3 = \sum x_i^3 y_i \end{cases} = \begin{cases} a * 8579 + b * 639 + c * 53 + 6 * d = 2266 \\ a * 121251 + b * 8579 + c * 639 + d * 53 = 19002 \\ a * 1765283 + b * 121251 + c * 8579 + d * 639 = 225106 \\ a * 26200059 + b * 1765283 + c * 121251 + d * 8579 = 3001122 \end{cases} = \begin{cases} a = 0,02 \\ b = 0,03 \\ c = -10,69 \\ d = 446,09 \end{cases}$$

где x – длина посылки;

y – расстояние передачи данных;

n – количество проведенных опытов.

Тогда уравнение (6) кубической регрессии, описывающее зависимость длины посылки от расстояния передачи данных, будет выглядеть следующим образом:

$$\hat{y} = 0,02 * x^3 + 0,03 * x^2 - 10,69 * x + 446,09.$$

Аппроксимируем график, изображенный на рисунке 1 (а) с помощью степенной регрессии. Уравнение степенной регрессии показано на формуле 7.

$$\hat{y} = a * x^b, \quad (7)$$

где x – длина посылки;

y – расстояние передачи данных.

Найдем коэффициент «b»:

$$b = \frac{n * \sum (\ln(x_i) * \ln(y_i)) - \sum \ln(x_i) * \sum \ln(y_i)}{n * \sum \ln^2(x_i) - (\sum \ln(x_i))^2} = \frac{6 * 65,6 - 11,14 * 35,58}{6 * 124,1 - 124,1^2} = -0,09,$$

где x – длина посылки;

y – расстояние передачи данных;

n – количество проведенных опытов.

Найдем коэффициент «а»:

$$a = \exp\left(\frac{1}{n} * \sum \ln(y_i) - \frac{b}{n} * \sum \ln(x_i)\right) = \exp\left(\frac{1}{6} * 35,58 + \frac{0,09}{6} * 11,14\right) = 441,71,$$

где x – длина посылки;

y – расстояние передачи данных;

n – количество проведенных опытов.

Тогда уравнение (7) степенной регрессии, описывающее зависимость длины посылки от расстояния передачи данных, будет выглядеть следующим образом:

$$\hat{y} = 441,71 * x^{-0,09}.$$

Аппроксимируем график, изображенный на рисунке 1 (а) с помощью показательной регрессии.

Уравнение показательной регрессии показано на формуле 8.

$$\hat{y} = a * b^x, \quad (8)$$

где x – длина посылки;

y – расстояние передачи данных.

Найдем коэффициент «b»:

$$b = \exp\left(\frac{n * \sum(x_i * \ln(y_i)) - \sum x_i * \sum \ln(y_i)}{n * \sum x_i^2 - (\sum x_i)^2}\right) = \exp\left(\frac{6 * 311,65 - 53 * 35,58}{6 * 639 - 2809}\right) = 0,99,$$

где x – длина посылки;

y – расстояние передачи данных;

n – количество проведенных опытов.

Найдем коэффициент «а»:

$$a = \exp\left(\frac{1}{n} * \sum \ln(y_i) - \frac{\ln(b)}{n} * \sum x_i\right) = \exp\left(\frac{1}{6} * 35,58 + \frac{0,01}{6} * 53\right) = 430,87,$$

где x – длина посылки;

y – расстояние передачи данных;

n – количество проведенных опытов.

Тогда уравнение (8) показательной регрессии, описывающее зависимость длины посылки от расстояния передачи данных, будет выглядеть следующим образом:

$$\hat{y} = 430,87 * 0,99^x.$$

Аппроксимируем график, изображенный на рисунке 1 (а) с помощью логарифмической регрессии.

Уравнение логарифмической регрессии показано на формуле (9).

$$\hat{y} = a + b * \ln(x), \quad (9)$$

где x – длина посылки;

y – расстояние передачи данных.

Найдем коэффициент «b»:

$$b = \frac{n * \sum(y_i * \ln(x_i)) - \sum \ln(x_i) * \sum y_i}{n * \sum \ln^2(x_i) - (\sum \ln(x_i))^2} = \frac{6 * 4025,8 - 11,14 * 2266}{6 * 124,1 - 124,1} = -33,59,$$

где x – длина посылки;
 y – расстояние передачи данных;
 n – количество проведенных опытов.

Найдем коэффициент «a» по формуле 10.

$$a = \frac{1}{n} * \sum y_i - \frac{b}{n} * \sum \ln(x_i) = \frac{1}{6} * 2266 + \frac{33,59}{6} * 11,14 = 440,04, \quad (10)$$

где x – длина посылки;
 y – расстояние передачи данных;
 n – количество проведенных опытов.

Тогда уравнение (9) логарифмической регрессии, описывающее зависимость длины посылки от расстояния передачи данных, будет выглядеть следующим образом:

$$\hat{y} = 440,04 - 33,59 * \ln(x).$$

Аппроксимируем график, изображенный на рисунке 1 (а) с помощью экспоненциальной регрессии.

Уравнение экспоненциальной регрессии показано на формуле 11.

$$\hat{y} = e^{a+b*x}, \quad (11)$$

где x – длина посылки;
 y – расстояние передачи данных.

Найдем коэффициент «b» по формуле 12.

$$b = \frac{n * \sum(x_i * \ln(y_i)) - \sum \ln(y_i) * \sum x_i}{n * \sum x_i^2 - (\sum x_i)^2} = \frac{6 * 311,65 - 35,58 * 53}{6 * 639 - 2809} = -0,02, \quad (12)$$

где x – длина посылки;
 y – расстояние передачи данных;
 n – количество проведенных опытов.

Найдем коэффициент «a» по формуле 13.

$$a = \frac{1}{n} * \sum \ln(y_i) - \frac{b}{n} * \sum x_i = \frac{1}{6} * 35,58 + \frac{0,02}{6} * 53 = 6,07, \quad (13)$$

где x – длина посылки;
 y – расстояние передачи данных;
 n – количество проведенных опытов.

Тогда уравнение (11) экспоненциальной регрессии, описывающее зависимость длины посылки от расстояния передачи данных, будет выглядеть следующим образом:

$$\hat{y} = e^{6,07-0,02*x}.$$

В таблице 3 приведены средние ошибки аппроксимаций в процентном соотношении. На рисунке 2 приведены график исходной функции и функциональных зависимостей, полученных различными методами аппроксимации.

Таблица 3. Средняя ошибка различных методов аппроксимации

Метода аппроксимации	Средняя ошибка, %
Линейная регрессия	3,33
Квадратичная регрессия	2,37
Кубическая регрессия	2,5
Степенная регрессия	2,77
Показательная регрессия	3,15
Логарифмическая регрессия	2,71
Экспоненциальная регрессия	3,15

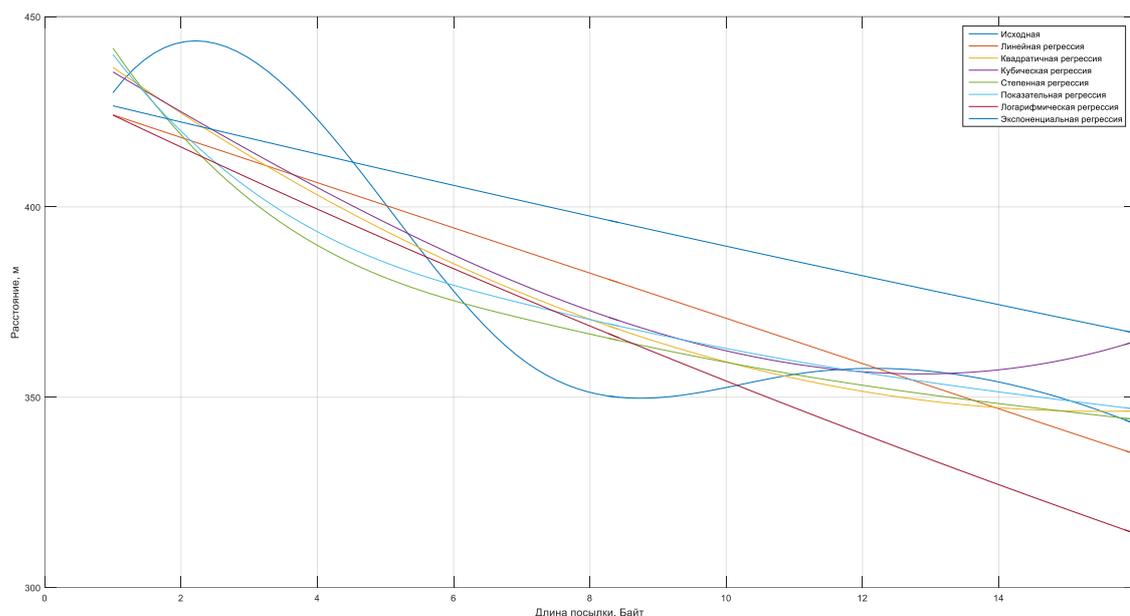


Рис. 2. График табличных экспериментальных значений и графики аппроксимированных функций

Аппроксимация зависимости времени передачи данных от длины посылки. Аппроксимируем график, изображенный на рисунке 1 (б) с помощью линейной регрессии. Уравнение линейной регрессии показано на формуле 1.

Найдем коэффициент «а» по формуле (2):

$$a = \frac{\sum x_i * \sum y_i - n * \sum x_i y_i}{(\sum x_i)^2 - n * \sum x_i^2} = \frac{53 * 427 - 6 * 4253}{53^2 - 6 * 639} = 2,82 .$$

Найдем коэффициент «b» по формуле (3):

$$b = \frac{\sum x_i * \sum x_i y_i - \sum x_i^2 * \sum y_i}{(\sum x_i)^2 - n * \sum x_i^2} = \frac{53 * 4253 - 639 * 427}{53^2 - 6 * 639} = 46,29 .$$

Тогда уравнение (1) линейной регрессии, описывающее зависимость длины посылки от времени передачи данных, будет выглядеть следующим образом:

$$\hat{y} = 2,82 * x + 46,29 .$$

Аппроксимируем график, изображенный на рисунке 1 (б) с помощью квадратичной регрессии. Уравнение кубической регрессии показано на формуле (5).

Найдем коэффициенты «а», «b» и «с» с помощью системы уравнений:

$$\begin{cases} a * \sum x_i^2 + b * \sum x_i + n * c = \sum y_i \\ a * \sum x_i^3 + b * \sum x_i^2 + \sum x_i = \sum x_i y_i \\ a * \sum x_i^4 + b * \sum x_i^3 + \sum x_i^2 = \sum x_i^2 y_i \end{cases} = \begin{cases} a * 639 + b * 53 + 6 * c = 427 \\ a * 8579 + b * 639 + c * 53 = 4253 \\ a * 121251 + b * 8579 + c * 639 = 54109 \end{cases} = \begin{cases} a = 0,13 \\ b = 0,55 \\ c = 52,26 \end{cases}.$$

Тогда уравнение (5) квадратичной регрессии, описывающее зависимость длины посылки от времени передачи данных, будет выглядеть следующим образом:

$$\hat{y} = 0,13 * x^2 + 0,55 * x + 52,26.$$

Аппроксимируем график, изображенный на рисунке 1 (б) с помощью кубической регрессии. Уравнение кубической регрессии показано на формуле (6). Найдем коэффициенты a, b, c и d с помощью системы уравнений:

$$\begin{cases} a * \sum x_i^3 + b * \sum x_i^2 + c * \sum x_i + n * d = \sum y_i \\ a * \sum x_i^4 + b * \sum x_i^3 + \sum x_i^2 + \sum x_i = \sum x_i y_i \\ a * \sum x_i^5 + b * \sum x_i^4 + \sum x_i^3 + \sum x_i^2 = \sum x_i^2 y_i \\ a * \sum x_i^6 + b * \sum x_i^5 + \sum x_i^4 + \sum x_i^3 = \sum x_i^3 y_i \end{cases} = \begin{cases} a * 8579 + b * 639 + c * 53 + 6 * d = 427 \\ a * 121251 + b * 8579 + c * 639 + d * 53 = 4253 \\ a * 1765283 + b * 121251 + c * 8579 + d * 639 = 54109 \\ a * 26200059 + b * 1765283 + c * 121251 + d * 8579 = 747647 \end{cases} = \begin{cases} a = -0,01 \\ b = 0,28 \\ c = -0,44 \\ d = 53,57 \end{cases}.$$

Тогда уравнение (6) кубической регрессии, описывающее зависимость длины посылки от времени передачи данных, будет выглядеть следующим образом:

$$\hat{y} = -0,01 * x^3 + 0,28 * x^2 - 0,44 * x + 53,57.$$

Аппроксимируем график, изображенный на рисунке 1 (б) с помощью степенной регрессии. Уравнение степенной регрессии показано на формуле (8).

Найдем коэффициент «b»:

$$b = \frac{n * \sum (\ln(x_i) * \ln(y_i)) - \sum \ln(x_i) * \sum \ln(y_i)}{n * \sum \ln^2(x_i) - (\sum \ln(x_i))^2} = \frac{6 * 48,33 - 11,14 * 25,45}{6 * 124,1 - 124,1} = 0,2.$$

Найдем коэффициент «a»:

$$a = \exp\left(\frac{1}{n} * \sum \ln(y_i) - \frac{b}{n} * \sum \ln(x_i)\right) = \exp\left(\frac{1}{6} * 25,45 - \frac{0,2}{6} * 11,14\right) = 48,28.$$

Тогда уравнение (7) степенной регрессии, описывающее зависимость длины посылки от времени передачи данных, будет выглядеть следующим образом:

$$\hat{y} = 48,28 * x^{0,2}.$$

Аппроксимируем график, изображенный на рисунке 1 (б) с помощью показательной регрессии. Уравнение показательной регрессии показано на формуле (8).

Найдем коэффициент «b»:

$$b = \exp\left(\frac{n * \sum (x_i * \ln(y_i)) - \sum x_i * \sum \ln(y_i)}{n * \sum x_i^2 - (\sum x_i)^2}\right) = \exp\left(\frac{6 * 231,61 - 53 * 25,45}{6 * 639 - 2809}\right) = 1,04.$$

Найдем коэффициент «a»:

$$a = \exp\left(\frac{1}{n} * \sum \ln(y_i) - \frac{\ln(b)}{n} * \sum x_i\right) = \exp\left(\frac{1}{6} * 25,45 - \frac{0,04}{6} * 53\right) = 48,94.$$

Тогда уравнение (8) показательной регрессии, описывающее зависимость длины посылки от времени передачи данных, будет выглядеть следующим образом:

$$\hat{y} = 48,94 * 1,04^x.$$

Аппроксимируем график, изображенный на рисунке 1 (б) с помощью логарифмической регрессии. Уравнение логарифмической регрессии показано на формуле (9).

Найдем коэффициент «b»:

$$b = \frac{n * \sum(y_i * \ln(x_i)) - \sum \ln(x_i) * \sum y_i}{n * \sum \ln^2 x_i - (\sum \ln(x_i))^2} = \frac{6 * 866,96 - 11,14 * 427}{6 * 124,1 - 124,1} = 13,66.$$

Найдем коэффициент «a» по формуле (10).

$$a = \frac{1}{n} * \sum y_i - \frac{b}{n} * \sum \ln(x_i) = \frac{1}{6} * 427 - \frac{13,66}{6} * 11,14 = 45,8.$$

Тогда уравнение (9) логарифмической регрессии, описывающее зависимость длины посылки от времени передачи данных, будет выглядеть следующим образом:

$$\hat{y} = 45,8 + 13,66 * \ln(x).$$

Аппроксимируем график, изображенный на рисунке 1 (б) с помощью экспоненциальной регрессии. Уравнение экспоненциальной регрессии показано на формуле (11). Найдем коэффициент «b» по формуле (12).

$$b = \frac{n * \sum(x_i * \ln(y_i)) - \sum \ln(y_i) * \sum x_i}{n * \sum x_i^2 - (\sum x_i)^2} = \frac{6 * 231,61 - 25,45 * 53}{6 * 639 - 2809} = 0,04.$$

Найдем коэффициент «a» по формуле 13.

$$a = \frac{1}{n} * \sum \ln(y_i) - \frac{b}{n} * \sum x_i = \frac{1}{6} * 25,45 - \frac{0,04}{6} * 53 = 3,89.$$

Тогда уравнение (11) экспоненциальной регрессии, описывающее зависимость длины посылки от времени передачи данных, будет выглядеть следующим образом:

$$\hat{y} = e^{3,89+0,04*x}.$$

В таблице 4 приведены средние ошибки аппроксимаций в процентном соотношении. На рисунке 3 приведены график исходной функции и функциональных зависимостей, полученных различными методами аппроксимации.

Таблица 4. Средняя ошибка различных методов аппроксимации

Метода аппроксимации	Средняя ошибка, %
Линейная регрессия	4,42
Квадратичная регрессия	1,84
Кубическая регрессия	1,65
Степенная регрессия	9,78
Показательная регрессия	2,86
Логарифмическая регрессия	11,42
Экспоненциальная регрессия	2,86

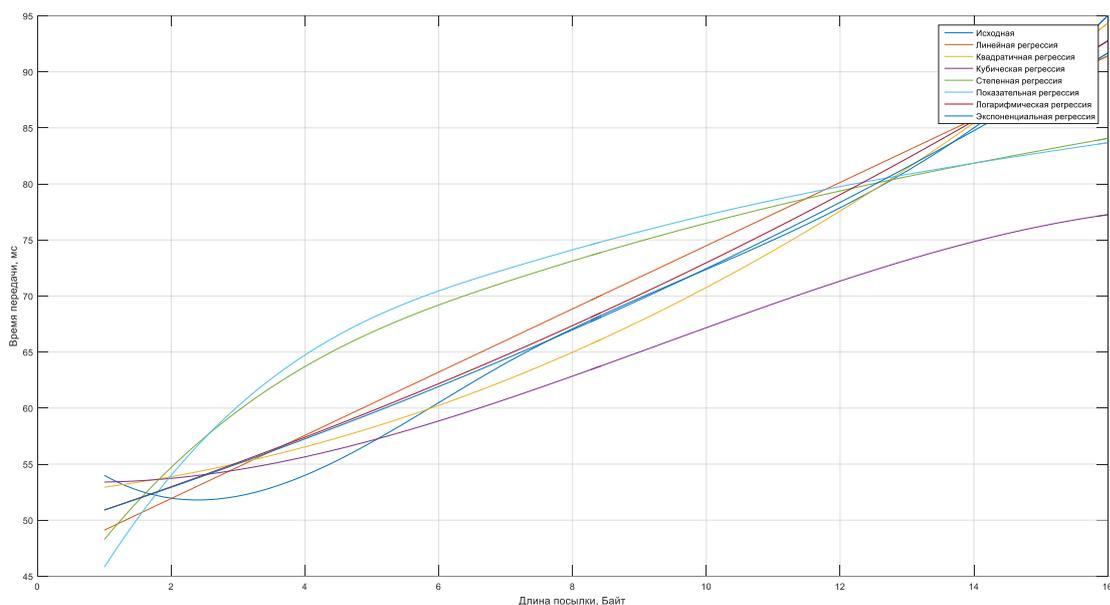


Рис. 3. График табличных экспериментальных значений и графики аппроксимированных функций

Аппроксимация зависимости скорости передачи данных от длины посылки. Аппроксимируем график, изображенный на рисунке 1 (в) с помощью линейной регрессии. Уравнение линейной регрессии показано на формуле (1).

Найдем коэффициент «а» по формуле (2):

$$a = \frac{\sum x_i * \sum y_i - n * \sum x_i y_i}{(\sum x_i)^2 - n * \sum x_i^2} = \frac{53 * 88,29 - 6 * 681,76}{2809 - 6 * 639} = -0,57.$$

Найдем коэффициент «b» по формуле (3):

$$b = \frac{\sum x_i * \sum x_i y_i - \sum x_i^2 * \sum y_i}{(\sum x_i)^2 - n * \sum x_i^2} = \frac{53 * 681,76 - 639 * 88,29}{2809 - 6 * 639} = 19,79.$$

Тогда уравнение (1) линейной регрессии, описывающее зависимость длины посылки от скорости передачи данных, будет выглядеть следующим образом:

$$\hat{y} = -0,57 * x + 19,79.$$

Аппроксимируем график, изображенный на рисунке 1 (в) с помощью квадратичной регрессии. Уравнение квадратичной регрессии показано на формуле (5).

Найдем коэффициенты «а», «b» и «с» с помощью системы уравнений:

$$\begin{cases} a * \sum x_i^2 + b * \sum x_i + n * c = \sum y_i \\ a * \sum x_i^3 + b * \sum x_i^2 + \sum x_i = \sum x_i y_i \\ a * \sum x_i^4 + b * \sum x_i^3 + \sum x_i^2 = \sum x_i^2 y_i \end{cases} = \begin{cases} a * 639 + b * 53 + 6 * c = 88,29 \\ a * 8579 + b * 639 + c * 53 = 681,76 \\ a * 121251 + b * 8579 + c * 639 = 7694,4 \end{cases} = \begin{cases} a = -0,01 \\ b = -0,43 \\ c = 19,42 \end{cases}$$

Тогда уравнение (5) квадратичной регрессии, описывающее зависимость длины посылки от скорости передачи данных, будет выглядеть следующим образом:

$$\hat{y} = -0,01 * x^2 - 0,43 * x + 19,42.$$

Аппроксимируем график, изображенный на рисунке 1 (в) с помощью кубической регрессии. Уравнение кубической регрессии показано на формуле (6). Найдем коэффициенты «а», «b», «с» и «d» с помощью системы уравнений:

$$\begin{cases} a * \sum x_i^3 + b * \sum x_i^2 + c * \sum x_i + n * d = \sum y_i \\ a * \sum x_i^4 + b * \sum x_i^3 + \sum x_i^2 + \sum x_i = \sum x_i y_i \\ a * \sum x_i^5 + b * \sum x_i^4 + \sum x_i^3 + \sum x_i^2 = \sum x_i^2 y_i \\ a * \sum x_i^6 + b * \sum x_i^5 + \sum x_i^4 + \sum x_i^3 = \sum x_i^3 y_i \end{cases} = \begin{cases} a * 8579 + b * 639 + c * 53 + 6 * d = 88,29 \\ a * 121251 + b * 8579 + c * 639 + d * 53 = 681,76 \\ a * 1765283 + b * 121251 + c * 8579 + d * 639 = 7694,4 \\ a * 26200059 + b * 1765283 + c * 121251 + d * 8579 = 99708 \end{cases} = \begin{cases} a = 0,004 \\ b = -0,1 \\ c = 0,16 \\ d = 18,63 \end{cases}$$

Тогда уравнение (6) кубической регрессии, описывающее зависимость длины посылки от скорости передачи данных, будет выглядеть следующим образом:

$$\hat{y} = 0,004 * x^3 - 0,1 * x^2 + 0,16 * x + 18,63.$$

Аппроксимируем график, изображенный на рисунке 1 (в) с помощью степенной регрессии. Уравнение степенной регрессии показано на формуле (7).

Найдем коэффициент «b»:

$$b = \frac{n * \sum (\ln(x_i) * \ln(y_i)) - \sum \ln(x_i) * \sum \ln(y_i)}{n * \sum \ln^2(x_i) - (\sum \ln(x_i))^2} = \frac{6 * 28,64 - 11,14 * 16}{6 * 130,87 - 130,86} = -0,2.$$

Найдем коэффициент «а»:

$$a = \exp\left(\frac{1}{n} * \sum \ln(y_i) - \frac{b}{n} * \sum \ln(x_i)\right) = \exp\left(\frac{1}{6} * 16 + \frac{0,2}{6} * 11,14\right) = 20,43.$$

Тогда уравнение (7) степенной регрессии, описывающее зависимость длины посылки от времени передачи данных, будет выглядеть следующим образом:

$$\hat{y} = 20,43 * x^{-0,2}.$$

Аппроксимируем график, изображенный на рисунке 1 (в) с помощью показательной регрессии. Уравнение показательной регрессии показано на формуле (8). Найдем коэффициент «b»:

$$b = \exp\left(\frac{n * \sum (x_i * \ln(y_i)) - \sum x_i * \sum \ln(y_i)}{n * \sum x_i^2 - (\sum x_i)^2}\right) = \exp\left(\frac{6 * 134,5 - 53 * 16}{6 * 639 - 2809}\right) = 0,96.$$

Найдем коэффициент «а»:

$$a = \exp\left(\frac{1}{n} * \sum \ln(y_i) - \frac{\ln(b)}{n} * \sum x_i\right) = \exp\left(\frac{1}{6} * 16 + \frac{0,04}{6} * 53\right) = 20,43.$$

Тогда уравнение (8) показательной регрессии, описывающее зависимость длины посылки от скорости передачи данных, будет выглядеть следующим образом:

$$\hat{y} = 20,43 * 0,96^x.$$

Аппроксимируем график, изображенный на рисунке 1 (в) с помощью логарифмической регрессии. Уравнение логарифмической регрессии показано на формуле (9). Найдем коэффициент «b»:

$$b = \frac{n * \sum(y_i * \ln(x_i)) - \sum \ln(x_i) * \sum y_i}{n * \sum \ln^2(x_i) - (\sum \ln(x_i))^2} = \frac{6 * 148,28 - 11,14 * 88,29}{6 * 124,1 - 124,1} = -2,89.$$

Найдем коэффициент «а» по формуле (10).

$$a = \frac{1}{n} * \sum y_i - \frac{b}{n} * \sum \ln(x_i) = \frac{1}{6} * 88,29 + \frac{2,89}{6} * 11,14 = 20,08.$$

Тогда уравнение (9) логарифмической регрессии, описывающее зависимость длины посылки от времени передачи данных, будет выглядеть следующим образом:

$$\hat{y} = 20,08 - 2,89 * \ln(x).$$

Аппроксимируем график, изображенный на рисунке 1 (в) с помощью экспоненциальной регрессии. Уравнение экспоненциальной регрессии показано на формуле (11).

Найдем коэффициент «b» по формуле (12):

$$b = \frac{n * \sum(x_i * \ln(y_i)) - \sum \ln(y_i) * \sum x_i}{n * \sum x_i^2 - (\sum x_i)^2} = \frac{6 * 134,5 - 16 * 53}{6 * 639 - 2809} = -0,04.$$

Найдем коэффициент «а» по формуле (13):

$$a = \frac{1}{n} * \sum \ln(y_i) - \frac{b}{n} * \sum x_i = \frac{1}{6} * 16 + \frac{0,04}{6} * 53 = 3,02.$$

Тогда уравнение (11) экспоненциальной регрессии, описывающее зависимость длины посылки от скорости передачи данных, будет выглядеть следующим образом:

$$\hat{y} = e^{3,02 - 0,04 * x}.$$

На рисунке 4 приведены график исходной функции и функциональных зависимостей, полученных различными методами аппроксимации. В таблице 5 приведены средние ошибки аппроксимаций в процентном соотношении.

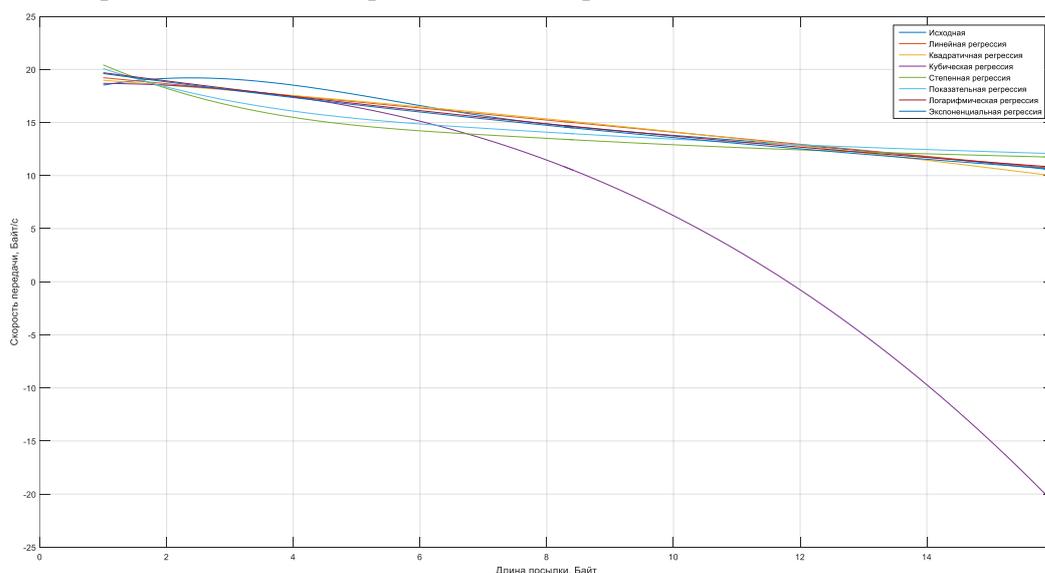


Рис. 4. График табличных экспериментальных значений и графики аппроксимированных функций

Таблица 5. Средняя ошибка различных методов аппроксимации

Метода аппроксимации	Средняя ошибка, %
Линейная регрессия	2,01
Квадратичная регрессия	2,33
Кубическая регрессия	2,04
Степенная регрессия	9,7
Показательная регрессия	2,86
Логарифмическая регрессия	8,5
Экспоненциальная регрессия	2,86

Выводы. Наиболее точным методом аппроксимации зависимости длины посылки от расстояния передачи данных является квадратичная регрессия. Не рекомендуется использовать линейную регрессию для аппроксимации зависимости длины посылки от расстояния передачи данных, так как этот метод аппроксимации имеет значительно высокую среднюю смещенную ошибку. Наиболее точным методом аппроксимации зависимости длины посылки от времени передачи данных является кубическая регрессия. Не рекомендуется использовать логарифмическую регрессию для аппроксимации зависимости длины посылки от времени передачи данных, так как этот метод аппроксимации имеет значительно высокую среднюю смещенную ошибку. Наиболее точным методом аппроксимации зависимости длины посылки от скорости передачи данных является линейная регрессия. Не рекомендуется использовать логарифмическую регрессию для аппроксимации зависимости длины посылки от скорости передачи данных, так как этот метод аппроксимации имеет значительно высокую среднюю смещенную ошибку.

В ходе проведения исследований выявлены зависимости между длиной посылки и расстоянием передачи данных, временем передачи посылки, а также скоростью передачи. Получены функциональные описания этих зависимостей различными методами аппроксимации экспериментальных данных, вычислены ошибки регрессии в процентном соотношении. Выделены наиболее точные методы аппроксимации данных, полученных в результате эксперимента.

СПИСОК ИСТОЧНИКОВ

1. Wireless, sensing & timing: Datasheet SX1276/77/78/79 / Semtech Corporation. – Camarillo: Semtech, 2015. – 131 p.
2. Naumov, M. A. Development of a control system for machines and mechanisms using lora wireless communication technology / M. A. Naumov, A. Yu. Karpovsky // Сборник материалов научно-практической конференции для преподавателей гуманитарных дисциплин Young scientists' researches and achievements in science. – Донецк: ДонНТУ, 2024. – С. 281-289.
3. Карповский, А. Ю. Разработка и исследование стенда для проверки беспроводной радиосвязи LoRa / А. Ю. Карповский, М. А. Наумов // Материалы XIX Всероссийской научно-практической конференции с международным участием Инженер настоящего и будущего: практика и перспективы развития партнерства в высшем техническом образовании. – Таганрог: ЮФУ, 2024. – Т. 2. – С. 330-336.
4. Наумов, М. А. Разработка системы управления машина и механизмами с применением беспроводной технологии связи LoRa / М. А. Наумов, А. Ю. Карповский, В. В. Казаков // Сборник научных трудов XXIV Международной научно-технической конференции Автоматизация Технологических процессов. Поиск молодых. – Донецк: ДонНТУ, 2024. – С. 123-126.
5. Малышева, Т. А. Численные методы и компьютерное моделирование. Лабораторный практикум по аппроксимации функций: учебно-методическое пособие / Т.А. Малышева. – СПб.: Университет ИТМО, 2016. – С. 33.

Поступила в редакцию 04.10.2024 г., рекомендована к печати 15.11.2024 г.

STUDYING THE EFFECT OF THE LENGTH OF A PARCEL ON THE DISTANCE, TIME AND SPEED OF DATA TRANSMISSION VIA LORA WIRELESS COMMUNICATION

Naumov M.A., Karpovskii A.IU., Kurnosov V.G.

This article analyzes the effect of the length of the parcel on the distance, time and data transfer rate over the LoRa wireless connection. The article examines the effect of packet length, an adjustable parameter in the transmission of LoRa data, on key performance indicators. It is shown that the length of the transmitted packet affects the distance traveled by the signal, the time spent on transmission, and the overall data transfer rate. Data on several experiments conducted under different conditions are presented. A comprehensive analysis of the experimental data is presented, which draws attention to the clear patterns and correlations observed during the experiment. The experimental data were approximated and corresponding graphs were constructed.

Keywords: wireless communication, parcel length, transmission distance, transmission time, transmission rate, LoRa, approximation, error.

Наумов Максим Александрович

младший научный сотрудник научно-исследовательского отдела автоматизации горных машин ГУ «Автоматгормаш им. В.А. Антипова», аспирант кафедры автоматизации и телекоммуникаций ФГБОУ ВО «Донецкий национальный технический университет», Российская Федерация, ДНР, г. Донецк.
E-mail: maksim-naumov-99@bk.ru

Naumov Maksim Aleksandrovich

Junior researcher of the Scientific Research Department of Automation of Mining machines of the «Automatgormash named after V.A. Antipov», postgraduate student of the Department of Automation and Telecommunications of the Donetsk National Technical University, Russian Federation, DPR, Donetsk.

Карповский Артур Юрьевич

заведующий научно-исследовательским отделом автоматизации горных машин ГУ «Автоматгормаш им. В.А. Антипова», Российская Федерация, ДНР, г. Донецк.
E-mail: oagm308@mail.ru

Karpovskii Artur Iurevich

Head of the Scientific Research Department of Automation of Mining Machines of the «Automatgormash named after V.A. Antipov», Russian Federation, DPR, Donetsk.

Курносков Вячеслав Григорьевич

доктор технических наук, профессор, первый заместитель директора по научной работе ГУ «Автоматгормаш им. В.А. Антипова», Российская Федерация, ДНР, г. Донецк.
E-mail: avtomatgormash@mail.ru

Kurnosov Viacheslav Grigorevich

Doctor of Technical Sciences, Professor, First Deputy Director for Scientific Work of the «Automatgormash named after V.A. Antipov», Russian Federation, DPR, Donetsk.

АНТЕННЫ НА ДИЭЛЕКТРИЧЕСКИХ РЕЗОНАТОРАХ© 2024 *Алексейчик Л.В., Курушин А.А.*

Предложены и промоделированы конструкции микроволновых антенн на диэлектрических резонаторах (ДР), предназначенных для применения в технике мобильной и спутниковой связи сантиметрового и миллиметрового диапазона длин волн. В результате численного моделирования получены основные технические параметры антенн на диэлектрических резонаторах, а также картины распределения векторов результирующего электромагнитного поля и частотные характеристики. Приведены диаграммы направленности ДРа, демонстрирующая свойства ДР как антенных элементов СВЧ диапазона. На этой основе выполнена волноводная антенная решетка, а также решетка на взаимно связанных диэлектрических резонаторах, установленных на подложке микрополосковой линии, симметрично по обе стороны полоскового проводника. Проведено также моделирование полуцилиндрического диэлектрического резонатора электромагнитным полем щели, выполненной в широкой стенке прямоугольного металлического волновода U -диапазона длин волн. Приведены результаты численного моделирования технических характеристик антенн на ДР.

Ключевые слова: микроволновые антенны ДРа, диэлектрические СВЧ резонаторы, численное моделирование, CSTSUITE.

Введение. Возможности применения диэлектрических резонаторов (ДР) в качестве антенных элементов рассматривались еще в первых работах по исследованию собственных типов колебаний диэлектрических тел с высокой диэлектрической проницаемостью, обладающих внутренними и внешними электромагнитными полями, энергия которых способна излучаться в окружающее свободное пространство, в частности при исследовании спектров колебаний диэлектрического шара [1, 2].

Однако в конструкциях частотно избирательных СВЧ устройств на ДР требовались меры по устранению излучения с помощью экранов, которыми могли служить металлические стенки линий передачи или специальные экранирующие устройства, что позволило реализовать новое поколение функциональных устройств СВЧ различного назначения, имеющих малые веса и габариты по сравнению с известными традиционными волноводными конструкциями [3, 4].

В первой публикации по реализации реальных конструкций микроволновой антенны на ДР [5] аналитический расчет её параметров и диаграммы направленности был проведен с помощью метода перевала и функции Грина для открытой конструкции микрополосковой линии при возбуждении ДР полосковым проводником линии. Позже в работе [6] были получены рекомендации по выбору диэлектрических материалов для оптимизации эффективности излучения антенн на ДР.

Все более возрастающий интерес к разработкам антенн и антенных решеток на ДР проявился в работах [7-9]. В настоящее время продолжают интенсивные исследования антенн на ДР (в англоязычной терминологии как DRA) в широкой области частот СВЧ диапазона, включая миллиметровый и субмиллиметровый диапазон волн, вплоть до оптического диапазона [10, 11]. С развитием мобильной и космической связи антенны на ДР получили распространение в средствах связи.

Основная часть. Остается важным исследование физических процессов дифракции ДР при возбуждении их сторонними источниками, что открывает новые возможности для совершенствования антенных СВЧ элементов и решеток на основе ДР.

Так в конструкции, предложенной в [12], достигнуто расширение функциональности, уменьшение тепловых потерь и повышение уровня излучения (рис. 1). Микроволновая антенна содержит микрополосковую линию 1, состоящую из диэлектрической подложки 2, металлизированного основания 3, полоскового проводника 4, и два диэлектрических резонатора 5 и 6, выполненных в виде прямоугольных призм, установленных на диэлектрической подложке 2 симметрично относительно полоскового проводника 4 с возможностью перемещения диэлектрических резонаторов 5 и 6 относительно полоскового проводника 4 и последующего их крепления на диэлектрической подложке 2, так, что широкие грани диэлектрических резонаторов 5 и 6 ориентированы вдоль полоскового проводника 4 и перпендикулярно поверхности диэлектрической подложки 2. ДР изготовлены из высокопроницаемого диэлектрического материала с относительной диэлектрической проницаемостью равной $\epsilon = 44$ и тангенсом угла потерь $\text{tg}\delta = 10^{-4}$. Геометрические размеры ДР: прямоугольная призма $4 \times 4 \times 2$ мм. Собственная резонансная частота ДР находится в окрестности 13,6 ГГц.

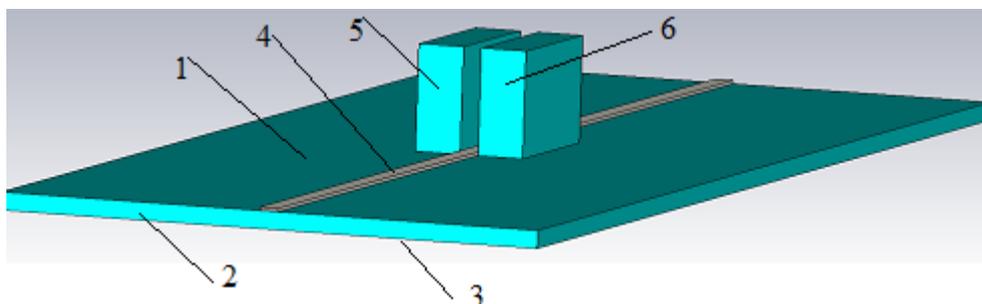


Рис. 1. Вид антенны на ДР

На входной порт МПЛ поступает сигнал электромагнитного поля распространяющейся ТЕМ-волны МПЛ, вызывающий возбуждение двух прямоугольных ДР на рабочем низшем типе колебаний $H_{11\delta}$, что приводит к накоплению энергии электромагнитного поля в них и излучению этой энергии из ДР в окружающее свободное пространство. Возбуждение ДР может быть осуществлено как в согласованном режиме МПЛ, так и в режиме «холостого хода» (ХХ), в этом случае ДР установлены в пучности магнитного поля МПЛ.

Электромагнитное поле пары ДР совместно с полем МПЛ наводит поверхностные токи на металлических частях конструкции, что дополнительно влияет на характеристики излучения антенны.

На рис. 2 и рис. 3 представлены картины векторного электрического поля E в плоскости широкой грани одного из ДР и векторного магнитного поля H в плоскости, параллельной плоскости диэлектрической подложки МПЛ на половинной высоте двух ДР, подтверждающих возбуждение ДР на резонансной частоте порядка 13,6 ГГц на низшем магнитном типе колебаний $H_{11\delta}$.

Из картин распределения электромагнитного поля следует, что за счет использования двух ДР со взаимной связью «согласного типа» включения резонаторов, достигается более широкое пространственное распределение результирующего магнитного момента (формируемого магнитным полем ДР) в сравнении с одиночным ДР, что приводит к усилению излучения электромагнитной энергии и к некоторому сужению лепестка диаграммы направленности (ДН) в плоскости, перпендикулярной полосковому проводнику МПЛ.

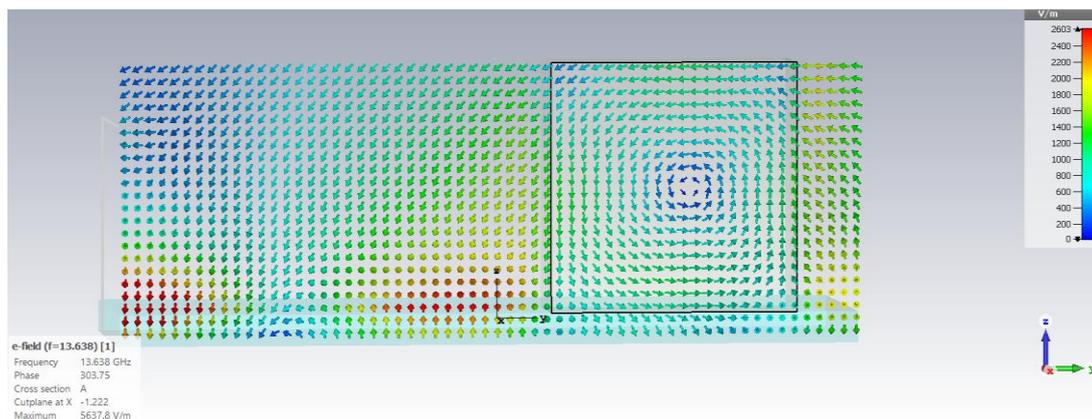


Рис. 2. Распределение векторов электрического поля в сечении ДР

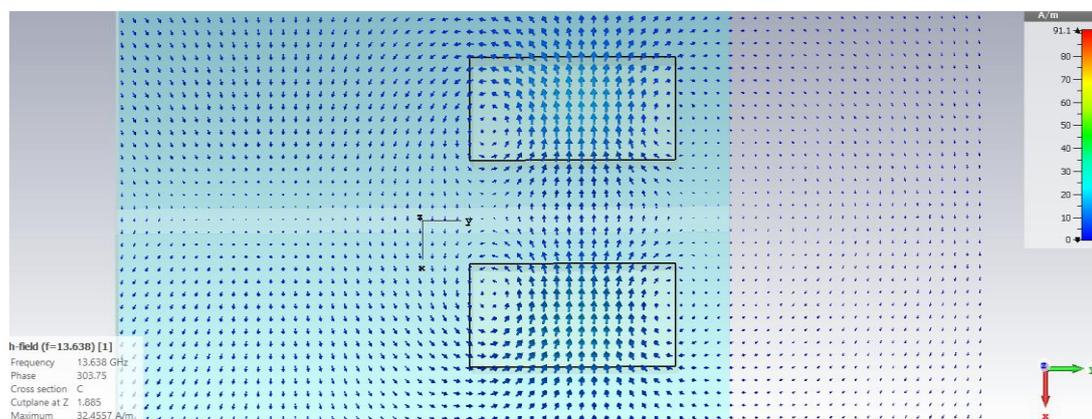


Рис. 3. Распределение векторов магнитного поля в сечении ДР

Благодаря размещению ДР на диэлектрической подложке вне полоскового проводника МПЛ достигается уменьшение тепловых потерь ДРа и следовательно повышение отношения потерь на излучение к тепловым потерям. Система двух ДР при этом образует звено двухконтурного фильтра, за счет чего достигается улучшение избирательности частотных характеристик ДРа.

На рис. 4 представлены диаграммы направленности (ДН) ДРа для режимов бегущей волны в МПЛ (режим согласованной нагрузки), аналогичная и для режима стоячей волны (режим холостого хода МПЛ).

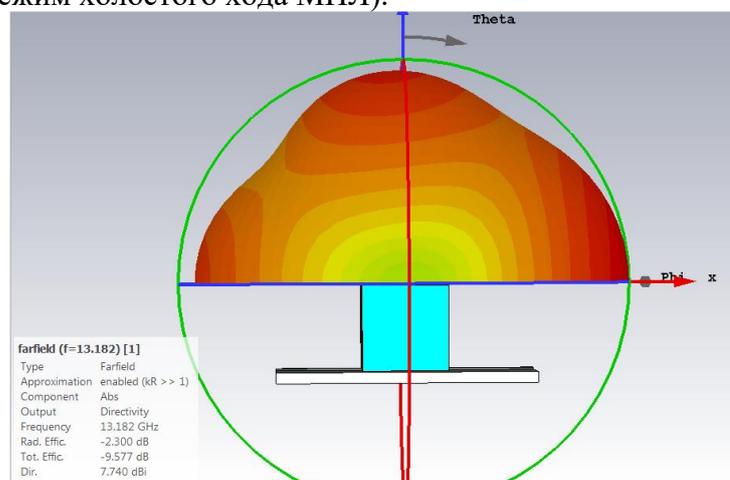


Рис. 4. ДН излучения ДРа (режим согласованной нагрузки МПЛ)

Так как ДР, помимо низшего рабочего типа колебаний, обладает множеством других высших типов колебаний, то для каждого из них ДРа формирует свою соответствующую ДН. В качестве примера дополнительно на рис. 5 показано распределение ДН для гибридного типа колебания двух ДР на высшем типе колебаний на частоте 18 ГГц. Однако, из данных рис. 5 следует, что эффективность излучения ДРа в этом случае значительно снижается по сравнению с ДРа, работающей на низшем магнитном типе колебаний ДР.

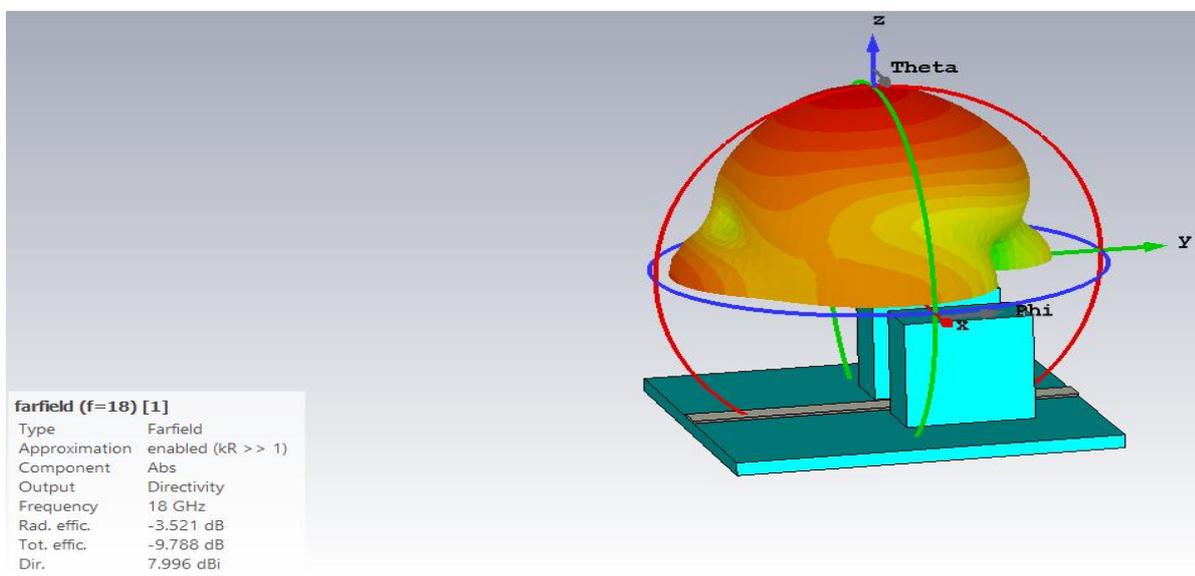


Рис. 5. ДН для гибридного типа колебания ДР на частоте 18 ГГц

Таким образом, конструкция ДРа с размещением двух прямоугольных ДР обеспечивает синфазное возбуждение ДР на H_{110} типе колебаний, необходимое для эффективного излучения энергии электромагнитного поля через апертуру раскрыва экрана МПЛ, при этом ДРа выполняет одновременно функции излучения антенного элемента и фильтрации двухзвенного фильтра.

Представленные результаты моделирования демонстрируют перспективность применения диэлектрических резонаторов в антенной СВЧ технике сантиметрового и миллиметрового диапазона длин волн, в том числе для создания малогабаритных антенных решеток (рис. 6).

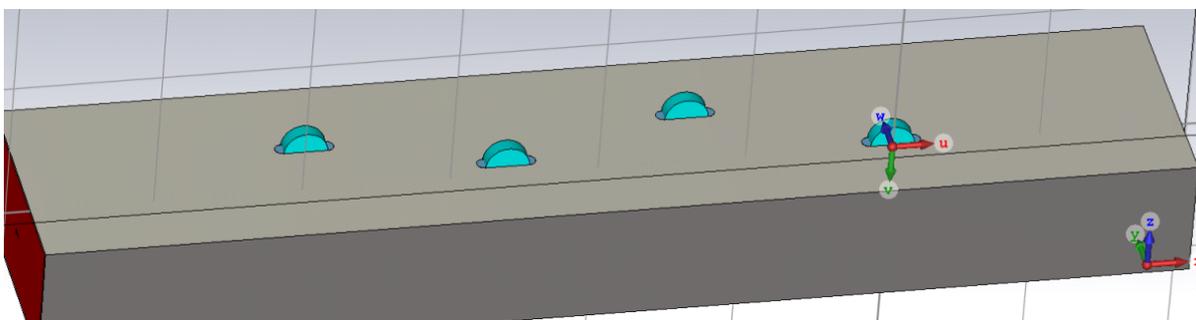


Рис. 6. Волноводная конструкция антенной решетки с цилиндрическими ДР

Вариант моделирования задачи возбуждения цилиндрического ДР, установленного в щели на металлической стенке прямоугольного волновода, полем падающей плоской электромагнитной волны, рассмотрен авторами в работе [13].

В результате численного моделирования с помощью программы CST Microwave Studio обнаружено преобразование низшего магнитного типа колебания ДР, размещенного в отверстии диафрагмы, в гибридный квази – $EH_{11\delta}$ тип колебания с поправкой на структуру электромагнитного поля, обусловленной влиянием отверстия связи. По АЧХ ДР и временным данным скорости затухания процесса возбуждения ДР были получены значения его резонансной частоты и нагруженной добротности, в том числе при изменении величины диэлектрической проницаемости материала ДР. Приведены диаграммы направленности в виде эффективных площадей рассеяния (RCS), демонстрирующие характеристики ДР как возможных антенных элементов СВЧ диапазона, подходящих для реализации простых конструкций антенных решеток (рис. 7).

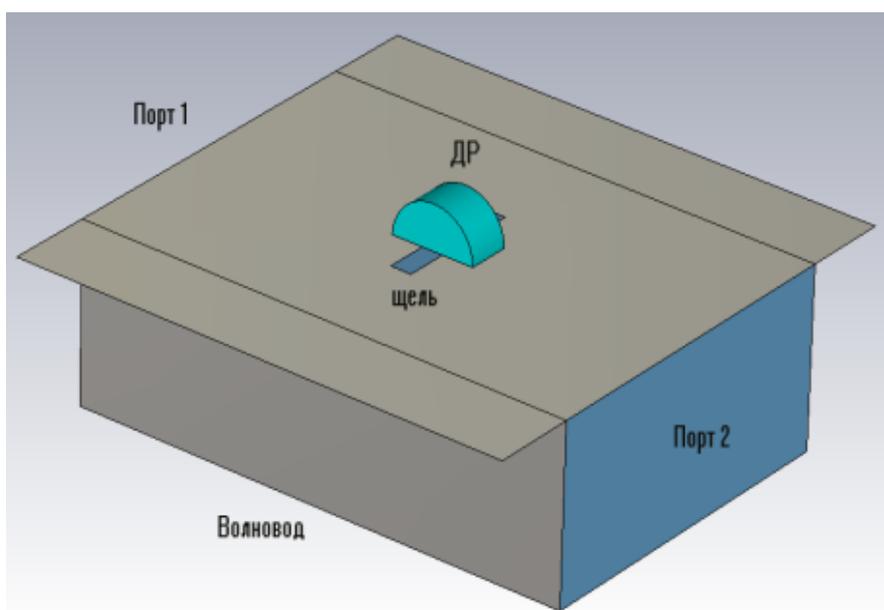


Рис. 7. Фрагмент антенной решетки на ДР

Дополнительным достоинством такой конструкции является возможность надежного крепления полуцилиндрического ДР к стенке волновода и меньший расход диэлектрического материала при реализации антенных решеток с большим числом ДР. В волноводе обеспечивается распространение рабочего H_{10} типа волны. В результате возбуждения ДР часть энергии его электромагнитного поля эффективно излучается в свободное пространство. Электромагнитное поле щели возбуждают резонансные колебания ДР на типе колебаний, близким по структуре поля к низшему типу колебания целого дискового ДР $H_{01\delta}$, однако преобразующиеся в гибридный квази- $EH_{11\delta}$ тип колебаний, из-за влияния щели и металлической поверхности волновода, как дополнительных неоднородностей, аналогично случаю размещения целого дискового ДР в отверстии металлической стенки, рассмотренного ранее в [12]. ДР изготовлен из высокопроницаемого диэлектрического материала с относительной диэлектрической проницаемостью $\epsilon = 44$ и тангенсом угла потерь $\text{tg}\delta = 10^{-4}$. Геометрические размеры ДР: диаметр 5 мм, толщина 2,5 мм. Собственная резонансная частота ДР находится в окрестности 15 ГГц и уточняется при моделировании.

Анализ полученных дифракционных картин электромагнитного поля рассеяния ДР позволил выявить характерные особенности, которые необходимо учитывать при разработке современных антенн. На рис. 8 приведена картина распределения векторов

электрического поля в области ДР и волновода, подтверждающие тип волны волновода и тип колебаний ДР, как указывалось ранее. При этом, судя по картине электрического поля, в ДР преобладает азимутальная составляющая векторов электрического поля, что подтверждает возбуждение типа колебаний ДР, близкого к гибричному квази- $EH_{11\delta}$ типу колебаний, где наблюдается минимум магнитного поля вблизи щели. При этом E -составляющая поля имеет выраженный минимум внутри полуцилиндрического ДР в области, близкой к 0,4-0,5 его радиуса, в то время как H -составляющая, напротив, имеет внутри этой области максимум. Учет этой особенности позволил уточнить рабочий тип колебаний как гибридный квази- $EH_{11\delta}$ с дополнительной вариацией электрического поля по радиусу полуцилиндра.

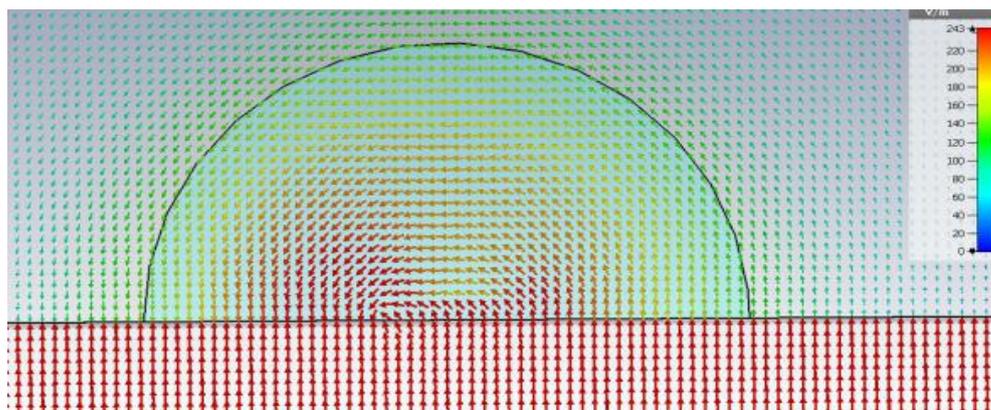


Рис. 8. Распределение векторов электрического поля ДР, щели и волновода в плоскости YOZ

При малых перемещениях ДР вдоль щели (или при регулировке размеров щели) достигается настройка антенны на оптимальные импедансные характеристики. При размещении ДР в волноводной конструкции линейной антенной решетки следует учитывать взаимные связи между ДР по распространяющейся волне для достижения требуемых фазовых соотношений между элементами решетки [14]. При численном моделировании значение резонансной частоты ДР определялось из графика АЧХ, полученного при введении пробника магнитного поля, установленного в центре полуцилиндрического ДР и ориентированного по направлению силовых линий магнитного поля, и ее величина составила 15,477 ГГц. При обработке данных АЧХ уточнялись значения резонансной частоты, полосы пропускания и вносимого ослабления. Полоса частот пропускания составила 360,0 МГц, нагруженная добротность ДР равна $Q_n = 15,477/360 = 43,0$. Добротность связи Q_c ДР со свободным пространством находится из соотношения:

$$1/Q_n = 1/Q_o + 1/Q_c,$$

где Q_o - собственная добротность ДР, обратно пропорциональная $\text{tg}\delta$ материала диэлектрика ДР. Для высокодобротных диэлектриков современных керамик тангенс угла потерь $\text{tg}\delta$ составляет не более 10^{-4} , при этом добротность связи на излучение практически равна нагруженной добротности ДР. Для снижения уровня добротности связи на излучение целесообразно использование диэлектрических материалов для изготовления ДР с более низкими значениями относительной диэлектрической проницаемости, однако обеспечивающих требуемое значение рабочей резонансной частоты на заданном типе колебаний ДР. Дополнительные возможности для эффективного излучения ДР связаны с вариацией формы ДР и включением различных неоднородностей, например, в виде воздушных зазоров [5, 15-17].

Выводы. Для случая цилиндрического ДР, установленного в отверстии металлической стенки и возбужденного гармоническим во времени электромагнитным полем падающей плоской электромагнитной волны были выявлены особенности распределения электромагнитного поля ДР, такие как эффекты "втягивания" пучности электрического поля падающей плоской волны в область объема ДР и соответствующее "втягивание" узла магнитного поля падающей плоской волны в окрестности ДР, связанных с наложением собственного поля ДР с полем стоячей волны в освещенном полупространстве от металлической стенки. Рассчитанные рабочие параметры антенны на ДР, включая АЧХ, значения резонансной частоты и добротности связи ДР со свободным пространством, а также 3D диаграммы направленности, согласовались с известными данными. В предложенной конструкции микрополосковой ДРа с парой симметрично размещенных относительно полоскового проводника ДР, благодаря размещению ДР на диэлектрической подложке вне полоскового проводника МПЛ достигается уменьшение тепловых потерь ДРа и следовательно повышение отношения потерь на излучение к тепловым потерям. Система двух ДР при этом образует звено двухконтурного фильтра, за счет чего достигается улучшение избирательности частотных характеристик ДРа.

В результате численного моделирования задачи возбуждения вынужденных колебаний ДР под действием электромагнитного поля щели в широкой стенке прямоугольного металлического волновода выявлены физические особенности явления дифракции и излучения ДР, получены картины распределения результирующего электрического поля ДР в ближней зоне и диаграмма направленности в дальней зоне. Впервые получены частотные характеристики амплитуд вынужденных колебаний полуцилиндрического ДР. Установлен низший тип колебаний как гибридный квази- $EH_{11\delta}$ тип колебания.

Приведены рекомендации по повышению эффективности излучения представленных конструкций микроволновых ДРа. Полученные в работе результаты представляют интерес при разработке антенн и антенных решеток на основе диэлектрических резонаторов, получивших широкое распространение в средствах космической и мобильной связи 5G поколения.

СПИСОК ИСТОЧНИКОВ

1. Richtmyer, R. D. Dielectric resonators / R. D. Richtmyer // Journal of Applied Physics. – 1939. – V. 10. – P.391-398.
2. Sager, O. On eigenmodes and forced resonance-modes of dielectric spheres / O. Sager, F. Tisi // Proceedings of the IEEE. – 1968. – V. 56, № 9. – P. 1593-1594.
3. Параметры и методы расчета диэлектрических резонаторов и генераторов и фильтров на их основе / Л. В. Алексейчик, И. И. Бродуленко, Н. Г. Гаврилюк [и др.]. – М.: Изд-во ЭТ, 1990. – 64 с.
4. Алексейчик, Л. В. Моделирование вносимых потерь диэлектрических резонаторов в конструкциях фильтров сверхвысоких частот / Л. В. Алексейчик // Известия Российской академии наук. Серия физическая. – 2018. – Т. 82, № 8. – С. 1012-1016.
5. Алексейчик, Л. В. Исследование диэлектрических СВЧ резонаторов в качестве антенных СВЧ элементов / Л. В. Алексейчик, В. М. Геворкян, Е. И. Калугин // Сб. науч. тр. МЭИ. – 1981. – № 528. – С. 31-40.
6. Long, S. A. The Resonant Cylindrical Dielectric Cavity Antenna / S. A. Long, M. W. McAllister, L. C. Shen // IEEE Transactions on Antennas and Propagation. – 1983. – V. 31, № 3. – P. 406-412.
7. Leung, K. W. State Key Laboratory of Millimeter Waves & Department of Electronic Engineering, City University of Hong Kong / K. W. Leung. – Presented to the IEEE LI Section Antenna and Propagation Society on Monday October 8th 2012. – 122 p.
8. Dielectric Resonator Antennas: Designs and Advances / D. Soren, R. Ghatak, R. K. Mishra, D. R. Poddar // Progress In Electromagnetics Research B. – 2014. –V. 60. – P. 195-213.
9. Maity, S. Closed form expressions to find radiation patterns of rectangular dielectric resonator antennas for various modes / S. Maity, B. Gupta // IEEE Transactions on Antennas and Propagation. – 2014. – V. 62, № 12. – P. 6524-6527.

10. Malheiros-Silveira, G. N. Dielectric resonator nanoantenna coupled to metallic coplanar waveguide / G. N. Malheiros-Silveira, H. E. Hernandez-Figueroa // IEEE Photonics Journal. – 2015. – V. 7, №. 1. – P. 1–7.
11. Keyrouz, L. Dielectric Resonator Antennas: Basic Concepts, Design Guidelines, and Recent Developments at Millimeter-Wave Frequencies / L. Keyrouz, D. Caratelli // International Journal of Antennas and Propagation. – V. 2016. – 20 p. –DOI 10.1155/2016/6075680.
12. Алексейчик, Л. В. Моделирование возбуждения диэлектрического резонатора полем плоской электромагнитной волны / Л. В. Алексейчик, А. А. Курушин // Журнал радиоэлектроники. – 2020. – № 11. – С. 6.
13. Алексейчик, Л. В. Моделирование частотных характеристик полосно-пропускающего фильтра мобильной радиорелейной связи на диэлектрических резонаторах СВЧ / Л. В. Алексейчик, М. П. Жохова, Г. А. Любимова // Известия Российской академии наук. Серия физическая. – 2020. – Т. 84, № 2. – С. 154-156.
14. Алексейчик, Л. В. Комплексное моделирование в программе CSTSUITE / Л. В. Алексейчик, А. А. Курушин. –М.: СОЛОН Пресс, 2022. – 276 с.
15. Алексейчик, Л. В. Моделирование диэлектрических резонаторов со взаимной связью по распространяющейся волне в волноводе / Л. В. Алексейчик // Известия Российской академии наук. Серия физическая. – 2022. – Т. 86, № 9. – С. 1283-1287.
16. Лу, Г. Моделирование и экспериментальные измерения многолучевой цифровой фазированной антенной решетки / Г. Лу, П. Н. Захаров, А. Ф. Королев // Известия Российской академии наук. Серия физическая. – 2019. – Т. 83, № 1. – С. 50-53.
17. Эффективная добротность миниатюрных диэлектрических резонаторов в экранированных линиях передачи / В. М. Геворкян, С. В. Вишняков, Ю. А. Казанцев, А. В. Шутов // Известия Российской академии наук. Серия физическая. – 2022. – Т. 86, № 9. – С. 1278-1282.

Поступила в редакцию 15.10.2024 г., рекомендована к печати 01.11.2024 г.

ANTENNAS ON DIELECTRIC RESONATOR

Alekseichik L.V., Kurushin A.A.

The designs of microwave antennas based on dielectric resonators (DR) designed for use in mobile and satellite communications technology in the centimeter and millimeter wavelength ranges are proposed and modeled. As a result of numerical simulation, the main technical parameters of antennas on dielectric resonators, as well as patterns of distribution of vectors of the resulting electromagnetic field and frequency characteristics are obtained. The radiation pattern of the DRa is shown, demonstrating the properties of the DR as antenna elements of the microwave range. On this basis, an antenna array is made on mutually connected dielectric resonators mounted on microstrip line substrate, symmetrically on both sides of the strip conductor. The simulation of a semi-cylindrical dielectric resonator by the electromagnetic field of a slit made in the wide wall of a rectangular metal waveguide of the U-wavelength range is carried out. The results of numerical modeling of the technical characteristics of the antenna and antenna array are presented.

Keywords: Microwave antennas DRa, dielectric microwave resonators, Antenna array, numerical modeling, CST SUITE.

Алексейчик Леонард Валентинович

доктор технических наук, профессор, профессор кафедры вычислительных машин, систем и сетей ФГБОУ ВО «Национальный исследовательский университет "Московский Энергетический Институт"»,

Российская Федерация, г. Москва.

E-mail: alexeychiklv@mpei.ru

Alekseichik Leonard Valentinovich

Doctor of Technical Sciences, Full Professor, Professor at Departments of Computing Machines, Systems and Networks of Moscow Power Engineering Institute,

Russian Federation, Moscow.

Курушин Александр Александрович

кандидат технических наук, старший научный сотрудник, доцент кафедры радиотехнических приборов и антенных систем ФГБОУ ВО «Национальный исследовательский университет "Московский Энергетический Институт"»,

Российская Федерация, г. Москва.

E-mail: kurushin@mail.ru

Kurushin Aleksandr Aleksandrovich

Candidate of Technical Sciences, Senior Researcher, Associate Professor at Department of Radio Technical Devices and Antenna Systems of Moscow Power Engineering Institute,

Russian Federation, Moscow.

СОВЕРШЕНСТВОВАНИЕ СТРУКТУРНОЙ МОДЕЛИ И ПАРАМЕТРОВ КОМПАКТНОГО ФУРЬЕ-СПЕКТРОМЕТРА ДЛЯ АНАЛИЗА ГАЗОВОГО СОСТАВА АТМОСФЕРЫ

© 2024 *Акапьев В.Л., Савотченко С.Е.*

Предложена модель усовершенствованной структурной модели и параметров компактного Фурье-спектрометра для анализа газового состава атмосферы. Схема спектрометра базируется на системе кодирования идентичных импульсов на основе перемещающегося клина (TWINS). Проанализированы зависимости разности хода полуволн в интерферометре от длины волны при варьировании толщины изменяемого TWINS блока. Произведен расчет и выполнен обоснованный подбор компонентов системы.

Ключевые слова: спектрометр, интерферометр, Фурье-спектрометр, фотоприёмник, спектральный анализ.

Введение. В настоящее время спектроскопия является одним из наиболее приоритетных направлений в научных исследованиях, поскольку сфера применения данного метода невероятно широка [1, 2]. Следует отметить, что существует сравнительно небольшое количество видов спектрометров, работающих в диапазоне длин волн, близком к оптическому [2-7]. Большинство из них основано на интерферометре Майкельсона («Bruker» IFS125HR, ИКФС2) [8], который обладает наилучшими характеристиками разности хода в области низких частот, когда минимальные перемещения подвижного зеркала меняют фазу не так кардинально, как при работе с излучением высокой частоты. На основании этого возникает необходимость внесения коррективов непосредственно в конструкцию интерферометра с целью перехода рабочего режима в диапазон более высоких частот.

Представленная в данной работе усовершенствованная модель компактного Фурье-спектрометра предназначена для проведения изучения газового состава атмосферы в диапазоне длин волн 800–1700 нм. В качестве основы выбрана принципиальная схема Фурье-спектрометра с интерферометром TWINS (Translating-Wedge-based Identical-pulses Ncoding System - система кодирования идентичных импульсов на основе перемещающегося клина) [9, 10]. Основной задачей является модернизация схемы спектрометра и оптимизация параметров его системы с помощью моделирования, на основе которого производится подбор ее элементной базы и проектирование.

При спектральном анализе газового состава атмосферы в диапазоне длин волн 800–1700 нм возникает необходимость детектирования такого ряда веществ как H₂O, CO₂, O₃, N₂O, CO, O₂, NO, NH₃, OH. Для детектирования спектров поглощения таких газов обычно используются спектрометры на интерферометре Майкельсона [11], а не на основе интерферометра TWINS. Здесь следует отметить, что, несмотря на сравнительную простоту конструкции интерферометра Майкельсона, по сравнению с ним преимущества интерферометра TWINS заключаются, в первую очередь, в существенно более высокой точности настройки разности хода. Кроме того, он может работать и в инфракрасной, и в ультрафиолетовой и в видимой областях спектра [12]. Несмотря на необходимость наличия в схеме спектрометра на интерферометре TWINS второго светоделительного куба, имеются возможности упрощения относительно сложной его конструкции путем совершенствования структурной схемы, результаты которого и представлены в данной работе.

Структурная схема. Интерферометр TWINS основан на введении разности хода между лучами, однако использует отличный от интерферометра Майкельсона метод контролирования задержки. Предлагаемая в данной работе усовершенствованная структурная схема Фурье-спектрометра с интерферометром TWINS приведена на рис. 1.

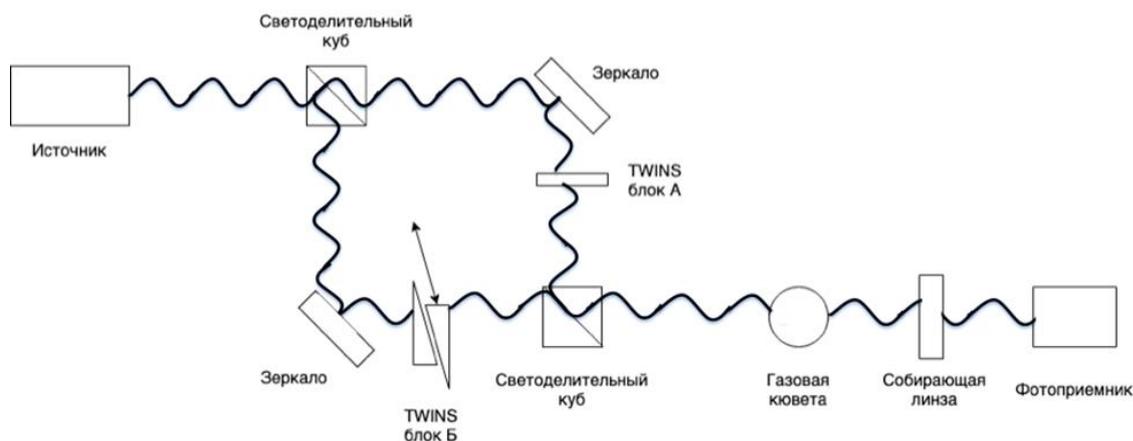


Рис. 1. Структурная схема Фурье-спектрометра с интерферометром TWINS при максимальной толщине блока Б

Широкополосный источник излучает волны определенного диапазона частот. После излучение приходит на светоделительный куб. В данной конфигурации устройство выполняет аналогичную функцию, как светоделитель в интерферометре Майкельсона. Разность хода уже отдельных лучей контролируется с помощью блока клиньев Б, который изготовлен из того же материала, что и блок А. Особенностью конструкции является то, что один из клиньев подвижный; направление его перемещения показано на рисунке 1. Сдвигая клин TWINS в блоке Б, можно изменять длину оптического пути в кристалле для одного из лучей, что, в свою очередь, позволяет регулировать «запаздывание» относительно второго луча. [10]. Далее лучи интерферируют на таком же светоделительном кубе.

Также, как и при работе интерферометра Майкельсона некоторые длины волн не будут «разрешены» ввиду взаимного ослабления при интерференции в разных фазах. Затем получившийся пучок проходит через изучаемый образец. И уже собирающая линза фокусирует луч на вход фотоприёмника, который регистрирует интерферограмму. Важно отметить, что оптические длины путей до и после зеркал в разделённом пучке должны быть одинаковыми для корректного определения разности хода.

Следует отметить, что в отличие от схемы, приведенной в [10], в предложенной в данной работе схеме (рис. 1), удалось избежать использования таких достаточно дорогостоящих элементов, как полуволновая пластина и поляризатор. Очевидно, что такое решение является более эффективным, так как не снижает функциональность спектрометра и его характеристики, но, в тоже время, снижает его себестоимость в целом.

Достоинство предложенной схемы прибора по сравнению со спектрометрами, основанными на интерферометре Майкельсона, заключается в более высокой точности настройки разности хода компонент луча за счёт малых углов при основании клина. Это связано с малым изменением оптического пути внутри кристалла, а, следовательно, и с малым изменением скорости излучения.

Особенности изменения спектральных параметров TWINS блоков. Важной характеристикой Фурье-спектрометра является вводимая и контролируемая разность хода разделённых лучей [13]:

$$\tau = (d_B - d_A) \left(\frac{1}{v_g} - \frac{1}{c} \right). \quad (1)$$

В (1) d_A и d_B – толщины блоков TWINS А и Б соответственно, v_g – групповая скорость лучей, c – скорость света в вакууме.

Тогда разность хода полуволн:

$$\Delta = \frac{\tau c}{2\lambda} = \frac{c(d_B - d_A)}{2\lambda} \left(\frac{1}{v_g} - \frac{1}{c} \right), \quad (2)$$

где λ – длина интерферирующей волны.

Групповая скорость лучей:

$$v_g = \frac{c}{n}, \quad (3)$$

где n – показатель преломления среды.

Поскольку разрабатываемый прибор предназначен для анализа газового состава атмосферы в диапазоне длин волн 800–1700 нм, то в качестве материала блоков TWINS было выбрано оптическое стекло N-BK7, которое достаточно хорошо пропускает излучение указанного диапазона: пропускание излучения в диапазоне 400–1700 нм составляет более 90% [14].

В системе TWINS толщина блока А $d_A = 2$ мм, блок Б имеет варьируемую клином толщину d_B в интервале от 2,1 мм до 5,4 мм.

Распределение групповой скорости луча (3) в N-BK7 в интервале 800–1700 нм показано на рис. 2, которое получено с использованием данных о зависимости показателя преломления от длины волны [15].

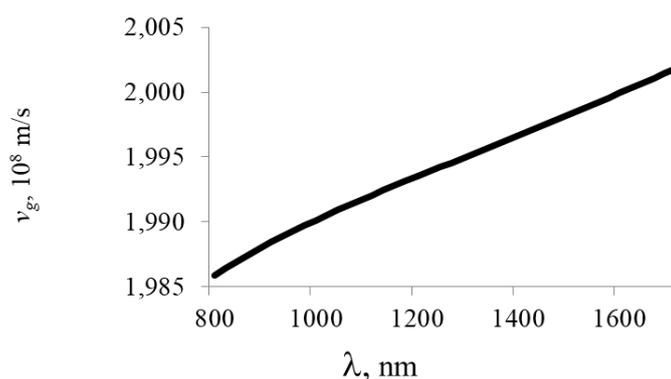


Рис. 2. Зависимость групповой скорости луча от длины волны

Видно, что групповая скорость монотонно возрастает с увеличением длины волны в рассматриваемом диапазоне от $1,985 \cdot 10^8$ до $2,002 \cdot 10^8$ м/с.

Изменение разности хода полуволн (2) в рассматриваемом диапазоне длин волн при базовой 0,036 мм, минимальной 0,021 мм и максимальной 0,054 мм толщинах приведена на рис. 3.

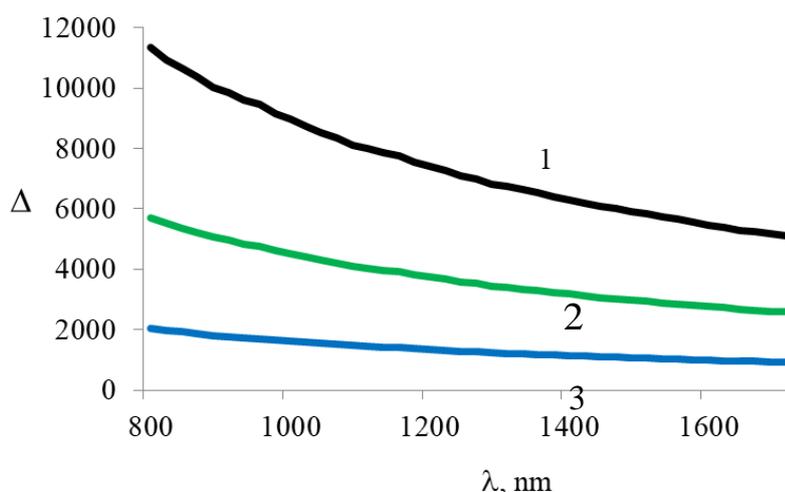


Рис. 3. Зависимость разности хода полуволн от длины волны при максимальной (1), базовой (2) и минимальной (3) толщинах блока Б

Видно, что разность хода полуволн монотонно убывает с ростом длины волны. Диапазон изменения разности хода полуволн расширяется при увеличении толщины блока Б. В частности, при минимальной толщине блока Б разность хода полуволн с ростом длины волны убывает от 2050 до 927 числа полуволн, а при максимальной – от 5610 до 2537.

Таким образом, в предложенной схеме спектрометра относительно небольшие клинья могут обеспечить разность хода от 2000 до 12000 полуволн. Поскольку обеспечиваемая разность хода является одной из важных характеристик прибора, та данный диапазон следует указать в спецификации изделия.

Основные компоненты системы. Для разработки модели проектируемой системы требуется точный подбор компонентов. К основным оптическим компонентам прибора относятся зеркала, линза, светоделительные кубы, многоходовая газовая кювета, фотоприёмник (см. рис. 1). В качестве оптических компонентов разрабатываемого прибора выбраны коммерчески доступные элементы, за исключением самостоятельно разработанных клиньев TWINS.

Постольку основным параметром для подбора линз является рабочая длина волны излучения, была выбрана собирающая линза LB1723 без покрытия производителя Thorlabs (диаметр = 25,4 мм; фокусное расстояние = 25 мм; рабочая апертура = 22,8 мм; диапазон рабочих длин волн = 320–2000 нм; пропускание излучения >96% [16]). С помощью такой линзы луч будет фокусироваться на приёмной площадке фотоприёмника.

В качестве зеркала, помогающего управлять ходом лучей в приборе, было выбрано зеркало PF20-03-M03 с золотым покрытием (диаметр = 25,4 мм; рабочая апертура = 22,8 мм; диапазон рабочих длин волн = 700–2000 нм; отражение излучения = 98%) того же производителя, от которого предлагается также взять светоделительный куб PBS513 (габаритные размеры, 25,4x25,4x25,4 мм; диапазон рабочих длин волн = 300–2300 нм; пропускание излучения >92% [16]).

В проекте был выбран фотоприёмник HSA-S-400M-IN производителя Femto с наилучшей чувствительностью в рабочем диапазоне длин волн (регистрируемый диапазон длин волн = 800–1700 нм; MinimumNEP = $6 \cdot 10^{-14}$, Вт/нм²; диаметр приёмной площадки = 2 мм, материал - InGaAs [17]).

В качестве газовой кюветы была выбрана многоходовая кювета КГ48 производства Инфраспек (габаритные размеры, 75x60x270 мм; длина кюветы = 0,2 м; длина оптического пути = 2,4 м; отражение зеркал >98% [18]).

Таким образом, при использовании схемы, представленной на рис. 1, с указанными компонентами, излучение претерпит 12 проходов (11 отражений), а длина оптического пути излучения в газовой среде увеличится до 2,4 метров.

Для определения подходящей температуры теплового источника был произведен расчёт функции Планка – светимости абсолютно чёрного тела. По закону распределения этой функции тела с более высокой температурой обладают излучением более высокой частоты (короткой длины волны). Это проявляется в смещении максимума функции, точки перегиба графика. Подходящей была принята температура, при которой максимум функции приходится примерно на середину диапазона 800–1700 нм. Именно при этом условии максимальная доля излучаемой энергии будет сконцентрирована в интересующем нас диапазоне длин волн. Если выбрать слишком высокую температуру, получается менее пологий график зависимости, однако сконцентрированная в площади под графиком мощность излучения неравномерно распределяется за пределами выбранного диапазона, в то время как в нужном диапазоне остается малая её часть. Если же выбрать слишком низкую температуру источника излучения, то в первую очередь можно столкнуться с проблемой недостаточной мощности. Подходящей была принята температура 2000 К. Поэтому была подобрана галогенная лампа QTH5 производителя Thorlabs, к преимуществам которой отнестись то, что излучение выходит из неё уже сколлимированным.

Согласованность технических параметров фотоприемника и с данной лампой основывается на следующем. Излучение претерпевает ослабление в многоходовой кювете и на остальной трассе прибора. В многоходовой кювете ослабления обусловлены: поглощением излучения газами и потерями на зеркалах. Коэффициент отражения зеркал кюветы 98%. Поглощение газами будем считать по максимальному ослаблению в выбранном интервале длин волн. Согласно эталону USAmode1 1976 summer, в диапазоне 800–1700 нм минимальным пропусканием излучения обладают водяные пары – около 0,979 на 10 мм в диапазоне длин волн 1382–1383 нм (округлено до 3 значащих цифры после запятой в сторону худшего пропускания) [19].

Минимальная мощность, регистрируемая выбранным приёмником рассчитывается по формуле $W_{\min} = \text{NEP} \cdot (\Delta f)^{1/2}$, где NEP – эквивалентная шумовая мощность выбранного приёмника, Δf – ширина спектральной линии, регистрируемой приёмником [20]. Так как для выбранного диапазона $\Delta f = 1,99 \cdot 10^{14}$ Гц и $\text{NEP} = 9,39 \cdot 10^{-6}$ Вт/нм, то получается минимальная мощность $W_{\min} = 8,45 \cdot 10^{-7}$ Вт.

Так как необходимая входную мощность есть $W_{\text{in}} = W_{\min} \cdot \text{SNR}$, где SNR – соотношения сигнал-шум, то для минимально необходимого для различения полезного сигнала, приняв $\text{SNR} = 10$, получаем входную мощность $W_{\text{in}} = 8,45 \cdot 10^{-6}$ Вт.

Далее, требуемая плотность мощности источника с учётом всех ослаблений линз, зеркал, призм и газов в диапазоне длин волн 1382–1383 нм [20] $W = W_{\text{in}} / T$, где T – произведение всех коэффициентов пропускания системы (отражение зеркал, пропускание линз и газов в кювете и тд). Таким образом, необходимая плотность мощности источника в диапазоне 1382–1383 нм равна 1,6 мВт/нм.

Для сравнения можно определить этот показатель с выбранным широкополосным источником излучения. Так как мощность 5 Вт является площадью под графиком функции Планка, определив цену деления одной относительной единицы интенсивности, получаем плотность мощности в диапазоне 1382–1383 нм составляет 10 мВт/нм, что, очевидно, больше требуемой величины, а значит, выбранная лампа удовлетворяет требованиям прибора.

Конструктивные особенности. В качестве держателей оптических компонентов разрабатываемого прибора выбраны коммерчески доступные элементы, за исключением самостоятельно разработанных держателей клиньев TWINS. Все держатели были выбраны, исходя исключительно из размеров оптических компонентов: держатель светоделительного куба BSH1/M производителя Thorlabs, держатель зеркал, линзы, TWINS-блока A SMR1 производителя Thorlabs, кольцо держателя зеркал, линзы, TWINS-блока A SM22RR производителя Thorlabs [16], моторизованный позиционер TWINS-блока Б с одноосным перемещением 30 мм SLC-2445 производителя SmarAct [21].

Конструкция имеет габаритные размеры 300x93x280 мм³, корпус прибора выполнен из конструкционного алюминия. Для удобства использования кромки скруглены, а шляпки винтов утоплены в корпусе. Необходимые отверстия (выходы каналов и провода кюветы, проводов моторизованного держателя, галогенной лампы и фотоприёмника) прорезинены прокладками для обеспечения герметичности корпуса (см. рис. 4).



Рис. 4. Внешний вид Фурье-спектрометра

Для обеспечения жёсткости конструкции корпус изготовлен методом лазерного спекания, наличие ячеек в корпусе также обуславливает лёгкость конструкции – 14 килограммов. Ход лучей в спектрометре схематически изображен на рис. 5.

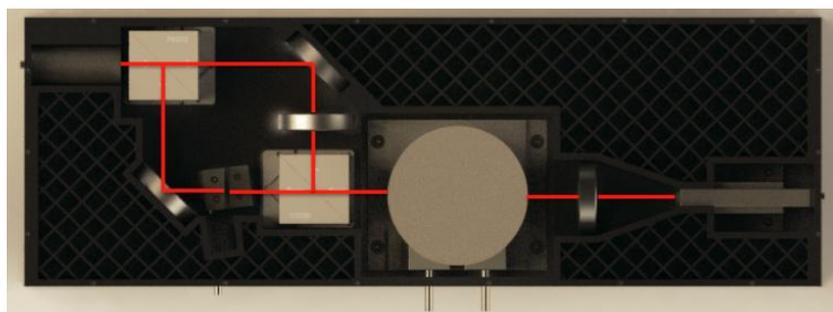


Рис. 5. Ход лучей в приборе

Выводы. В работе приведены результаты разработки усовершенствованной схемы компактного Фурье-спектрометра для анализа газового состава атмосферы в диапазоне длин волн 800–1700 нм.

Проведен анализ, в ходе которого была определена ключевая характеристика разрабатываемого модуля интерферометра – обеспечиваемая им разность хода. Показано, что при выборе оптического стекла N-BK7 в качестве материала блоков TWINS групповая скорость монотонно возрастает, а разность хода полуволн монотонно убывает увеличением длины волны. Кроме того, диапазон изменения разности хода полуволн увеличивается при увеличении толщины изменяемого блока TWINS. Показано, что в предложенной схеме спектрометра относительно небольшие клинья могут обеспечить разность хода от 2000 до 12000 полуволн.

Были проведены расчёты, на основании которых подобраны подходящие оптические элементы конструкции. В качестве компонентов разрабатываемого спектрометра были выбраны коммерчески доступные элементы, за исключением самостоятельно разработанных клиньев TWINS и их держателей.

В разработанной схеме удалось избежать использования полуволновой пластины и поляризатора, которые представляют собой дорогостоящие элементы, что привело к снижению себестоимости спектрометра в целом без потерь функциональности. Также, по сравнению со спектрометрами, основанными на интерферометре Майкельсона, предложенная схема обеспечивает более высокую точности настройки разности хода компонент луча за счёт малых углов при основании клина.

СПИСОК ИСТОЧНИКОВ

1. Белл, Р. Дж. Введение в Фурье-спектроскопию: перевод с англ. / Р. Дж. Белл. – М.: Мир, 1975. – 382 с.
2. Lagos, M. Fourier Transform Spectrometry with Fourier Analysis of the Interferogram as Just an Optional Tool / M. Lagos, R. Paredes, C. Retamal // ACS Omega. – 2018. – V. 3, № 12. – P. 18258-18262.
3. Волоконный флуоресцентно-отражательный спектрометр с многоволновым возбуждением / Г. В. Папаян, В. М. Журба, А. А. Кишалов [и др.] // Оптический журнал. – 2014. – Т. 81, № 1. – С. 38-43. – EDN TQJMLR.
4. Вагин, В. А. Система регистрации интерферограмм в двухзондовом (двухканальном) Фурье-спектрометре / В. А. Вагин, А. И. Хорохорин // Физические основы приборостроения. – 2020. – Т. 9, № 1(35). – С. 100-107. – DOI 10.25210/jfor-2001-0100107. – EDN JMDFGW.
5. Вагин, В. А. Современная Фурье-спектроскопия и быстрый нейроподобный метод снижения размерности спектральных данных / В. А. Вагин, А. Е. Краснов // Физические основы приборостроения. – 2020. – Т. 9, № 3(37). – С. 86-91. – DOI 10.25210/jfor-2003-086091. – EDN OTBXOC.
6. The POLARBEAR Fourier transform spectrometer calibrator and spectroscopic characterization of the POLARBEAR instrument. Review of Scientific Instruments / F. Matsuda, L. Lowry, A. Suzuki [et al.] // – 2019. – V. 90. – P. 115115. – DOI 10.1063/1.5095160.
7. Hashimoto, K. Phase-controlled Fourier-transform spectroscopy / K. Hashimoto, T. Ideguchi // Nature Communications. – 2018. – V. 9. – P. 4448. – DOI 10.1038/s41467-018-06956-x.
8. Рамановский спектрометр на основе статического интерферометра Майкельсона / И. Б. Винтайкин, Н. С. Васильев, И. С. Голяк [и др.] // Известия Российской академии наук. Энергетика. – 2016. – № 6. – С. 144-152. – EDN UZVBMW.
9. Brida, D. Phase-locked pulses for two-dimensional spectroscopy by a birefringent delay line / D. Brida, C. Manzoni, G. Cerullo // Optics Letters. – 2012. – V. 37. – P. 3027-3029. – DOI 10.1364/OL.37.003027.
10. Scanning Fourier transform spectrometer in the visible range based on birefringent wedges / A. Oriana, J. Rhault, F. Preda [et al.] // Journal of the Optical Society of America. – 2016. – V. 33, № 7. – P. 1415-1420. – DOI 10.1364/JOSAA.33.001415.
11. Long open-path measurements of greenhouse gases in air using near-infrared Fourier transform spectroscopy / D. W. T. Griffith, D. Pohler, S. Schmitt [et al.] // Atmospheric Measurement Techniques. – 2018. – V. 11. – P. 1549-1563.

12. On-chip parallel Fourier transform spectrometer for broadband selective infrared spectral sensing A. Fathy, Y. M. Sabry, S. Nazeer [et al.] // *Microsystems & Nanoengineering*. – 2020. – Vol. 6. – P.10. – DOI 10.1038/s41378-019-0111-0.
13. Тымкул, В. М. Оптико-электронные приборы и системы. Теория и методы энергетического расчёта / В. М. Тымкул, Л. В. Тымкул. – Новосибирск: СГГА, 2005. – 215 с.
14. База данных показателей пропускания сред [Электронный ресурс]. – URL: <http://pmoptics.com> (дата обращения 06.09.2024).
15. База данных показателей преломления сред [Электронный ресурс]. – URL: <http://refractiveindex.info> (дата обращения 06.09.2024).
16. Thorlabs [Электронный ресурс]. – URL: <https://www.thorlabs.com/> (дата обращения 20.09.2024).
17. Femto [Электронный ресурс]. – URL: <https://www.femto.de/en/products/photoreceivers/fixed-gain-400-mhz-hca-s-400m.html>. (дата обращения 06.09.2024).
18. Инфраспек. ИК Фурье-спектрометры ФСМ. Кювета газовая многоходовая КГ48 [Электронный ресурс]. – URL: <http://www.infraspek.ru/produktsiya/kyuvetyi/kyuvetyi-gazovyye/kyuveta-gazovaya-mnogohodovaya-kg48/> (дата обращения 20.09.2024).
19. Филачёв, А. М. Фотоприёмники в оптико-электронных приборах и системах / А. М. Филачёв, И. И. Таубкин, М. А. Трищенко. – М.: Физматкнига, 2016. – 105 с.
20. Mattson, D. Sensitivity of a Fourier transform infrared spectrometer / D. Mattson // *Applied Spectroscopy*. – 1978. – V. 32, №4. – P. 335-338.
21. SmarAct. Products [Электронный ресурс]. – URL: <https://www.smaract.com/products/> (дата обращения 06.09.2024).

Поступила в редакцию 16.10.2024 г., рекомендована к печати 05.11.2024 г.

IMPROVEMENT OF THE STRUCTURAL MODEL AND PARAMETERS OF A COMPACT FOURIER SPECTROMETER FOR ANALYSIS OF THE GAS COMPOSITION OF THE ATMOSPHERE

Akapev V.L., Savotchenko S.E.

A model of an improved structural model and parameters of a compact Fourier spectrometer for analyzing the gaseous composition of the atmosphere is proposed. The spectrometer scheme is based on the moving wedge identical pulse coding system (TWINS). The dependences of the half-wave path difference in the interferometer on the wavelength are analyzed by varying the thickness of the variable TWINS block. A calculation was made and a reasonable selection of system components was made.

Keywords: spectrometer, interferometer, Fourier spectrometer, photodetector, spectral analysis.

Акапьев Виктор Львович

кандидат педагогических наук, доцент кафедры информационно-компьютерных технологий в деятельности органов внутренних дел ФГКОУ ВО «Белгородский юридический институт Министерства Российской Федерации имени И.Д. Путилина»,
Российская Федерация, г. Белгород.
E-mail: akapevv1@yandex.ru

Akapev Viktor Lvovich

Candidate of Pedagogical Sciences, Associate Professor at Department of Information and Computer Technologies in the Activities of Internal Affairs Bodies of Belgorod Law Institute of the Ministry of the Russian Federation named after I.D. Putilin,
Russian Federation, Belgorod.

Савотченко Сергей Евгеньевич

доктор физико-математических наук, доцент, профессор кафедры высшей математики и физтики ФГБОУ ВО «Российский государственный геологоразведочный университет имени Серго Орджоникидзе»,
Российская Федерация, г. Москва.
E-mail: savotchenkose@mgi.ru

Savotchenko Sergei Evgenevich

Doctor of Physical and Mathematical Sciences, Associate Professor, Professor at Department of High Mathematics and Physics Sergo Ordzhonikidze Russian State University for Geological Prospecting Russian Federation, Moscow.

ОЦЕНКА МОДЕЛЕЙ ПОЛЕВЫХ ТРАНЗИСТОРОВ ДЛЯ ШИРОКОПОЛОСНЫХ УСИЛИТЕЛЕЙ КЛАССА D

© 2024 Казаков Е.А., Остапенков П.С., Соловьев Е.Д., Щукин А.В.

В настоящее время большинство разработок сложных систем осуществляется на основе расчета и предварительного функционального компьютерного моделирования в специализированных системах автоматизированного проектирования. В статье описываются результаты моделирования на основе SPICE-моделей ключевых усилителей мощности, которые используются для возбуждения гидроакустических антенн широкополосных систем подводной связи. В качестве активных элементов таких усилителей часто используются мощные МОП-транзисторы (MOSFET). В статье проводится сравнительный анализ транзисторов, выполненных по различным технологиям, приводятся результаты моделирования переходных процессов усилителей класса D, реализованных на основе мощных полевых транзисторов на основе кремния (Si), карбида-кремния (SiC), нитрида галлия (GaN). Полученные результаты могут быть использованы при проектировании широкополосных усилителей мощности класса D, в том числе и для систем гидроакустической связи.

Ключевые слова: ключевые усилители мощности, усилители класса D, широтно-импульсная модуляция, широкополосные усилители, гидроакустическая связь.

Введение. В современных системах гидроакустической связи для формирования акустических сигналов широкое применение находят широкополосные усилители мощности. Использование усилителей класса D в гидроакустических системах формирования сигналов позволяет существенно сократить весо-габаритные параметры элементов усилительного тракта, уменьшить потребляемую мощность по сравнению с использованием усилителей классов А-С. Применение компьютерного моделирования и специализированных систем автоматизированного проектирования (САПР), позволяет сократить время проектирования устройства, подобрать оптимальные параметры работы.

В статье рассматриваются вопросы моделирования ключевых усилителей мощности на основе SPICE-моделей для трактов формирования акустических сигналов систем гидроакустической связи. В качестве активных элементов усилителей используются мощные МОП-транзисторы (MOSFET). Проводится сравнительный анализ и результаты моделирования процессов переключения полевых транзисторов, выполненных по различным технологиям на основе кремния (Si), карбид-кремния (SiC), нитрид-галлия (GaN); проводится моделирование одноканальных и многоканальных ключевых усилителей мощности.

Модель двухтактного усилителя класса D. Обобщенная структурная схема усилителя мощности с ШИМ-модулятором изображена на рис. 1.

В общем случае усилитель широтно-импульсной модуляцией (ШИМ) сигнала состоит из следующих элементов: блок входной гальванической развязки, драйвер, усилитель, блок выходной развязки. Основное внимание в статье уделяется описанию моделирования блока Усилитель, показанному на рис. 1 в рамке, образованной сплошной линией.

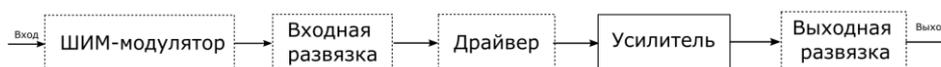


Рис. 1. Структурная схема усилителя

Рассмотрим структуру усилителя класса D на основе ШИМ более подробно. ШИМ-модулятор не входит в состав усилителя, но на основе этого блока осуществляется формирование требуемого информационного сигнала. Для формирования информационного сигнала можно использовать напрямую колебание с ШИМ, сформированное на основе цифровых методов с использованием процессоров, программируемых логических интегральных схем (ПЛИС), микроконтроллеров (МК).

Входная гальваническая развязка может быть выполнена на основе специализированных микросхем изоляторов на трансформаторах или оптронах. В нашем случае развязка реализована на основе микросхемы четырехканального изолятора трансформаторного типа ADUM340E.

Драйверы обеспечивают согласование токов и напряжений управляющего устройства (формирователя ШИМ сигнала) с рабочими напряжениями и токами затворов (баз) транзисторов усилителя. Драйверы включают схемы защиты для управляемых ими активных элементов и могут поддерживать работу источников напряжения противоположного знака для более эффективного запираания силовых ключей.

Блок выходной развязки реализуется на основе трансформатора, выполняющего функцию согласования и развязки усилителя с антенной. Усилитель может быть реализован по двухтактной схеме класса D или в смешанных режимах [1]: AD, BD, ABD и пр.

Ключевой режим является базовым режимом работы усилителей класса D. В этом режиме основное потребление энергии связано с динамическими потерями при переключении активных элементов, в то время, как в стационарных режимах потребление таких усилителей незначительно, поэтому использование усилителей класса D энергетически более эффективно по сравнению с усилителями других классов. Отдельно стоит отметить более компактные весо-габаритные параметры усилителей D класса и меньшее тепловыделение по сравнению с другими классами усилителей.

Применение полевых транзисторов в усилителях класса D обусловлено их крайне малыми токами потребления в закрытом состоянии и более высокими скоростями переключения, по сравнению с биполярными транзисторами. Отсутствие эффектов накопления и рассасывания неосновных носителей заряда в полевых транзисторах, обеспечивает их более высокое быстродействие по сравнению с другими видами транзисторов, например, биполярных транзисторов с изолированным затвором (IGBT).

Перечисленные преимущества усилителей D класса на полевых транзисторах для задач гидроакустической связи позволяют существенно повысить энергетическую эффективность передающих гидроакустических устройств, при одновременном уменьшении их габаритов.

Рассмотрим более подробно блок усилителя. Усилители ШИМ сигнала могут реализовываться по мостовой или полумостовой схемам [2]. Выходная нагрузка в обоих случаях подключается через гальваническую развязку или трансформатор.

Выходной трансформатор рассчитывается таким образом, чтобы сердечник не претерпевал насыщение во всей рабочей полосе частот работы усилителя, помимо этого необходимо добиться минимизации индуктивности рассеяния трансформатора и, как следствие, уменьшения обратного выброса тока при переключении активных элементов [3]. При моделировании усилителя следует учитывать и параметры нагрузки.

В САПР в проекте ключевого усилителя и при анализе быстродействия его активных элементов использовались SPICE-модели производителей, соответствующих компонентов. В представленных моделях не учитывалось влияние входных цепей усилителя и драйверов, поскольку данные блоки практически не оказывают влияния на параметры входного ШИМ-сигнала.

Моделирование усилителей можно проводить в различных САПР, поддерживающих SPICE формат моделей электронных компонентов, таких как ADS, LTspice, TINA TI, Qucs-S, Micro-Cap и др., но следует учитывать специфические особенности SPICE-моделей, предоставляемых разными производителями (SPICE, PSPICE, XSPICE, PLES и другие форматы) и их взаимодействие с соответствующей САПР.

В статье представлены результаты моделирования, которое проводилось с использованием САПР LTspice.

Формирование ШИМ. Усилители класса D работают в ключевом режиме, поэтому активный(ы) элемент(ы) может находиться только в открытом или закрытом состоянии и управление такими усилителями осуществляется с помощью ШИМ.

Качество выходного сигнала определяется верхней частотой рабочего диапазона f_B , частотой ШИМ $f_{\text{ШИМ}}$, а также частотной характеристикой нагрузки.

В гидроакустических системах связи частота f_B может варьироваться от десятков кГц до сотен кГц, при этом частота $f_{\text{ШИМ}}$ должна подбираться приблизительно на порядок больше верхней частоты f_B спектра сигнала.

Для того чтобы входной управляющий ШИМ-сигнал не влиял на качество усиления, должно выполняться следующее соотношение [1, 4]: $f_{\text{ШИМ}}/f_B \approx 15 - 20$. На практике усилители с ШИМ широко применяются на частотах до 100-120 кГц, но в ряде случаев возможны и более высокочастотные реализации ШИМ.

Следует учитывать, что с ростом частоты следования ШИМ импульсов возрастают динамические потери в активных элементах и, соответственно, повышаются требования к их задержкам переключения.

Существуют разные схемотехнические реализации усилителей класса D. Наиболее широкое применение получила симметричная двухсторонняя ШИМ первого рода (ДШИМ-1). Двухсторонняя ШИМ означает, что размах и уровни модулирующего и модулируемого сигналов совпадают. Понятие симметричной ШИМ относится к таким, где в качестве модулирующего используются симметричные относительно полупериода треугольные импульсы, как показано на рис. 2.

Род ШИМ, первый второй или третий, описывает процедуру сравнения модулирующего и модулируемого сигналов в компараторе.

Структурная схема ДШИМ-1 представлена на рис. 2.

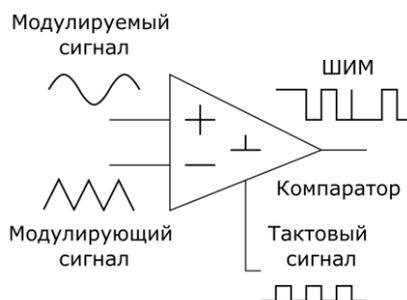


Рис. 2. Структурная схема блока формирования симметричной ДШИМ-1

Отличием ДШИМ-1 является низкий уровень комбинационных составляющих спектра модулированной последовательности импульсов и отсутствие комбинационных составляющих в спектре полезного низкочастотного (НЧ) сигнала [4, 5].

На практике формирование модулированных импульсов при ДШИМ-1 осуществляется посредством сравнения усиливаемого сигнала с симметричным линейно изменяющимся модулирующим сигналом, который формируется на тактовой частоте ШИМ-модулятора.

ШИМ-модулятор может быть реализован на основе цифровых схем, например, на базе микроконтроллеров (МК), процессоров или программируемых логических интегральных схем (ПЛИС).

Для реализации многоканальных усилителей с ШИМ более предпочтительным будет выбор ПЛИС, ввиду их большей гибкости и возможности одновременной параллельной обработки всех каналов ШИМ. Другим возможным вариантом является формирование ШИМ на основе специализированных сигнальных процессоров, с аппаратной поддержкой прецизионной ШИМ таких, как TMS320x280x.

Выбор активных элементов. В настоящее время доступны полевые транзисторы, выполненные из различных полупроводниковых материалов. Наиболее широко распространенные - это кремний (Si), карбид-кремния (SiC), нитрид-галлия (GaN) [6, 7]. С точки зрения быстродействия и динамических потерь, при равных рабочих токах, напряжениях и температурных режимах, карбил-кремниевые (SiC МОП) транзисторы обладают рядом преимуществ перед кремниевыми (Si МОП):

- меньшая входная емкость и заряд затвора,
- меньшее время включения/выключения,
- меньшее время восстановления обратного диода,
- меньшее сопротивление сток-исток в открытом состоянии.

В свою очередь уже нитрид-галлиевые транзисторы, которые относятся к классу транзисторов с высокой подвижностью электронов – ТВПЭ (HEMT), имеют ряд преимуществ перед карбид-кремниевыми:

- меньшая входная и проходная емкость и заряд затвора,
- меньшее сопротивление сток-исток в открытом состоянии.

При моделировании использовались транзисторы с близкими основными характеристиками, а именно: максимальное напряжение сток-исток $V_{DS} = 650$ В, ток стока $I_D = 20$ А и максимальный импульсный ток стока $I_{D(pulse)} = 48$ А.

Для сравнения использовались полевые транзисторы на основе кремния IPW65R190C7 с сопротивлением канала в открытом состоянии $R_{DS(on)} = 168$ мОм и зарядом затвора $Q_G = 23$ нКл, карбид-кремниевые IMW65R107M1H $R_{DS(on)} = 107$ мОм $Q_G = 15$ нКл, нитрид-галлиевые GS66506T $R_{DS(on)} = 67$ мОм $Q_G = 4,5$ нКл.

Все транзисторы - производства компании Infineon. Следует отметить, что на текущий момент, последняя из перечисленных серий транзисторов, находится в разработке и производитель пока не предоставил для неё окончательные SPICE-модели всех уровней.

Моделирование процессов переключения. Для обоснования выбора активных элементов усилителя проводилось моделирование процессов переключения полевых транзисторов на основе схемы, предложенной производителем в САПР LTspice [8].

На рис. 3 показана модель схемы для анализа импульсов на затворе и стоке полевого транзистора Si МОП IPW65R190C7. Модель включает в себя транзистор с обозначением M1. На рис. 3 изображен кремниевый транзистор IPW65R190C7 с моделью уровня L1.

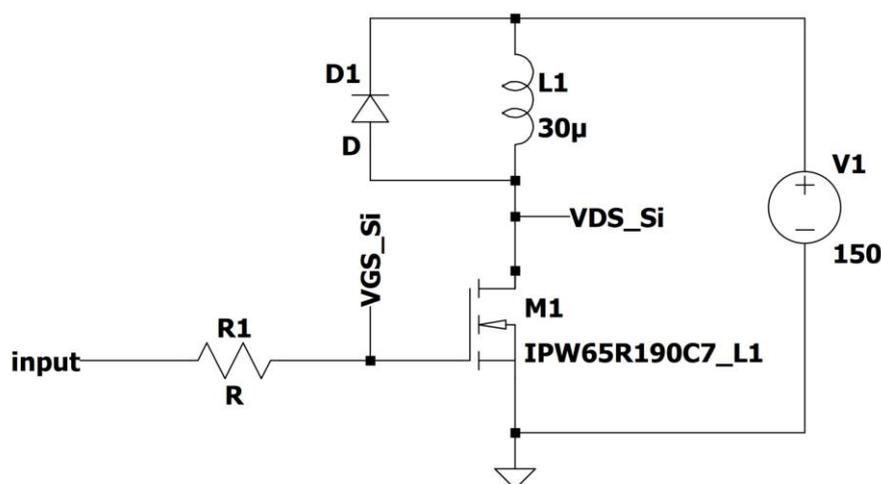


Рис. 3. Модель для измерения динамических характеристик транзистора в САПР LTspice

Далее в тексте модели транзисторов будут отмечаться суффиксом Lx и в тексте статьи выделяться курсивом. На рисунках из САПР уровень модели указан в конце названия соответствующего транзистора (IPW65R190C7_L1). Уровни моделей транзисторов будут подробнее рассмотрены в следующей части статьи.

Ток затвора ограничивается резистором R1. Номинал резистора R1 подбирается из значений для измерения быстродействия транзистора, приводимых в документации. Для кремниевого транзистора $R1=10$ Ом, для карбид-кремниевого и нитрид-галлиевого - $R1=2$ Ом.

Исследуемые транзисторы нагружены индуктивностью L1 с параллельно включенным диодом D1. Диод D1 полагается идеальным и необходим для предотвращения выброса напряжения на стоке в момент переключения. Питание схемы осуществляется от источника V1 напряжением 150 В.

На основе модели рис. 3. анализируются временные зависимости напряжений на затворах $V_{GS}(t)$ и стоках $V_{DS}(t)$ соответствующих транзисторов с индуктивной нагрузкой. В силовой электронике при работе с реактивной нагрузкой, индуктивный вид нагрузки встречается достаточно часто.

На рис. 4 приведены результаты моделирования для импульса напряжения на затворе длительностью 1 мкс и длительностями фронта, среза по 6 нс.

На рис. 4, 5, 6 эпюры напряжений (токов) на затворе и стоке нормируются по оси ординат и приводятся к единому масштабу для удобства анализа.

Результаты моделирования кремниевых IPW65R190C7, карбид-кремниевых IMW65R107M1H транзисторов соотносятся с параметрами из документации.

При моделировании времени включения t_{on} измерения проводились между уровнями от 0,1 фронта импульса напряжения затвор-исток (V_{GS}) до уровней 0,1 спада импульса напряжения сток-исток (V_{DS}) относительно максимального значения.

Полученные при моделировании значения времен включения t_{on} для кремниевых и карбид кремниевых транзисторов получилось близким, и как показано на рис.4(а, в) составляют 8 нс для IPW65R190C7, 9 нс для IMW65R107M1H. В документации производителя приводятся значения максимальных времен включения до $t_{on} = 22$ нс для IPW65R190C7 и до 20,6 нс для IMW65R107M1H.

Расхождения с параметрами из документации могут быть связаны с тем, что при моделировании учитывались следующие характеристики входного импульса: длительностью 1 мкс, длительности среза и среза 6 нс, но в документации производителя информация о длительностях фронта и среза для входного импульса не указана, что может сказываться на представлении точности модели.

Следует также отметить, что время переключения, приводимые в документации, измеряются при различных токах стока $I_{D\ Si} = 5,7\text{ А}$ и $I_{D\ SiC} = 8,9\text{ А}$ и напряжениях затвор-исток $V_{GS\ Si} = 13\text{ В}$ $V_{GS\ SiC} = 18\text{ В}$. При моделировании и сравнении параметров транзисторов, использовалась единая схема, представленная на рис. 3, за исключением номинала сопротивления затвора. Параметры транзисторов измерялись при индуктивной нагрузке $L1=30\text{ мкГн}$ с входным импульсом напряжения длительностью 1 мкс, током стока $I_D = 5\text{ А}$.

По сравнению с кремниевыми транзисторами, карбид-кремниевые транзисторы закрываются значительно быстрее. В документации приводится время выключения $t_{off} = 63\text{ нс}$ для кремниевого IPW65R190C7 и $t_{off} = 18,8\text{ нс}$ для карбид-кремниевого IMW65R107M1H транзисторов.

Время выключения, рис. 4 (б), (г), измеряется по уровню 0,9 от спада напряжения затвор-исток (V_{GS}) до уровня 0,9 фронта импульса напряжения сток-исток (V_{DS}).

При моделировании время выключения кремниевого транзистора IPW65R190C7 составляет 33 нс, для карбид-кремниевого IMW65R107M1H – 19 нс.

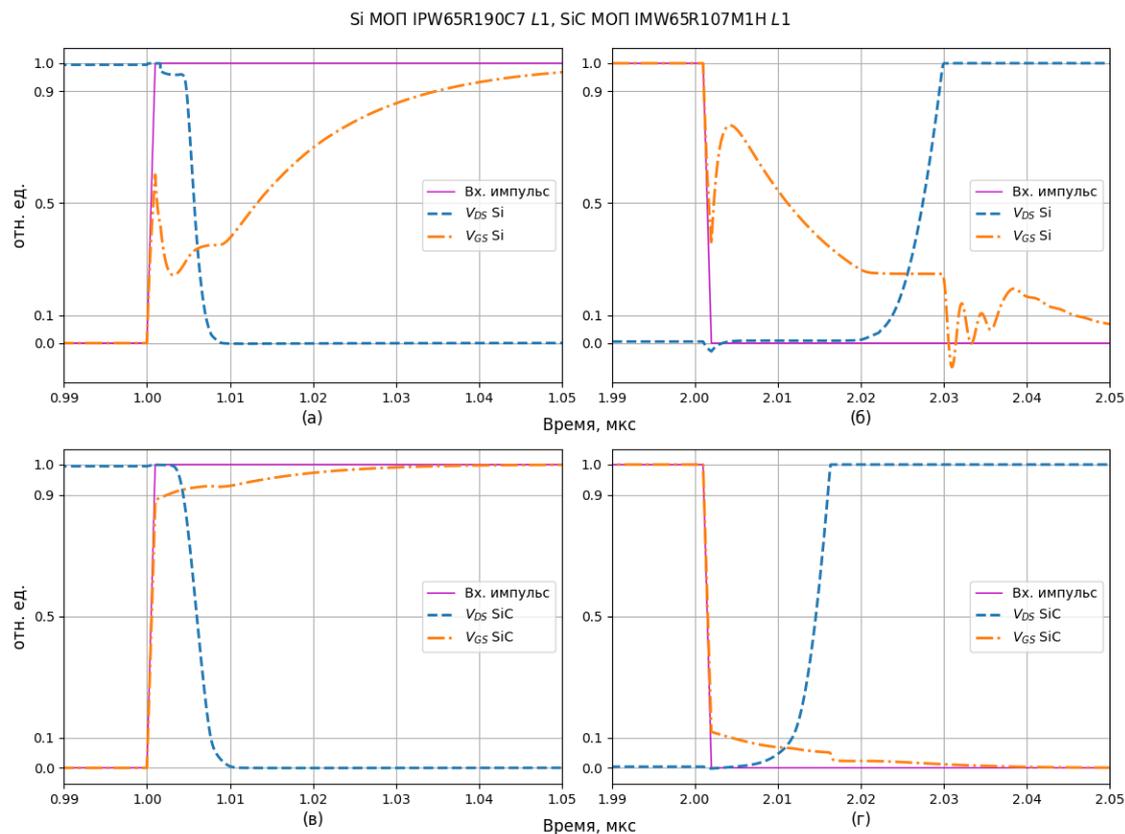


Рис. 4. Эпюры напряжений в момент переключения кремниевого (IPW65R190C7) (а), (б) и карбид-кремниевого (IMW65R107M1H) (в), (г) транзисторов: $V_{DS} = 150\text{ В}$ $V_{GS} = 18\text{ В}$ $I_D = 5\text{ А}$

Моделирование карбид-кремниевых IMW65R107M1H и нитрид-галлиевых GS66506T транзисторов имеет ряд отличий.

Транзисторы GS66506T имеют другой диапазон допустимых напряжений затвористок, поэтому $V_{GS\text{ GaN}} = 18\text{ В}$, также в документации не приводятся значений задержек включения и выключения, отсутствует схема измерения этих параметров, поэтому для проверки использовалась схема подключения аналогичная предыдущему эксперименту (см. рис. 3 при длительности входного импульса 1 мкс). Результаты моделирования представлены на рис. 5.

По результатам моделирования получились следующие значения для времен включения и выключения нитрид-галлиевых транзисторов $t_{on\text{ GaN}} = 2,3\text{ нс}$, $t_{off\text{ GaN}} = 12,5\text{ нс}$.

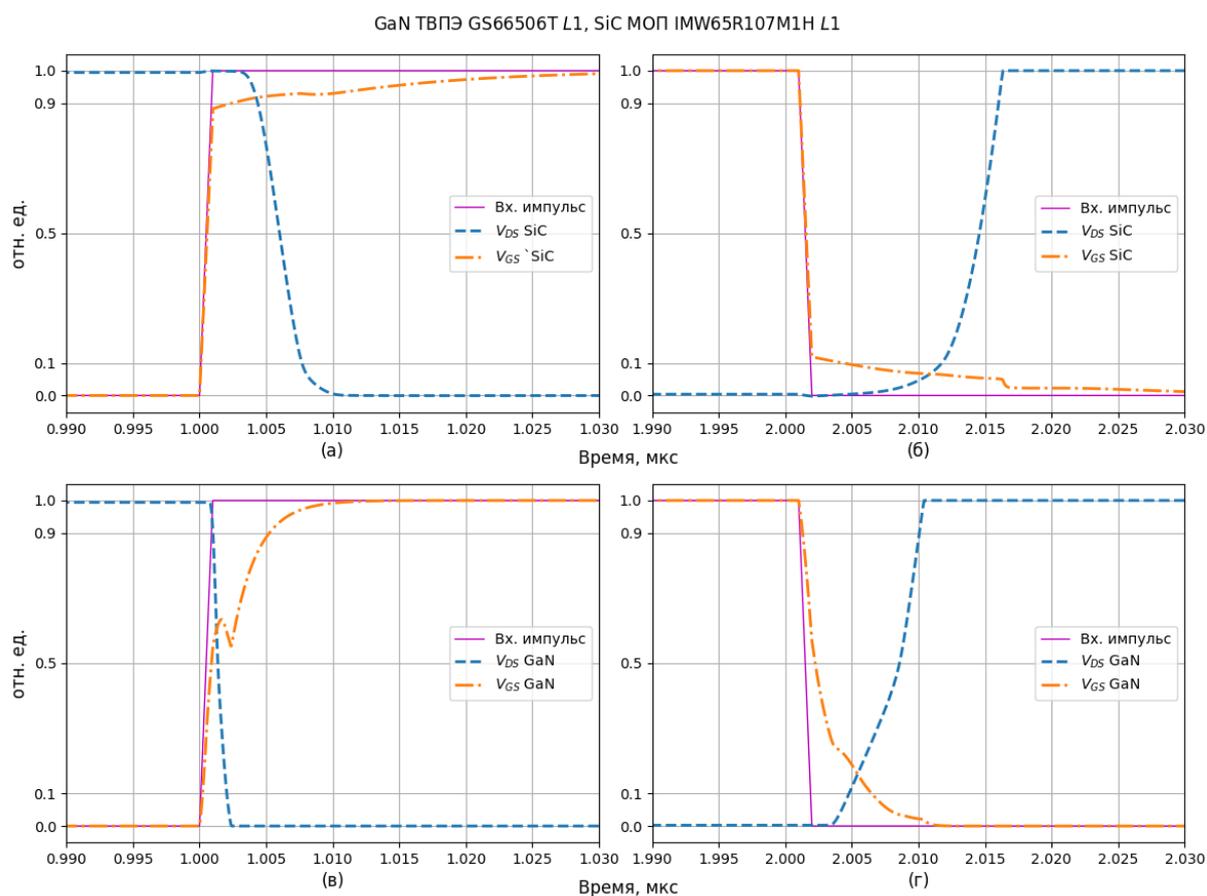


Рис. 5. Эпоуры напряжений в момент переключения карбид-кремниевое (IMW65R107M1H) (а), (б) и нитрид-галлиевое (GS66506T) (в), (г) транзисторов: $V_{DS} = 150\text{ В}$ $V_{GS} = 18\text{ В}$ $I_D = 5\text{ А}$

По итогам моделирования можно сделать вывод, что при необходимости повышенных требований к быстродействию усилителей класса D, при реализации например, прецизионных ШИМ сигналов и для минимизации динамических потерь предпочтительнее использовать карбид-кремниевые или нитрид-галлиевые транзисторы.

Уровни моделирования. Многие производители электронных компонентов предоставляют SPICE-модели для схемотехнического моделирования, но каждый производитель может придерживаться своих собственных подходов по описываемым параметрам и форматам SPICE-моделей.

В случае компании Infineon для полевых транзисторов возможны 4 уровня детализации моделей с 0 по 3. Каждый уровень отмечается суффиксом Lx в имени модели в библиотеке.

Нулевой уровень $L0$ является базовым. Модели нулевого уровня состояются из директив SPICE и других вложенных компонентов и библиотек.

Модели первого уровня $L1$ учитывают больше функциональных зависимостей. Например, для полевых транзисторов учитываются зависимости входной C_{iss} , выходной C_{oss} , обратной C_{oss} емкостей от приложенного напряжения, влияние температуры, которая устанавливается неизменной на протяжении всей симуляции.

SPICE-модели уровня $L2$ и выше описывают параметры транзисторов более детально, но представлены в открытом доступе для сравниваемых полевых транзисторов Infineon. Уровни детализации моделей $L2-L4$ предназначены для моделирования устройств с различными температурами в различных областях кристалла, и учетом изменений на протяжении длительности симуляции. Например, в модели учитываются влияние температуры корпуса T_p , температуры p-n-перехода T_j и другие параметры.

Следует отметить, что модели Infineon в первую очередь разрабатываются для симуляторов SiMetrix™, LTspice, MultiSim.

На рис. 6 показаны напряжения на затворах V_{GS} и стоках V_{DS} при переключении кремниевого МОП транзистора IPW65R190C7 для моделей уровней $L0$, $L1$. Различия в поведении моделей обусловлены более точным, зависящим от напряжения, моделированием емкостей транзистора для модели уровня $L1$.

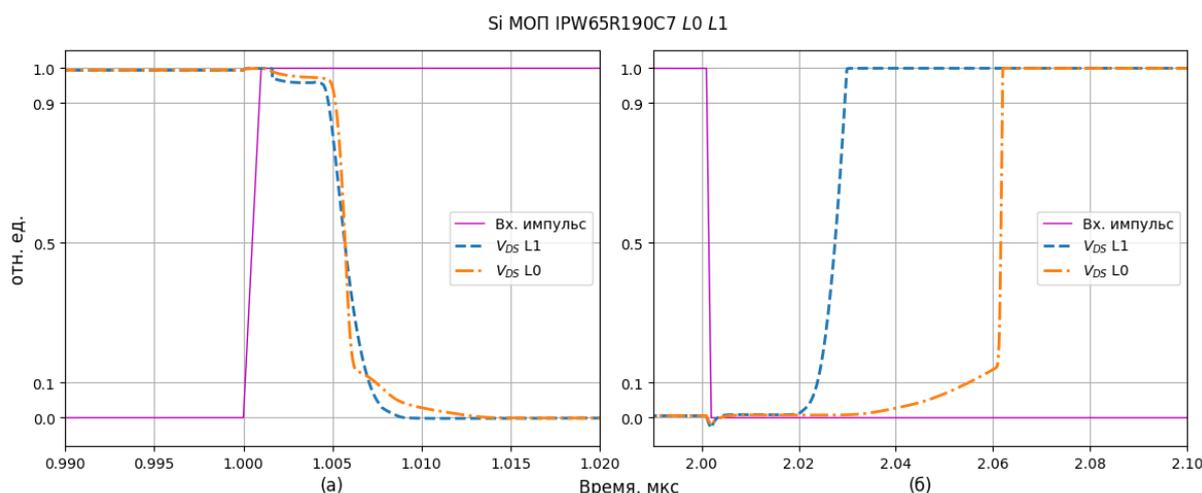


Рис. 6. Моделирование процесса переключения кремниевого IPW65R190C7 транзистора для моделей уровней $L0$, $L1$: $V_{DS} = 150$ В $V_{GS} = 18$ В $I_D = 5$ А. Фронт входного импульса (а), срез (б)

Моделирование двухтактных ключевых усилителей с ШИМ. При проведении моделирования использовались идентичные схемы подключения двухтактных усилителей.

В моделях использовались полумостовая (комплемментарную) и симметричная схемы. Использование двух разных схем позволяет провести моделирование процессов переключения транзисторов, включенных как последовательно, так и параллельно [2].

В моделях не использовались дополнительные фильтрующие цепи. Моделирование схем проводилось с учетом развязывающего трансформатора и индуктивной нагрузки, подключенной к стокам соответствующих транзисторов усилителя.

На рис. 7 показана модель двухтактного ключевого усилителя, выполненного по полумостовой схеме.

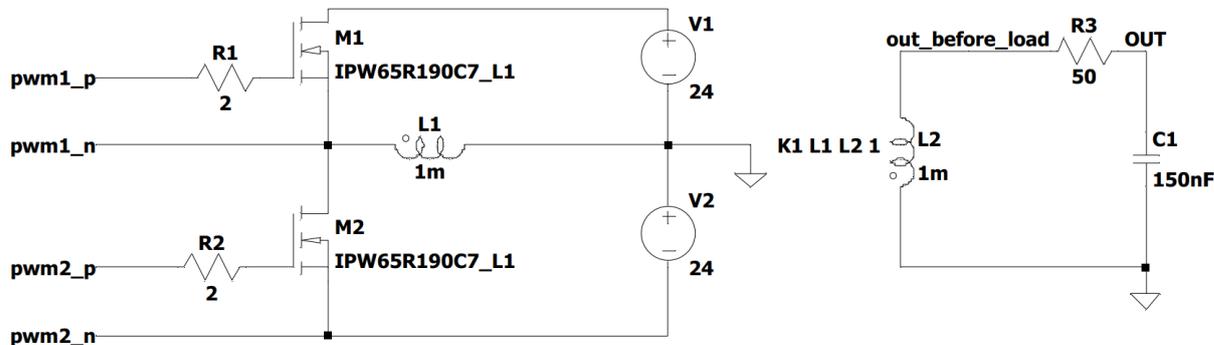


Рис. 7. Модель полумостового двухтактного ключевого усилителя в САПР LTspice

Модель состоит из двух транзисторов M1, M2 с резисторами в цепях затвора R1, R2 и выходного трансформатора с обмотками L1, L2, сердечником K1.

Аналогичные обозначения приняты и для модели симметричной схемы включения транзисторов, представленной на рис. 8. В этом случае выходной трансформатор состоит уже из двух первичных обмоток L1, L2 и вторичной обмотки L3. Оба трансформатора моделируются как идеальные с линейной связью между обмотками. В качестве широкополосной нагрузки для всех моделей выступает RC-цепь с частотой среза 20 кГц. На вход каждого из двухтактных усилителей подается гармонический сигнал с частотой 15 кГц, который модулируется ШИМ на частоте $f_{\text{ШИМ}} = 200$ кГц.

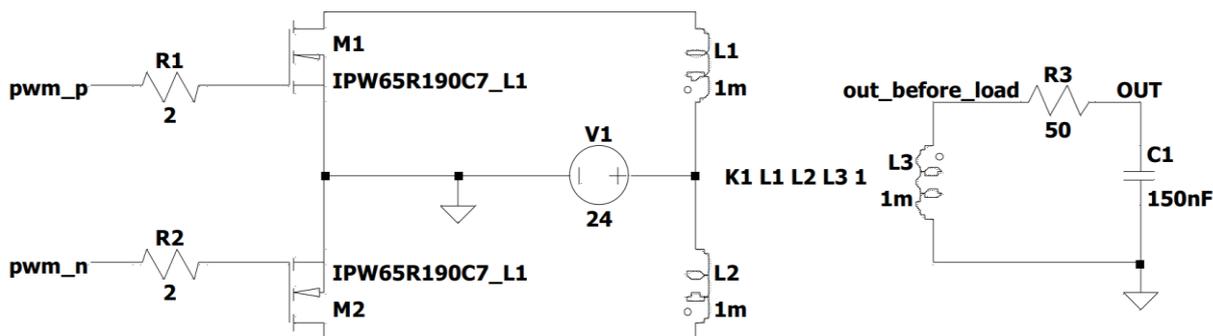


Рис. 8. Модель симметричного двухтактного ключевого усилителя в САПР LTspice

При моделировании импульсов учитывалось напряжение на вторичных обмотках трансформаторов для каждой из схем с кремниевыми и карбид-кремниевыми полевыми транзисторами. На схемах рис. 7, 8, обозначения L2, L3 соответствуют вторичным обмоткам трансформаторов.

Результаты моделирования представлены на рис. 9, 10.

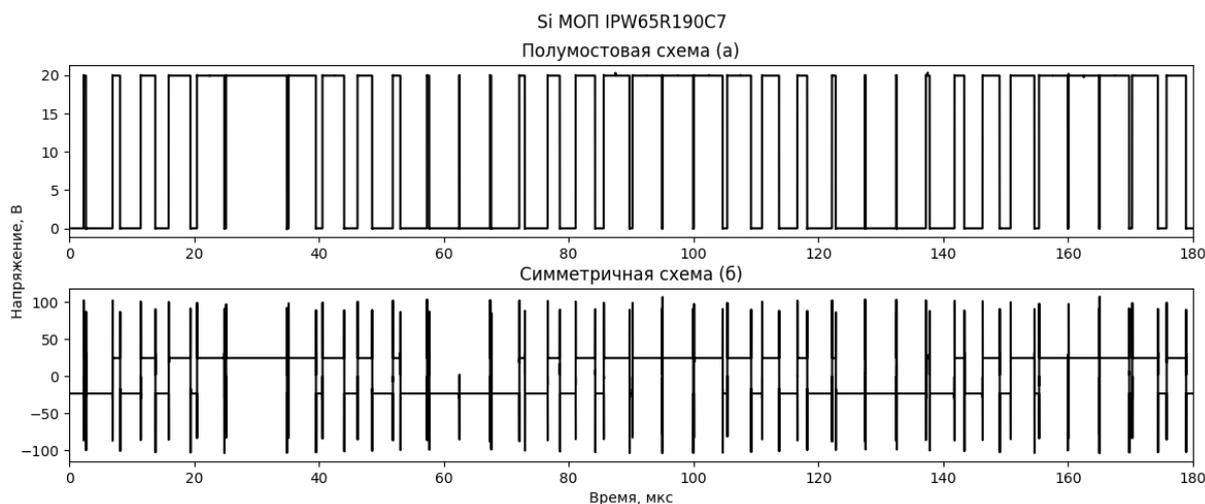


Рис. 9. Эпюры напряжений на вторичных обмотках выходных трансформаторов для кремниевого транзистора IPW65R190C7: (а) полумостовая схема, (б) симметричная схема

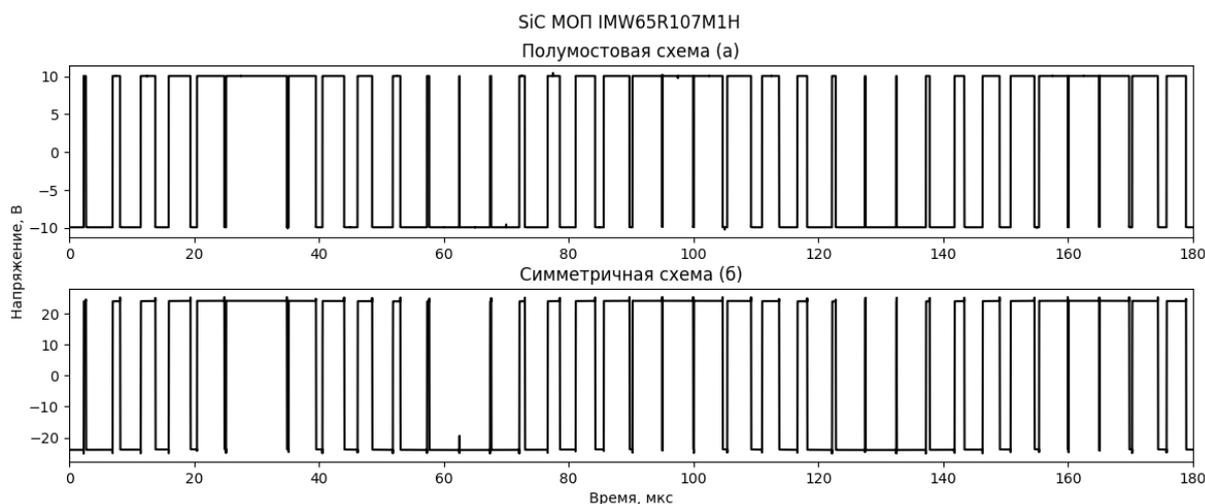


Рис. 10. Эпюры напряжений на вторичных обмотках выходных трансформаторов с карбид-кремниевыми транзисторами IMW65R107M1H: (а) полумостовая схема, (б) симметричная схема

Повышение быстродействия ключей положительно сказывается на работе симметричной схемы, максимальные выбросы выходного напряжения при переключении плеч двухтактного каскада с карбид-кремниевыми транзисторами практически отсутствуют (см. рис. 9 (б), рис. 10 (б)). Такой положительный эффект объясняется снижением выходной емкости C_{iss} карбид-кремниевых транзисторов по сравнению с кремниевыми транзисторами.

В то же время, полумостовая схема работает одинаково стабильно без выбросов выходного напряжения (см. рис. 9 (а), рис. 10 (а)), поскольку в полумостовой схеме используется выходной трансформатор с одной первичной обмоткой L1 и выбросов тока с соседних обмоток L1, L2 не происходит (см. рис. 7).

Дополнительным достоинством полумостовой схемы является возможность непосредственного подключения нагрузки без использования трансформатора.

С другой стороны, использование симметричной схемы более предпочтительно для реализации согласования и развязки на основе трансформаторной связи с нагрузкой. Следует обратить внимание, что в реальных устройствах на выходах транзисторов устанавливаются фильтрующие цепи, для фильтрации высокочастотных составляющих напряжения (тока) при переключении.

Выводы. Проведено моделирование процессов переключения трех высокоомощных полупроводниковых транзисторов, реализованных по разным технологиям: кремниевый (Si), карбид-кремниевый (SiC) и нитрид-галлиевый (GaN).

На основе моделирования в САПР LTspice показано, что полевые карбид-кремниевые (SiC МОП) транзисторы обладают лучшим быстродействием, по сравнению с близкими по параметрам кремниевыми (Si МОП) транзисторами.

При сопоставимом времени включения были получены следующие соотношения: $t_{on Si}/t_{on SiC} = 9 \text{ нс}/8 \text{ нс} = 1,1$ раза, для времени выключения в $t_{on Si}/t_{on SiC} = 33 \text{ нс}/19 \text{ нс} = 1,7$ раза меньше у карбид-кремниевых транзисторов (см. рис. 4). Однако, несмотря на практически одинаковое время включения, у карбид-кремниевых транзисторов менее выражен эффект Миллера из-за меньших значений входной емкости: $C_{iss Si}/C_{iss SiC} = 1150 \text{ пФ}/496 \text{ пФ} = 2,5$ раза (см. рис. 4 (а), (в)).

Моделирование ключевых усилителей показало, что использование транзисторов на основе карбида-кремния (SiC) более предпочтительно в мощных ключевых усилителях, где более быстродействующие ключи способствуют уменьшению выбросов напряжения (тока) в выходном трансформаторе ключевого каскада (см. рис. 9(б), рис.10(б)). В тоже время нитрид-галлиевые (GaN) транзисторы могут использоваться в перспективных разработках ключевых усилителей мощности с прецизионной ШИМ, поскольку потенциально обладают лучшим быстродействием, чем карбид-кремниевые (SiC) полевые транзисторы, но оценка этого потенциала, осложняется тем, что многие производители (GaN) транзисторов не предоставляют подробную документацию и SPICE модели для моделирования в САПР.

СПИСОК ИСТОЧНИКОВ

1. Динамические потери энергии в ключевых усилителях мощности в составе гидроакустического передающего тракта / А. В. Александров, О. В. Воробьев, Ю. В. Казаков, Л. В. Маркова // Труды учебных заведений связи. – 2022. – № 3. – С. 14–26. – DOI 10.31854/1813-324X-2022-8-3-14-26.
2. Grebennikov, A. Switchmode RF and microwave power amplifiers /A. Grebennikov, N. O. Sokal, M. J. Franco // Academic Press, 2012. – 677 p.
3. Михеев, Г. М. Методы определения индуктивности рассеяния обмоток силового трансформатора / Г. М. Михеев, В. М. Шевцов, Т. Г. Иванова // Вестник Чувашского университета. – 2009. – № 2. – С. 147-153.
4. Исследование цифровых методов генерации сигналов гидроакустических фазированных антенных решеток / В. А. Александров, А. П. Буянов, Л. В. Маркова, М. А. Сиверс // Труды учебных заведений связи. – 2021. – № 1. – С. 42–53. – DOI 10.31854/1813-324X-2021-7-1-42-53.
5. Александров, В. А. Применение периодических широтно-модулированных импульсных последовательностей в цифровых генераторных устройствах высокочастотных гидроакустических передающих трактов / А. В. Александров // Гидроакустика. – 2019. – № 39(3). – С. 72–80.
6. Wang, J. A Comparison between Si and SiC MOSFETs / J. Wang // IOP Conference Series: Materials Science and Engineering. – 2020. – V. 729(1). – DOI 10.1088/1757-899x/729/1/012005.
7. Langpoklakpam, C. Review of Silicon Carbide Processing for Power MOSFET / C. Langpoklakpam, A.-C. Liu, K.-H. Chu // Crystals. – 2022. – V. 12(2). – P. 245. – DOI <https://doi.org/10.3390/cryst12020245>.
8. Alhalabi, M. Modelling of SiC power MOSFET in Matlab, Simulink, and LTspice / M. Alhalabi // 2018 IEEE International Conference on Environment and Electrical Engineering and 2018 IEEE Industrial and Commercial Power Systems Europe (EEEIC/I&CPS Europe). – 2018. – P. 1-6.

Поступила в редакцию 08.11.2024 г., рекомендована к печати 20.11.2024 г.

MODELS ANALYSIS OF SWITCHING MOSFETS FOR WIDEBAND D CLASS AMPLIFIERS

Kazakov E.A., Ostapenkov P.S., Solovev E.D., SHCHukin A.V.

The developing of the complex radio systems strongly related to preliminary computer modeling stage with CAD systems. The purpose of this paper is SPICE modeling switching power amplifiers for wideband underwater acoustic communication systems. This paper provides D class amplifiers transient analysis on high power MOSFETs based on different technology semiconductors: silicon (Si), silicon carbide (SiC) and gallium nitride (GaN). The results can be applied for hydroacoustic wideband communication systems.

Keywords: switching power amplifiers, D class amplifiers, pulse-width modulation, wideband amplifiers, hydroacoustic system.

Казиков Евгений Александрович

аспирант кафедры формирования и обработки радиосигналов ФГБОУ ВО «Национальный исследовательский университет "Московский Энергетический Институт"»,
Российская Федерация, г. Москва.
E-mail: KazakovYevA@mpei.ru

Kazakov Evgenii Aleksandrovich

Postgraduate Student at Department of Formation and Processing of Radio Frequency Signals of Moscow Power Engineering Institute,
Russian Federation, Moscow.

Остапенков Павел Сергеевич

кандидат технических наук, доцент, заведующий кафедрой формирования и обработки радиосигналов ФГБОУ ВО «Национальный исследовательский университет "Московский Энергетический Институт"»,
Российская Федерация, г. Москва.
E-mail: OstapenkovPS@mpei.ru

Ostapenkov Pavel Sergeevich

Candidate of Technical Sciences, Associate Professor, Head of the Department of Formation and Processing of Radio Frequency Signals of Moscow Power Engineering Institute,
Russian Federation, Moscow.

Соловьев Евгений Дмитриевич

ассистент кафедры формирования и обработки радиосигналов ФГБОУ ВО «Национальный исследовательский университет "Московский Энергетический Институт"»,
Российская Федерация, г. Москва.
E-mail: SolovyevYD@mpei.ru

Solovev Evgenii Dmitrievich

Assistant at Department of Formation and Processing of Radio Frequency Signals of Moscow Power Engineering Institute,
Russian Federation, Moscow.

Щукин Александр Викторович

старший преподаватель кафедры формирования и обработки радиосигналов ФГБОУ ВО «Национальный исследовательский университет "Московский Энергетический Институт"»,
Российская Федерация, г. Москва.
E-mail: ShchukinAV@mpei.ru

SHCHukin Aleksandr Viktorovich

Senior Lecturer at Department of Formation and Processing of Radio Frequency Signals of Moscow Power Engineering Institute,
Russian Federation, Moscow.

ИЗМЕРЕНИЕ ИНТЕНСИВНОСТИ ТУРБУЛЕНТНОСТИ ПО УПРОЩЕННОЙ МЕТОДИКЕ ГРАДУИРОВКИ ПРОВОЛОЧНОГО ДАТЧИКА ТЕРМОАНЕМОМЕТРА

© 2024 *Пометун Е.Д., Лебедев В.Н.*

В работе приводится описание упрощенной методики градуировки термоанемометра постоянной температуры (сопротивления) с проволочным первичным преобразователем, не требующей получения массива данных по двум независимым параметрам: скорости и температуре газового потока. Показано, что после ряда преобразований и сокращений уравнения теплообмена, для измерения интенсивности турбулентности достаточно измерить выходное напряжение термоанемометра при отсутствии скорости потока и установить оптимальный показатель степени, используя уравнение Кинга для нагретой нити.

Ключевые слова: термоанемометр постоянной температуры, градуировка датчика, интенсивность турбулентности, проволочный чувствительный элемент.

Введение. Термоанемометры постоянной температуры (ТА ПТ) широко используются в различных областях науки и техники [1]. Градуировка термоанемометров выполняется на специальных аэродинамических стендах, которые позволяют создавать установившийся поток воздуха, с возможностью независимого регулирования скорости и температуры. Данные устройства и поток в рабочей части аэродинамической трубы необходимо аттестовать в соответствии со специальными методиками. Необходимость индивидуальной градуировки ТА ПТ, учитывая низкую надежность проволочных (нитяных) датчиков, приводит к большим материальным затратам: во-первых, необходимо создавать заново или адаптировать существующие аэродинамические стенды и во-вторых, обработка массивов данных требует значительного времени. В данной работе описан способ измерения интенсивности турбулентности газового потока по упрощенной градуировке ниточного датчика ТА ПТ, не требующий специального оборудования.

Основная часть. Для измерения интенсивности турбулентности в качестве чувствительных элементов датчиков ТА ПТ в основном используется нити диаметром от 1 мкм до 8 мкм. Ниточные датчики обладают малой инерционностью и высокой чувствительностью к пульсациям скорости, что позволяет выполнять измерения в широком диапазоне чисел Струхала [1]. Расширение частотного диапазона, связано с уменьшением диаметра нити и, следовательно, снижением надежности датчика: в результате загрязнения и/или обрыва нити. При обрыве нити конструкция датчика не изменяется, восстановление работоспособности осуществляется путем замены нити. Восстановленные датчики нуждаются в повторной градуировке.

В данной работе приведены результаты измерения интенсивности турбулентности воздушного потока при упрощенной и стандартной методике градуировки ТА ПТ.

Градуировка датчика основана на предварительном измерении напряжений (токов) при различных скоростях и температурах воздушного потока. Дальнейшая обработка выходных данных ТА может выполняться по различным алгоритмам. В работах [2, 3] приводится способ, основанный на гипотезе о возможности представления сложной функции, описывающей теплообмен нагретого датчика ТА и зависящей от двух параметров: скорости и температуры, в виде произведения двух более простых функций (скорости потока воздуха и его температуры).

В общем случае, уравнение теплообмена для нагретого цилиндра (нити) описывается уравнением Кинга [1]:

$$Nu = 0.42 \cdot Pr^{0.2} + 0.5 \cdot Pr^{0.33} \cdot Re^{0.5}, \quad (1)$$

где Pr - число Прандтля, Re – число Рейнольдса.

Влияние различных параметров на показания ТА можно оценить, анализируя уравнение теплового баланса нити (1). Пренебрегая потерями на излучение и теплоотвод в державки, уравнение теплового баланса можно записать в виде [1]:

$$Q = I^2 R_w = \alpha \pi d l (\theta_w - \theta_g), \quad (2)$$

где: I – сила тока, R_w - сопротивление нити, α - коэффициент теплоотдачи, l – длина нити, d – диаметр нити, θ_w - температура нити, θ_g - температура газа.

Используя критерий Нуссельта $Nu = \frac{\alpha d}{\lambda_g}$, уравнение теплового баланса нити (2), с учетом уравнения (1), можно представить в виде:

$$I^2 R = \pi l \lambda_g (\theta_w - \theta_g) * 0,42 Pr^{0.2} + 0,57 Pr^{0.33} Re^{0.5}, \quad (3)$$

где λ_g - теплопроводность газа.

Для большинства двухатомных газов число Прандтля слабо зависит от температуры и для воздуха в диапазоне 18 – 45 °С и может быть принято равным $Pr=0,7$, что позволяет при обработке экспериментальных данных полагать $Nu = f(Re)$ [1].

В ограниченном диапазоне скоростей зависимость между теплообменом и числом Рейнольдса может быть представлена соотношением вида (4):

$$Nu = (A + B Re^n) \cdot \left(\frac{\theta_g}{\theta_w} \right)^{-m}, \quad (4)$$

где A, B – градуировочные коэффициенты.

Или с учетом (3):

$$\frac{I^2 R_w d}{\lambda_g S (\theta_w - \theta_g)} \left(\frac{\theta_g}{\theta_w} \right)^m = \frac{E_{TA}^2 d}{R_w \lambda_g S (\theta_w - \theta_g)} \left(\frac{\theta_g}{\theta_w} \right)^m = A + B Re^n. \quad (5)$$

где S – площадь поверхности нити.

Если принять число Рейнольдса равным нулю ($Re=0$), с учетом $A = f(I_0^2 R_w)$, соотношение (5) можно представить в виде:

$$\frac{(E_{TA}^2 - E_0^2)}{R_w \lambda_g S (\theta_w - \theta_g)} \left(\frac{\theta_w}{\theta_g} \right)^m = B Re^n, \quad (6)$$

где E_0 - выходное напряжение ТА ПТ при отсутствии скорости потока.

В соотношении (6) геометрические параметры, теплопроводность и сопротивление нити слабо изменяется, т.о. $\frac{1}{R_w \lambda_g S} \approx const$ и если принять показатель степени $m = 0$ в $\left(\frac{\theta_g}{\theta_w}\right)^m$, тогда операция термокомпенсации может быть выполнена путем деления левой части на безразмерный коэффициент $\frac{1}{(\theta_w - \theta_g)}$, причем полученное уравнение не будет зависеть от температуры потока.

Линеаризация полученной функции основана на преобразовании нелинейной зависимости, между термокомпенсированными значениями выходного электрического параметра и скоростью потока, с помощью полинома. В работах [2, 3] было определено, что оптимальная степень полинома (7) равна 4, причем при дальнейшем увеличении степени полинома погрешность не уменьшалась, тогда [2]:

$$V = \left(\frac{1}{B_1} (E_{TA}^2 - E_0^2) \right)^{1/n} = C_0 + C_1 E + C_2 E^2 + C_3 E^3 + C_4 E^4, \quad (7)$$

где B_1 - нормирующий коэффициент.

Вычисление коэффициентов полинома и значений термокомпенсированной функции выполняется по массивам данных, полученных при различной скорости и температуре потока, и требует значительных материальных затрат. Следовательно, измерение интенсивности турбулентности можно свести к решению уравнения (7) и подстановки полученных значений скорости V в уравнение (8) [4]:

$$\varepsilon = \frac{\sqrt{(V_i - \bar{V})^2}}{\bar{V}} \cdot 100\% . \quad (8)$$

где V_i - текущее значение скорости, \bar{V} - среднее значение скорости.

В данной работе был предложен упрощенный способ линеаризации, основанный на особенности уравнения (8): интенсивность турбулентности равна не абсолютному, а относительному значению величин. При этом постоянные величины, входящие в уравнение (6), включая функцию температуры, сокращаются, окончательно:

$$\varepsilon = \frac{\sqrt{(E_{TA_i}^2 - E_0^2)^{1/n} - (\bar{E}_{TA}^2 - E_0^2)^{1/n}}}{(\bar{E}_{TA}^2 - E_0^2)^{1/n}} \cdot 100\% . \quad (9)$$

где: E_0 - величина напряжения ТА при скорости $V=0$, E_{TA_i} - текущее значение напряжения, \bar{E}_{TA} - среднее значение напряжения.

Расчет интенсивности турбулентности по уравнению (9), заключается в градуировке (тарировке) датчика по одной скорости потока близкой к нулю, и не требует дополнительных данных для температурной компенсации.

На практике, измерения пульсаций скорости выполнялись следующим образом. Предварительно, проградуированный проволочный датчик ТА ПТ размещался на срезе конфузора аэродинамического стенда АДС 200/250 [5]. С помощью координатного устройства датчик перемещался от центра конфузора к его стенкам (в область развитого турбулентного течения) с шагом 10 мм. В каждой точке выполнялась регистрация данных ТА с помощью АЦП NI USB-6001 (время записи 10 секунд, частота дискретизации 1кГц). Было выполнено 3 серии измерений: при фиксированной температуре потока воздуха $\theta_g = 30^\circ\text{C}$ и скоростях 7 м/с, 9 м/с, 14 м/с.

Для проверки работоспособности упрощенного способа измерения интенсивности турбулентности, вначале был выполнен расчет с помощью стандартного способа по соотношению (7). На рисунке 1 приведены зависимости интенсивности турбулентности от координаты и при различных скоростях газового потока.

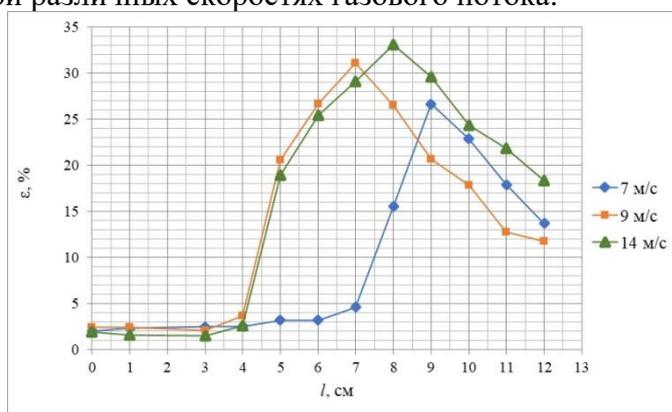


Рис. 1. Зависимость интенсивности турбулентности, от координаты при различных скоростях газового потока

На рисунке 2 представлена аналогичная зависимость, полученная после обработки массива данных по упрощенной методике по соотношению (9).

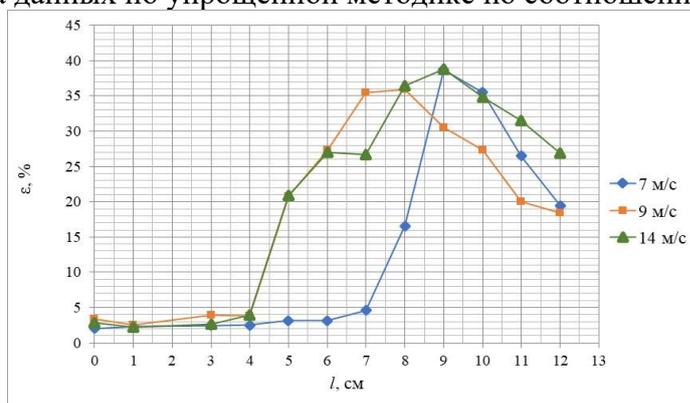


Рис. 2. Зависимость интенсивности турбулентности, по упрощенной методике градуировке, от координаты при различных скоростях газового потока

Сравнение результатов показывает, что интенсивность турбулентности, вычисленная по упрощенной методике больше, чем при полиномиальной аппроксимации. Эти различия связаны с рядом допущений, принятых при выводе уравнений, погрешностью температурной компенсации и линеаризации. Однако основной причиной, является выбор степени n при числе Re в уравнении (9). Например, в работе [1] приводится степень $n \geq 0,5$. При точном определении степени n расхождение между двумя способами, вероятно, уменьшится.

Выводы. Предложенный способ упрощенной градуировки проволочного датчика ТА ПТ основан на измерении выходного напряжения при отсутствии потока, и не требует наличия специализированных аэродинамических стендов, позволяющих независимо варьировать два параметра: скорость и температуру потока.

Различия между двумя способами измерения интенсивности турбулентности обусловлено рядом факторов, в том числе отличием показателем степени при числе Рейнольдса от фактического значения.

СПИСОК ИСТОЧНИКОВ

1. Ярин, Л. П. Термоанемометрия газовых потоков / Л. П. Ярин, А. Л. Генкин, В. И. Кулес. – Л.: Машиностроение, 1983. – 198 с.
2. Лебедев, В. Н. Термокомпенсация выходного сигнала термоанемометра на основании уравнения теплового баланса / В. Н. Лебедев, Е. Д. Пометун, А. В. Васильева // Оригинальные исследования. – 2022. – Т. 12, № 5. – С. 406-412. – EDN OWAUUQ.
3. Пометун, Е. Д. Исследование функций аппроксимации градуировочной характеристики термоанемометра в неизотермическом газовом потоке / Е. Д. Пометун, В. Н. Лебедев // Системный анализ и информационные технологии в науках о природе и обществе. – 2015. – № 1-2(8-9). – С. 88-91. – EDN XSASJB.
4. Пометун, Е. Д. Термоанемометрический метод измерения пульсаций скорости турбулентных потоков / Е. Д. Пометун // Вестник Донецкого национального университета. Серия Г: Технические науки. – 2023. – № 4. – С. 15-19. – EDN NBPFBG.
5. Кузнецов Д. Н. Исследование поля скоростей рабочего участка аэродинамической установки для градуировки датчиков / Д. Н. Кузнецов, А. А. Зори // Научные труды Донецкого национального технического университета. Серия: Вычислительная техника и автоматизация. – 2004. – Вып. 74. – С. 300-305.

Поступила в редакцию 11.11.2024 г., рекомендована к печати 22.11.2024 г.

MEASUREMENT OF TURBULENCE INTENSITY USING A SIMPLIFIED CALIBRATION METHOD FOR A WIRE SENSOR OF A THERMOANEMOMETER

Pometun E.D., Lebedev V.N.

The paper describes a simplified technique for calibration of a constant temperature (resistance) thermoanemometer with a wire primary transducer, which does not require obtaining an array of data on two independent parameters: velocity and temperature of the gas flow. It is shown that after a number of transformations and reductions of the heat transfer equation, it is sufficient to measure the output voltage of the thermoanemometer in the absence of flow velocity and to establish the optimal degree exponent using King's equation for a heated filament to measure the turbulence intensity.

Keywords: constant temperature thermoanemometer, sensor graduation, turbulence intensity, wire sensing element.

Пометун Екатерина Дмитриевна

кандидат технических наук, доцент кафедры физики неравновесных процессов, метрологии и экологии им. И.Л. Повха ФГБОУ ВО «Донецкий государственный университет»,
Российская Федерация, ДНР, г. Донецк.
E-mail: alternativa19031992@rambler.ru

Pometun Ekaterina Dmitrievna

Candidate of Technical Sciences, Associate Professor at Department of Physics of Nonequilibrium Processes, Metrology and Ecology named after I.L. Povkh of Donetsk State University,
Russian Federation, DPR, Donetsk.

Лебедев Владимир Николаевич

старший преподаватель, заведующий лабораторией кафедры физики неравновесных процессов метрологии и экологии им. И.Л. Повха ФГБОУ ВО «Донецкий государственный университет»,
Российская Федерация, ДНР, г. Донецк.

Lebedev Vladimir Nikolaevich

Senior Lecturer, Head of Laboratory at Department of Physics of Nonequilibrium Processes, Metrology and Ecology named after I.L. Povkh of Donetsk State University,
Russian Federation, DPR, Donetsk.

ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ И ТЕЛЕКОММУНИКАЦИИ

EDN VEQGYZ
УДК 65.011.56:681.5.08

DOI 10.5281/zenodo.14514529

ОПТИМАЛЬНОЕ ДИСКРЕТНОЕ ПРЕДСТАВЛЕНИЕ НЕПРЕРЫВНЫХ ИЗМЕРИТЕЛЬНЫХ ПРОЦЕССОВ В АСУ И АСНИ

© 2024 Подлесный В.Ю., Данилов В.В., Третьяков И.А., Колесник Е.В., **Тимченко В.И.**

В настоящей работе, представлены результаты исследования основ информационно-измерительных технологий в АСУ и АСНИ, в частности, дискретной формы представления непрерывных измерительных процессов: дискретное представление непрерывных процессов, погрешность восстановления при дискретном представлении непрерывного процесса, оптимальное дискретное представление, дискретизация полиномами Лежандра.

Ключевые слова: АСУ, АСНИ, ИИС, измерительные сообщения, информационные характеристики, дискретные представления, интерполяция, погрешности, выборки.

Введение. Повсеместный рост применения автоматизированных систем управления (АСУ) и автоматизированных систем научных исследований (АСНИ) в современных производственных процессах невозможен без постоянного качественного улучшения их информационно-измерительных систем (ИИС) и подсистем [1-5]. Наиболее продуктивно совершенствование ИИС не путем улучшения ее элементной базы, устройств, блоков и т.п., а путем внедрения результатов синтеза знаний фундаментальных и прикладных наук, что может привести к внедрению информационно-измерительных технологий. Особую актуальность приобретают информационные технологии при создании АСНИ, где остро стоит вопрос о точности обрабатываемой информации [3, 6].

В работе [7] рассматривалась общая тема исследований, основы информационно-измерительных технологий, в частности, информационные характеристики измерительных сообщений, точнее, модели измерительных сообщений, модели стационарных случайных измерительных процессов, модели нестационарных измерительных процессов, погрешности при измерениях и передаче.

Процесс преобразования сообщений, функций непрерывного времени $\lambda(t)$ в виде функций дискретного времени λ_i , как совокупности координат, представляет собой дискретное представление непрерывных величин, по значениям которых может быть получена оценка исходного непрерывного сообщения $\lambda^*(t)$ [8, 9].

Непрерывное сообщение в дискретной форме. Дискретное представление непрерывного сообщения $\lambda(t)$ на интервале T совокупностью координат сообщения $\lambda_0, \dots, \lambda_{N_k}$ и последующее восстановление по ним исходного сообщения $\lambda^*(t)$ представим в виде:

Исследование проводилось в рамках научной г/б темы «Исследование природы каналов побочных электромагнитных излучений и наводок элементов и устройств офисной вычислительной техники» (регистрационный номер 124012400347-2).

$$(\lambda_0, \lambda_1, \dots, \lambda_{N_k}) = A\lambda(t); \quad (1)$$

$$\lambda^*(t) = B(\lambda_0, \lambda_1, \dots, \lambda_{N_k}). \quad (2)$$

где A - оператор представления; B - оператор восстановления; $\varepsilon_{np} = \lambda_i(t) - \lambda_i^*(t)$ - текущая погрешность дискретного представления. Математические операция (1) реализуется дискретизатором, операция (2), устройством восстановления сообщений источника, когда координаты $i = 0, 1, 2, \dots, N_k$, являются выборками (отсчетами) исходного сообщения. Операторы A и B в общем случае, могут быть как нелинейными, так и линейными, причем с одним и тем же оператором представления могут быть использованы разнообразные операторы восстановления, и наоборот. Если оператор A соответствует представлению выборками, когда $\lambda_i(t) = \lambda(t_i)$, то оператор B может осуществлять аппроксимацию, интерполяцию и т. д. Аппаратная реализация линейных операторов проще, чем нелинейных. При линейном представлении и восстановлении, которыми далее и ограничимся, операторы A и B примут следующий вид:

$$A\lambda(t) = \int_T V_i(t)\lambda(t)dt = \lambda_i; \quad i = 0, 1, 2, \dots, N_k; \quad (3)$$

$$B(\lambda_0, \lambda_1, \dots, \lambda_{N_k}) = \sum_{i=0}^{N_k} \lambda_i W_i(t) = \lambda^*(t), \quad (4)$$

где $V_i(t)$, $W_i(t)$ - соответственно весовые и базисные (координатные) функции, которые должны быть всегда нормированы так, чтобы произведение $[V_i(t), W_i(t)]$ имело размерность, обратную времени t . В зависимости от выбора системы весовых функций $V_i(t)$ в (3) координаты сообщения (рис. 1а) могут представлять собой:

- коэффициенты некоторого ряда λ_i (рис. 1б), при этом выражение (4) определяет обобщенное дискретное представление, в частном случае $V_i(t) \equiv W_i(t)$;

- текущие мгновенные значения (выборки) $\lambda(t_i)$ когда $V_i(t) = \delta(t - t_i)$, где $\delta(t - t_i)$ - дельта-функции Дирака, при этом выражение (3) определяет дискретное представление по выборкам (рис. 1в);

- конечные разности:

$$\Delta^N \lambda(t_i) = \sum_{k=0}^N (-1)^k C_N^k \lambda(t_i, kT_0); \quad (5)$$

когда:

$$V_i(t) = \sum_{k=0}^N (-1)^k C_N^k \delta(t - t_i + kT_0), \quad (6)$$

т. е. $V_i(t)$ является линейной комбинацией функций Дирака. При этом выражение (3) определяет дискретное разностное представление [8]. Представление разностями 1-го порядка ($N=1$), для которого согласно (5), (6) иллюстрируется на рис. 1 г, д:

$$\begin{aligned} \Delta \lambda(t_i) &= \lambda(t_i) - \lambda(t_i - T_0) = \lambda(t_i) - \lambda(t_{i-1}); \\ V_i(t) &= \delta(t - t_i) - \delta(t - t_i + T_0) = \delta(t - t_i) - \delta(t - t_{i-1}), \end{aligned}$$

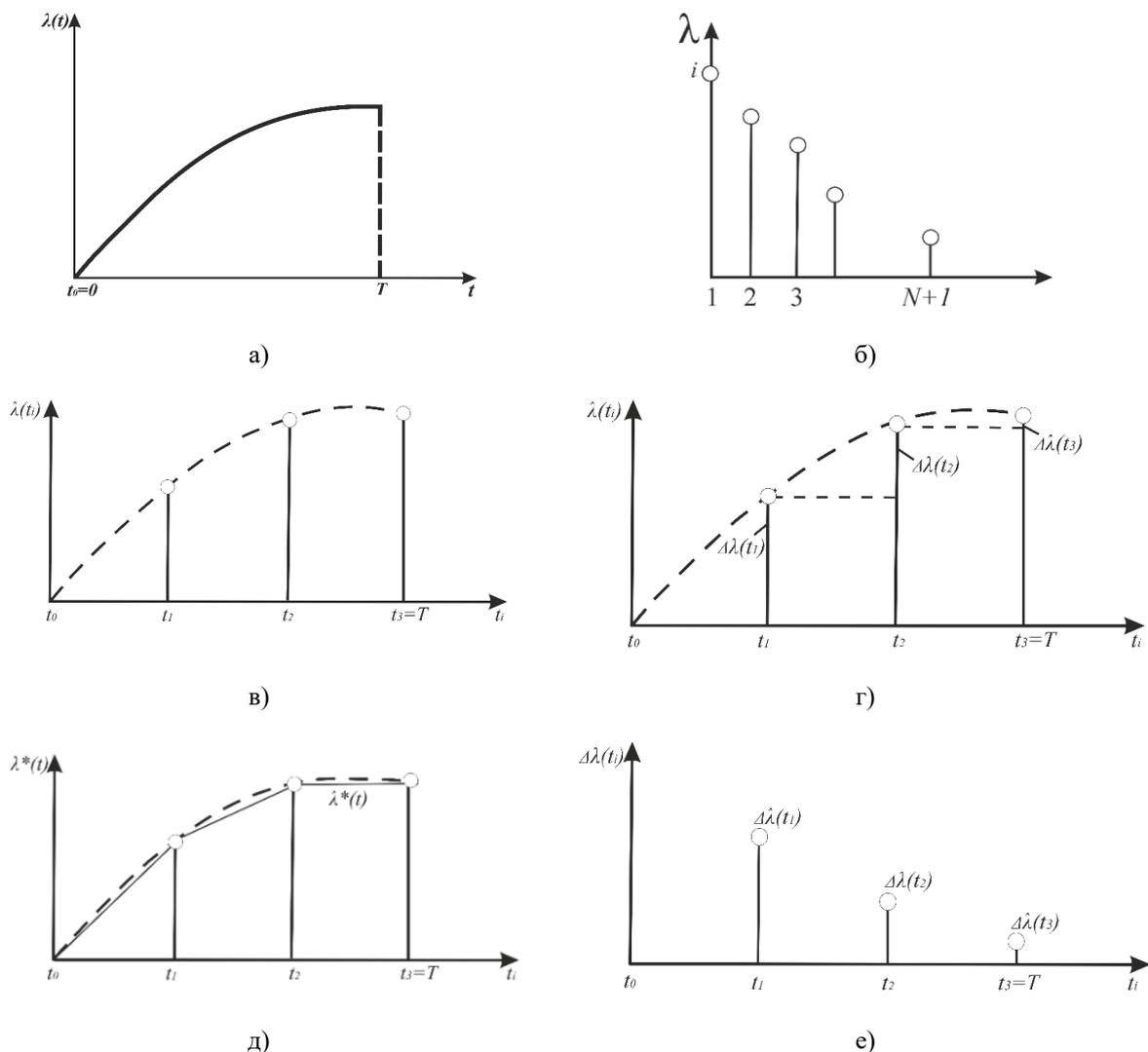


Рис. 1. Дискретные представления непрерывного сообщения:

- а) исходное сообщение на интервале представления T
- б) представление коэффициентами ряда
- в) представление регулярными выборками
- г) формирование конечных разностей первого порядка
- д) представление конечными разностями первого порядка
- е) восстановление исходного сообщения по выборкам путем интерполяции полиномами первого порядка (линейная интерполяция)

Процесс восстановления сообщения согласно (4) осуществляется при всех представлениях в виде полинома $\lambda^*(t)$ степени N_k , который по отношению к исходному сообщению $\lambda_i(t)$ аппроксимирующий (приближающий). В отдельном случае, когда в качестве координат используются выборки, а базисные функции выбраны так, что значения аппроксимирующего полинома совпадают со значениями выборок в моменты их отсчета, этот полином - интерполирующий (рис. 1е). При разностных представлениях восстановление выполняется, в два этапа: вначале по разностям вычисляются значения последовательных выборок, а затем по выборкам строится аппроксимирующий (или чаще интерполирующий) полином. Восстановление (4) можно рассматривать также как линейную фильтрацию с импульсными функциями $W_i(t)$.

Погрешность восстановления при дискретном представлении. В качестве показателя верности восстановления предлагается использовать квадрат средней квадратической погрешности:

$$\begin{aligned} \delta_{\Sigma}^2 &= (1/T) \int_T m[E(t)]^2 dt = (1/T) \int_T m[\Lambda(t) - \Lambda^*(t)]^2 dt = \\ &= (1/T) \int_T m[\Lambda(t) - \sum_{i=0}^{N_k} \Lambda_i^* W_i(t)]^2 dt, \end{aligned} \quad (7)$$

где $\Lambda_i^* = \Lambda_i + Z_i$ - оценка значений координат с учетом воздействия шума $Z(t)$, вызванного квантованием, помехами в канале связи, записью/воспроизведением и др. После возведения в квадрат (7) с учетом линейности операторов, математического ожидания и интерполирования получим:

$$\begin{aligned} \delta_{\Sigma}^2 &= (1/T) \int_T m[\Lambda(t)]^2 dt - (2/T) \int_T \sum_{i=0}^{N_k} m[\Lambda_i^* \Lambda(t)] W_i(t) dt + \\ &+ (1/T) \int_T \sum_{i=0}^{N_k} \sum_{l=0}^{N_k} m[\Lambda_i^* \Lambda_l^*] W_i(t) W_l(t) dt. \end{aligned} \quad (8)$$

Выражение (8) принимает следующий вид, если $\Lambda(t)$ и $Z(t)$ не коррелированы:

$$\delta_{\Sigma}^2 = \delta_{ПП}^2 + \delta_{П}^2, \quad (9)$$

где:

$$\begin{aligned} \delta_{ПП}^2 &= (1/T) \int_T m[\Lambda(t)]^2 dt - (2/T) \int_T \sum_{i=0}^{N_k} m[\Lambda_i \Lambda(t)] W_i(t) dt + \\ &+ (1/T) \int_T \sum_{i=0}^{N_k} \sum_{l=0}^{N_k} m[\Lambda_i \Lambda_l] W_i(t) W_l(t) dt; \end{aligned} \quad (10)$$

$$\delta_{П}^2 = \sum_{i=0}^{N_k} \sum_{k=0}^{N_k} m[Z_i Z_k] (1/T) \int_T W_i(t) W_k(t) dt, \quad (11)$$

$$m[\Lambda_i \Lambda_k] = \int_T \int_T m[\Lambda(t) \Lambda(t')] dt dt'. \quad (12)$$

В уравнениях (9-12), $\delta_{ПП}^2$ - погрешность дискретного представления по не зашумленным координатам; $\delta_{П}^2$ - погрешность, вызванная воздействием внешних помех. Если сообщение $\lambda(t)$ представляет собой реализацию случайного стационарного в широком смысле процесса $\Lambda(t)$ с нулевым математическим ожиданием, то первое слагаемое в (10) равно - δ_{λ}^2 , а выражение (12) принимает вид:

$$m[\Lambda_i \Lambda_k] = \int_T \int_T R_{\lambda}(t-t') V_i(t) V_k(t') dt dt',$$

Функция $Z(t)$ - белый импульсный шум с нулевым математическим ожиданием, т.е.

$$m[Z_i Z_k] = \begin{cases} \sigma_{ш}^2; & i = k; \\ 0; & i \neq l. \end{cases} \quad (13)$$

где $\sigma_{ш}^2$ - дисперсия шумовой погрешности по одной координате, возникающая вследствие действия помех.

Дисперсию шумовой составляющей погрешности на основании (11) и (13) для рассматриваемого случая определим выражением вида:

$$\begin{aligned} \sigma_{ш}^2 &= \sum_{i=0}^{N_k} \sum_{k=0}^{N_k} m[Z_i Z_k] (1/T) \int_T W_i(t) W_k(t) dt = \\ &= (1/T) \sum_{i=0}^{N_k} \sigma_{ш}^2 \|W_i(t)\|^2 = \sigma_{ш}^2 (1/T) \sum_{i=0}^{N_k} \|W_i(t)\|^2 = \gamma_{\phi} \sigma_{ш}^2, \end{aligned} \quad (14)$$

где γ_{ϕ} - коэффициент фильтрующей способности выбранных базисных функций по отношению к внешним помехам типа дискретного белого шума; $\|W_i(t)\|$ - норма базисной функции, которые определим выражениями (15):

$$\begin{cases} \gamma_{\phi} = (1/T) \sum_{i=0}^{N_k} \|W_i(t)\|^2 \\ \|W(t)\| = [\int W_i^2(t) dt]^{0.5} \end{cases}. \quad (15)$$

В частном случае при нормировке вида:

$$(1/T) \|W_i(t)\| = 1, \quad (16)$$

коэффициент фильтрации будет равен $\gamma_{\phi} = N_k + 1$.

Из (14) следует вывод, что тогда $\Delta f_{\text{за}} = 1/(T^2) \sum_{i=1}^{N_k} \|W_i(t)\|^2$ можно рассматривать как эффективную полосу пропускания аппроксимирующего (восстанавливающего) полинома (4); $\gamma_{\phi} = \Delta f_{\text{за}} T$ - коэффициент фильтрации.

Далее, рассмотрим два основных источника помех: шумы квантования с дисперсией $\sigma_{кв}^2$ и шумы в системе передачи с дисперсией $\sigma_{ПЕР}^2$. В предположении некоррелированности этих шумов в цифровых системах передачи $\sigma_{ш}^2 \approx \sigma_{кв}^2 + \sigma_{ПЕР}^2 = \sigma_{кв}^2 + \sigma_{ан}^2$ и на основании (9), (14) получим для стационарных передаваемых сообщений $\lambda(t)$:

$$\delta_{\Sigma}^2 \approx \delta_{ПР}^2 + \gamma_{\phi} (\delta_{кв}^2 + \delta_{ПЕР}^2) = \delta_{ПР}^2 + \delta_{кв\Sigma}^2 + \delta_{ПЕР\Sigma}^2, \quad (17)$$

где $\delta_{\Sigma}^2 = \sigma_{\Sigma}^2 / \sigma_{\lambda}^2$; $\delta_{кв}^2 = \sigma_{кв}^2 / \sigma_{\lambda}^2$, $\delta_{ПЕР\Sigma}^2 = \sigma_{ПЕР}^2 / \sigma_{\lambda}^2$ - дисперсии соответствующих погрешностей, приведенные к дисперсии сообщения σ_{λ}^2 ; $\delta_{кв\Sigma}^2 = \gamma_{\phi} \delta_{кв}^2$, $\delta_{ПЕР\Sigma}^2 = \gamma_{\phi} \delta_{ПЕР}^2$ - суммарные погрешности, вызванные соответственно квантованием и передачей всех координат сообщения на интервале представления T . В аналоговых импульсных системах $\sigma_{ПЕР}^2 = \sigma_{П}^2 + \sigma_{ан}^2$.

Оптимальное дискретное представление. Нахождение оптимального дискретного представления функции $\Lambda(t)$ представляет отыскание системы весовых $V_i(t)$ и базисных $W_i(t)$ функций, которая обеспечивает минимальную погрешность дискретного представления $\delta_{ПР}^2$ при фиксированном числе координат или минимальное число координат при заданной погрешности представления.

Условия, определяющие систему оптимальных весовых и координатных функций, при использовании $(N_k + 1)$ координат, при которых $\delta_{\text{ПР}}^2$ будет минимальной, могут быть найдены из (10). Когда $\lambda(t)$ представляет собой реализацию стационарного случайного процесса с корреляционной функцией $R_\lambda(\tau)$, число членов суммы $(N_k + 1)$ будет минимальным, если [10-13]:

- $V_i(t) = W_i(t)$;

- в качестве координат используются обобщенные коэффициенты Фурье, определяемые в соответствии с (3);

- оптимальные координатные функции $\{W_i(t)\}$ удовлетворяют однородному интегральному уравнению Фредгольма второго рода

$$\sigma_i^2 W_i(t) = (1/T) \int_0^T R_\lambda(t, t') W_i(t') dt', \text{ где } W_i(t) - \text{собственные функции; числа } \sigma_i^2 -$$

собственные значения ядра уравнения, т.е. $R_\lambda(t, t')$. При таком представлении будут выполняться условия ортонормированности базисных функций и некоррелированности координат. Следует отметить, что при ортогональности базисных функций

$$\int_T W_i(t) W_k(t) dt = \begin{cases} 0; & i \neq k \\ \|W_i\|^2; & i = k \end{cases} \text{ выражение для квадрата средней квадратической}$$

погрешности $\delta_{\text{ПР}}^2$, приведенной к дисперсии процесса σ_λ^2 на основании (10) принимает вид:

$$\delta_{\text{ПР}}^2 = \sigma_{\text{ПР}}^2 / (\sigma_\lambda^2) = 1 - \sum_{i=0}^{N_k} \delta_i^2, \tag{18}$$

где
$$\delta_i^2 = \sigma_i^2 / (\sigma_\lambda^2) = \frac{1}{T \sigma_\lambda^2 \|W_i\|^2} \iint_T R_\lambda(t, t') W_i(t) W_i(t') dt dt' \tag{19}$$

приведенная дисперсия i-й координаты. Выражение (19) получено из (12) с учетом того, что мощность i-й координаты $\sigma_i^2 = (1/T) m[\Lambda_i^2]$.

Близкие к оптимальным ортогональные базисные функции для некоторых случайных процессов при различной длительности интервала представления приведены в таблице 1 [10, 13, 14].

Таблица 1. Оптимальные ортогональные базисные функции

Порядок дифференцируемости процесса	$W_i(t)$		
	$T \ll \tau_k$	$T \approx \tau_k$	$T > \tau_k$
0	$0,5; \cos(\pi it)/T$	$0,5; \cos(\pi it)/T$	$\sin(\pi it)/T$ (для убывающего спектра) или $\frac{\sin \Omega_c(t - iT_0)}{\Omega_c(t - iT_0)}$ (для неубывающего спектра)
1	Полиномы Лежандра		
2		Полиномы Лежандра	
∞			

Полином Лежандра и дискретизация. Дискретное представление по ортогональным полиномам Лежандра (дискретизация) выполняется по формуле (3), а восстановление непрерывного сообщения - согласно (4), где в качестве базисных функций $W_i(t)$ используются полиномы Лежандра. Определим их на интервале $[-T/2, T/2]$ и подчиним; условию (16). Тогда $V_i(t) = W_i(t) \|W_i\|^2$. Полученные таким образом полиномы Лежандра степеней $N_k = 0; 1; 2$ имеют следующий вид: $W_0(t) = 1; W_1(t) = 2\sqrt{3}t/T; W_2(t) = (\sqrt{5}/2)[(12t^2)/T^2 - 1]$.

Определим квадрат средней квадратической погрешности (дисперсии) представления для моделей стационарных процессов, ориентируясь на работы [11, 15, 16]. Конечные формулы и промежуточные выкладки существенно упрощаются, если ограничиться получением приближенных выражений, справедливых в области величин погрешности $\delta_{\text{пр}} \leq 0,1$, что соответствует практической потребности. При этом ограничении и условии, что степень аппроксимирующего полинома не выше второй, интервал аппроксимации T значительно меньше интервала корреляции исходного процесса τ_k [12, 17], что позволяет использовать разложение корреляционной функции в ряд Тейлора вблизи нулевого значения аргумента и ограничиться несколькими начальными членами.

Ориентируясь на практическую значимость, проведем моделирование бесконечно дифференцируемого процесса с нормальной корреляционной функцией вида:

$$r_\lambda(\tau) = \sin \Omega_c \tau / (\Omega_c \tau); \quad \sigma_\lambda^2 = 1,$$

при аппроксимации процесса по первым двум полиномам Лежандра. Введем следующие обозначения: $F_0 = 1/T$ - частота регулярных обобщенных координат (частота дискретизации) в предположении, что все $(N_k + 1)$ координаты на интервале T передаются как одна обобщенная; $\nu = F_0 / (1F_c) = \pi / (\Omega_c T)$, где $F_c = \Omega_c / (2\pi)$ - коэффициент избыточности, показывающий, во сколько раз частота дискретизации превышает величину $2F_c$ определенную по теореме Котельникова. Используя далее эти обозначения, запишем $R_\lambda(\tau) = \sigma_\lambda^2 [\sin \pi \tau / (\nu T)] / [\pi \tau / (\nu T)]$. Разлагая корреляционную функцию в ряд Тейлора по степеням аргумента, имеем:

$$R_\lambda(\tau) = \sigma_\lambda^2 [1 - \frac{1}{6} (\frac{\pi \tau}{\nu T})^2 + \frac{1}{120} (\frac{\pi \tau}{\nu T})^4 - \dots].$$

Если поместить начало отсчета времени в середине интервала аппроксимации, то пределы интегрирования в формуле (19) будут $[-T/2, T/2]$. Используя (18), (19), а также найденные выражения для полинома нулевого порядка и его нормы, получим формулу для дисперсии координаты с номером $i = 0$:

$$\sigma_0^2 = 1/T^2 \cdot \int_{-T/2}^{T/2} \int_{-T/2}^{T/2} \{1 - 6^{-1} (\frac{\pi}{\nu T})^2 (t-t')^2 + 120^{-1} (\frac{\pi}{\nu T})^4 (t-t')^4 - \dots\} dt dt'$$

Расчеты показывают, что $\sigma_0^2 = 1 - 36^{-1} (\pi/\nu)^2 + 1800^{-1} (\pi/\nu)^4 - \dots$. При $\pi/\nu \leq 1$, это соответствует оговоренному выше условию малости T по сравнению с интервалом корреляции. Ряд сходится, так как его члены монотонно убывают по модулю. Поэтому, если ограничиться первыми двумя членами, то остаток ряда по модулю не превысит первого отброшенного члена.

Таким образом, с погрешностью вычисления не более 0,1% в качестве расчетной формулы для дисперсии старшей координаты можно принять выражение:

$$\sigma_0^2 = 1 - 36^{-1}(\pi/v)^2. \quad (20)$$

Путем аналогичного подхода, используя полином первого порядка $W_1(t)$, получаем:

$$\sigma_1^2 \approx 1 - 36^{-1}(\pi/v)^2 - 3600^{-1}(\pi/v)^4. \quad (21)$$

Последовательно подставляя (20), затем (21) вместе в формулу (18), найдем приведенные погрешности аппроксимации $\sigma_{ПР}^2$ при использовании соответственно одного ($N_k = 0$) и двух полиномов ($N_k = 1$) Лежандра:

- при $N_k = 0$ $\delta_{ПР}^2 \approx \pi^2 / (36v^2)$;
- при $N_k = 1$ $\delta_{ПР}^2 \approx \pi^4 / (3600v^4)$.

Аналогичным образом по результатам моделирования были получены остальные расчетные формулы для дисперсий коэффициентов и приведенных погрешностей при других видах корреляционных функций сообщений и различных степенях полиномов (таблицы 2 и 3). Погрешности счета по этим приближенным формулам при $v \geq \pi$, не превышают 3% для не дифференцируемого и существенно меньше для дифференцируемых процессов.

Таблица 2. Дисперсия коэффициентов

Номер координаты i	δ_i^2		
	$r_\lambda(\tau) = e^{(-\Omega_c \tau)}$	$(1 + \Omega_c \tau)e^{(-\Omega_c \tau)}$	$\frac{\sin \Omega_c \tau}{\Omega_c \tau}$
0	$1 - \pi / (3v)$	$1 - \pi^2 / (12v^2)$	$1 - \pi^2 / (36v^2)$
1	$\pi / (5v)$	$\pi^2 / (12v^2) - \pi^3 / (105v^3)$	$\frac{\pi^2}{36v^2} - \frac{\pi^4}{3600v^4}$
2	$\pi / (21v)$	$\pi^3 / (126v^3)$	$\frac{\pi^4}{3600v^4} - \frac{\pi^6}{705,6 \cdot 10^3 v^6}$

Таблица 3. Дисперсия погрешности

Метод представления	Степень представления полинома N_k	Коэффициент фильтрации $\Omega_c \tau$	Приведенная дисперсия погрешности дискретного представления $\delta_{ПР}^2$			
			$r_\lambda(\tau) = \exp(-\Omega_c \tau)$	$r_\lambda(\tau) = (1 + \Omega_c \tau) \times \exp(-\Omega_c \tau)$	$r_\lambda(\tau) = e^{\frac{-\Omega_c \tau }{\sqrt{2}}} \times (\cos \frac{\Omega_c \tau }{\sqrt{2}} + \sin \frac{\Omega_c \tau }{\sqrt{2}})$	$r_\lambda(\tau) = \sin \frac{\Omega_c \tau}{\Omega_c \tau}$
Коэффициенты ряда по полиномам Лежандра	0	1	$\pi / (3v)$	$\pi^2 / (12v^2)$	$\pi^2 / (12v^2)$	$\pi^2 / (36v^2)$
	1	2	$2\pi / (15v)$	$\pi^3 / (105v^3)$	$\pi^3 / (105\sqrt{2}v^3)$	$\pi^4 / (3600v^4)$
	2	3	$3\pi / (35v)$	$\pi^3 / (630v^3)$	$\pi^3 / (630\sqrt{2}v^3)$	$\pi^6 / (705,6 \cdot 10^3 v^6)$
	0	1	$\pi / (2v)$	$\pi^2 / (12v^2)$	$\pi^2 / (12v^2)$	$\pi^2 / (36v^2)$
Выборками с интерполяцией полиномом Лагранжа	1	2/3	$\pi / (3v)$	$2\pi^3 / (45v^3)$	$\sqrt{2}\pi^3 / (45v^3)$	$\pi^4 / (600v^4)$
	2	4/5	$\pi / (3v)$	$\pi^3 / (30v^3)$	$\pi^3 / (30\sqrt{2}v^3)$	$\pi^6 / (3308v^6)$

Данные таблицы 2 показывают, что дисперсии координат убывают весьма быстро с ростом номеров координат, причем дисперсия старшей координаты близка к дисперсии исходного процесса. Следовательно, уменьшение погрешности аппроксимации с добавлением координат (увеличением степени полинома) происходит все меньшими порциями.

Результаты моделирования зависимости $v = f(\delta_{пп})$ при использовании полиномов Лежандра приведены в табл. 4 и на рис. 2. Предполагалось, что составляющие суммарной погрешности (17) вносят одинаковый вклад, т.е. $\delta_{пп}^2 = 3^{-1} \delta_{\Sigma}^2$.

Таблица 4. Результаты моделирования зависимости

Метод представления	N_k	Коэффициент избыточности представления v		
		$r_{\lambda}(\tau) = e^{(-\Omega_c \tau)}$	$r_{\lambda}(\tau) = (1 + \Omega_c \tau) \times \exp(-\Omega_c \tau)$	$r_{\lambda}(\tau) = \sin(\Omega_c \tau) / (\Omega_c \tau)$
Коэффициентами ряда по полиномам Лагранжа	0	$\pi / (\delta_{\Sigma}^2)$	$\pi / (2\delta_{\Sigma}^2)$	$\pi / (2\sqrt{3}\delta_{\Sigma}^2)$
	1	$2\pi / (5\delta_{\Sigma}^2)$	$\pi / (35\delta_{\Sigma}^2)^{1/3}$	$\pi / (1200\delta_{\Sigma}^2)^{1/4}$
	2	$9\pi / (35\delta_{\Sigma}^2)$	$\pi / (210\delta_{\Sigma}^2)^{1/3}$	$\pi / (235,2 \cdot 10^3 \delta_{\Sigma}^2)^{1/6}$
Выборками с интерполяцией полиномом Лагранжа	0	$3\pi / (2\delta_{\Sigma}^2)$	$\pi / (2\delta_{\Sigma}^2)$	$\pi / (2\sqrt{3} \cdot \delta_{\Sigma}^2)$
	1	$\pi / (\delta_{\Sigma}^2)$	$[2\pi^3 / (15\delta_{\Sigma}^2)]^{1/3}$	$\pi / (200\delta_{\Sigma}^2)^{1/4}$
	2	$11\pi / (10\delta_{\Sigma}^2)$	$\pi / (10\delta_{\Sigma}^2)^{1/3}$	$[3\pi^6 / (3308\delta_{\Sigma}^2)]^{1/6}$

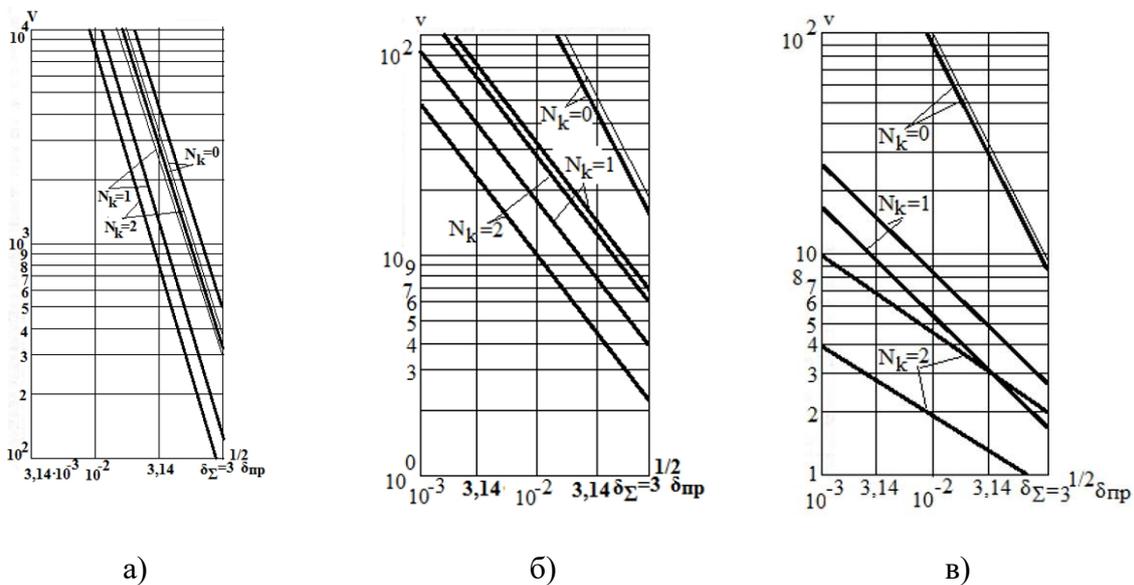


Рис. 2. Сравнение частоты дискретизации при представлениях сообщения коэффициентами ряда с восстановлением на основе полиномов Лежандра и выборками при интерполяции полиномом Лагранжа для случайного процесса с корреляционной функцией (N_k - степень восстанавливающего полинома):

а) $r_{\lambda}(\tau) = \exp(-\Omega_c |\tau|)$ б) $r_{\lambda}(\tau) = e^{\frac{-\Omega_c |\tau|}{\sqrt{2}}} \times (\cos \frac{\Omega_c |\tau|}{\sqrt{2}} + \sin \frac{\Omega_c |\tau|}{\sqrt{2}})$ в) $r_{\lambda}(\tau) = \sin \frac{\Omega_c \tau}{\Omega_c \tau}$

Выводы. Работа посвящена разработке основ информационно измерительных технологий в АСУ и АСНИ, где, показаны результаты исследования дискретной формы представления непрерывных измерительных процессов, а именно:

- дискретное представление непрерывных процессов показало, что при выбранном методе представления задача восстановления сводится к выбору аппроксимирующего или интерполирующего оператора. Обычно при обобщенном дискретном представлении восстановление ведется на основе аппроксимации, а при дискретном представлении выборками и разностных представлениях - путем интерполяции;

- оценка погрешности восстановления при дискретном представлении непрерывного процесса показала целесообразность использования в качестве показателя верности восстановления применение квадрата средней квадратической погрешности;

- оптимальное дискретное представление случайных процессов представляет собой сложную техническую задачу, поэтому для практических приложений желательно иметь универсальные координатные функции, применение которых возможно при не очень сложной аппаратуре и в то же время обеспечивает близость предоставления к оптимальному.

СПИСОК ИСТОЧНИКОВ

1. Кузнецов, С. В. Модель единой централизованной автоматизированной системы управления техническим состоянием вооружения, военной и специальной техники / С. В. Кузнецов, А. М. Винограденко // Научные технологии в космических исследованиях Земли. – 2018. – Т. 10, № 4. – С. 48-54. – DOI 10.24411/2409-5419-2018-10096. – EDN YATSQX.
2. Полтавский, А. В. Оптимизация информационно-измерительной системы беспилотного воздушного судна / А. В. Полтавский, А. А. Тюгашев, Н. К. Юрков // Надежность и качество сложных систем. – 2021. – № 4(36). – С. 44-55. – DOI 10.21685/2307-4205-2021-4-6. – EDN UOUUMU.
3. Полтавский, А. В. Модель управления состоянием оптико-электронного тракта информационно-измерительных и управляющих систем / А. В. Полтавский, Н. А. Кузин, Н. К. Юрков // Измерение. Мониторинг. Управление. Контроль. – 2023. – № 3(45). – С. 5-15. – DOI 10.21685/2307-5538-2023-3-1. – EDN ZBTOQU.
4. Третьяков, И. А. Применение оптических методов обработки информации в АСНИ радиосигналов / И. А. Третьяков, В. В. Данилов, В. А. Ступак // Вестник Донецкого национального университета. Серия Г: Технические науки. – 2024. – № 1. – С. 42-49. – DOI 10.5281/zenodo.10974622. – EDN LSKSPA.
5. Синтез методов цифровой регистрации в системах сбора и обработки измерительной информации для обеспечения достоверности в информационно-управляющих системах / Ж. У. Севинов, Б. М. Темербекова, У. Б. Мамазаров, Б. М. Бежимбетов // Потомки Аль-Фаргани. – 2024. – №4. – С. 91-96. – DOI 10.5281/zenodo.14555403.
6. Третьяков, И. А. Вариации моделей измерительных процессов в АСУ и АСНИ / И. А. Третьяков, В. В. Данилов // Сборник статей VII Междунар. науч.-техн. конф. «Минские научные чтения-2024» в 3 т. (Минск, 03–05 декабря 2024 г.). – Т. 1. – Минск: БГТУ, 2024. – С. 319-323. – EDN KWWZPG.
7. Основы информационно-измерительных технологий в АСУ и АСНИ / В. Ю. Подлесный, В. В. Данилов, И. А. Третьяков, Е. В. Колесник // Вестник Донецкого национального университета. Серия Г: Технические науки. – 2024. – № 3. – С. 64-74. – DOI 10.5281/zenodo.14018604. – EDN SDSQAP.
8. Цифровые информационно-измерительные системы: Теория и практика / А. Ф. Фомин, О. Н. Новоселов, К. А. Победоносцев [и др.]. – М.: Энергоатомиздат, 1996. – 445 с.
9. Раннев, Г. Г. Измерительные информационные системы / Г. Г. Раннев. – М.: Академия, 2010. – 329 с.
10. Сурогина, В. А. Информационно-измерительная техника и электроника / В. А. Сурогина, В. И. Калашников, Г. Г. Раннев. – М.: Высшая школа, 2006. – 512 с.
11. Цапенко, М. П. Измерительные информационные системы / М. П. Цапенко. – М.: Энергоатомиздат, 1985. – 440 с.
12. Гаранин, М. В. Системы и сети передачи информации / М. В. Гаранин, В. И. Журавлев, С. В. Кунегин – М.: Радио и связь, 2001. – 336 с.
13. Прохоров, С. А. Математическое описание и моделирование случайных процессов / С. А. Прохоров. – Самара: Изд-во СамГАУ, 2001. – 209 с.
14. Ольховский, Ю. Б. Сжатие данных при телеизмерениях / Ю. Б. Ольховский, О. Н. Новоселов, А. П. Мановцев. – М.: Сов. радио, 1971. – 303 с.

15. Андриевский, Б. Р. Элементы математического моделирования в программных средах МАТЛАБ и Scilab / Б. Р. Андриевский, А. Л. Фрадков. – СПб.: Наука, 2001. – 286 с.
16. Чернявский, Е. А. Анализ информационных процессов измерительно-вычислительных средств. / Е.А. Чернявский. – СПб.: Энергоатомиздат, 1999. – 191 с.
17. Волков В.Л. Проектирование цифровых алгоритмов информационно-измерительных систем / В. Л. Волков. – Москва: Изд-во МАИ, 1991. – 33 с.

Поступила в редакцию 03.09.2024 г., рекомендована к печати 20.09.2024 г.

OPTIMAL DISCRETE REPRESENTATION OF CONTINUOUS MEASUREMENT PROCESSES IN ACS AND ASRS

*Podlesnyi V.IU., Danilov V.V., Tretiakov I.A., Kolesnik E.V., **Тимченко В.И.***

This paper presents the results of a study of the fundamentals of information and measurement technologies, in particular, the discrete form of representation of continuous measurement processes: discrete representation of continuous processes, the error of reconstruction in the discrete representation of a continuous process, optimal discrete representation, discretization by Legendre polynomials.

Keywords: ACS, ASRS, IMS, measurement messages, information characteristics, discrete representations, interpolation, errors, samples.

Подлесный Владимир Юрьевич

аспирант кафедры радиофизики и инфокоммуникационных технологий ФГБОУ ВО «Донецкий государственный университет», Российская Федерация, ДНР, г. Донецк.
E-mail: pikaayzek@yandex.ru

Podlesnyi Vladimir IUrevich

Postgraduate Student at Department of Radiophysics and Infocommunication Technologies of Donetsk State University, Russian Federation, DPR, Donetsk.

Данилов Владимир Васильевич

доктор технических наук, профессор, заведующий кафедрой радиофизики и инфокоммуникационных технологий ФГБОУ ВО «Донецкий государственный университет», Российская Федерация, ДНР, г. Донецк.
E-mail: ut5iv@mail.ru

Danilov Vladimir Vasilevich

Doctor of Technical Sciences, Full Professor, Head of the Department of Radiophysics and Infocommunication Technologies of Donetsk State University, Russian Federation, DPR, Donetsk.

Третьяков Игорь Александрович

кандидат технических наук, доцент, заместитель декана по научной работе, доцент кафедры радиофизики и инфокоммуникационных технологий ФГБОУ ВО «Донецкий государственный университет», Российская Федерация, ДНР, г. Донецк.
E-mail: i.tretiakov@mail.ru

Tretiakov Igor Aleksandrovich

Candidate of Technical Sciences, Associate Professor, Deputy Dean for Research, Associate Professor at Department of Radiophysics and Infocommunication Technologies of Donetsk State University, Russian Federation, DPR, Donetsk.

Колесник Екатерина Витальевна

аспирант кафедры радиофизики и инфокоммуникационных технологий ФГБОУ ВО «Донецкий государственный университет», Российская Федерация, ДНР, г. Донецк.
E-mail: ketrinkilli@ya.ru

Kolesnik Ekaterina Vitalevna

Postgraduate Student at Department of Radiophysics and Infocommunication Technologies of Donetsk State University, Russian Federation, DPR, Donetsk.

Тимченко Владимир Иванович

кандидат физико-математических наук, доцент, доцент кафедры радиофизики и инфокоммуникационных технологий ФГБОУ ВО «Донецкий государственный университет», Российская Федерация, ДНР, г. Донецк.

Timchenko Vladimir Ivanovich

Candidate of Physico-Mathematical Sciences, Associate Professor, Associate Professor at Department of Radiophysics and Infocommunication Technologies of Donetsk State University, Russian Federation, DPR, Donetsk.

ЦИФРОВЫЕ ПЛАТФОРМЫ ДЛЯ ОНЛАЙН-КОНСУЛЬТИРОВАНИЯ: СРАВНИТЕЛЬНЫЙ АНАЛИЗ И ПЕРСПЕКТИВЫ

© 2024 *Нижникова О.Л., Дорохина Г.В., Большакова С.А.*

В статье рассматривается онлайн-консультирование как эффективный инструмент для получения экспертных рекомендаций через цифровые платформы. Акцент сделан на структуру таких платформ и их функциональные возможности, включая возможность общения с консультантом в реальном времени. Это позволяет повысить качество обслуживания клиентов и улучшить взаимодействие с ними. Анализ различных видов платформ помогает выявить лучшие практики для оптимизации процессов консультаций и совершенствования сервиса, что способствует более эффективному решению проблем пользователей и организаций.

Ключевые слова: онлайн-консультирование, вопросно-ответный механизм, веб-чат, консалтинг.

Введение. Онлайн-консультирование – это виртуальная версия традиционного консультирования, реализуемая с помощью цифровых платформ. Она является более гибкой и доступной, может выполняться как в реальном времени, так и отсрочено, снимает географические и временные рамки, может сохранять историю обращений [1]. Такое консультирование допускает разные формы общения: индивидуальное и групповое; в виде текстового чата, звонка, видео-связи, видеоконференции, форума, электронного письма, заполнения электронных форм, обращения на «горячую линию». Цифровые платформы онлайн-консультирования предполагают как консультирование человеком-оператором, так и автоматизацию этого процесса посредством компьютерных программ и технологий искусственного интеллекта [2-4].

В последние годы, в условиях стремительной цифровизации и технологических изменений, онлайн-консультирование становится все более популярным, что подчеркивает его важность для различных областей, включая производственные процессы [5], бизнес [6], науку, образование [7], здравоохранение [8, 9], исполнительная система [10], государственные услуги [11–13]. В связи с этим возникает необходимость в анализе функциональных возможностей таких платформ для последующего применения в государственном и муниципальном секторе, в сфере образования и науки.

Цифровые технологии уже сейчас облегчают взаимодействие граждан с государственными органами и обеспечивают более быстрый и удобный доступ к различным услугам [13]. Например, на портале «Госуслуги» подключение системы онлайн-консультирования с людьми-операторами одновременно с виртуальным ассистентом позволило бы расширить круг задач, с которыми он справляется. Внедрение таких систем было бы полезно для взаимодействия с подведомственными организациями, с контролирующими органами. Автоматизированное онлайн-консультирование позволило бы сократить издержки на взаимодействие с гражданами, увеличить их вовлеченность и качество жизни, что обуславливает актуальность исследования.

Исследование проводилось в рамках научной г/б темы «Исследование и разработка методов обработки данных и естественно-языковых текстов с применением онтологий (FREN-2023-0007)» (№ Г/Р 1230926000030-4).

Постановка задачи. Целью данного исследования является выявление ключевых характеристик платформ онлайн-консультирования для последующего внедрения в образовательной и научной, государственной и муниципальной сферах.

Для достижения цели в работе решены следующие задачи.

1. Проанализировать преимущества, ограничения и вызовы, связанные с онлайн-консультированием, инфраструктурные особенности, влияющие на его распространение.

2. Выполнить анализ структуры и функциональных возможностей современных платформ онлайн-консультирования.

Преимущества, ограничения и вызовы онлайн-консультирования, особенности современной цифровой инфраструктуры. Преимущества онлайн-консультирования:

– доступность и удобство: возможность получить консультацию в любое время, независимо от местоположения;

– снижение затрат: уменьшаются расходы на командировки, аренду помещений и другие организационные моменты;

– экономия времени: отсутствие необходимости в личной встрече позволяет ускорить процессы и улучшить доступность консультаций.

Ограничения и вызовы онлайн-консультирования:

– проблемы с технологической инфраструктурой: в некоторых регионах или странах проблемы с интернет-соединением могут ограничивать доступ к онлайн-консультациям;

– безопасность и конфиденциальность: вопросы защиты данных становятся особенно важными, когда речь идет о личных данных граждан или конфиденциальной научной информации;

– недостаток личного общения: в некоторых ситуациях личная встреча может быть предпочтительнее для более глубокого обсуждения сложных или деликатных вопросов.

В целом, онлайн-консультирование в науке и государственной структуре продолжает развиваться и становится важной частью современной инфраструктуры обслуживания и сотрудничества, несмотря на существующие вызовы.

Для увеличения использования онлайн-консультирования можно принять ряд стратегических и организационных мер. Важно учитывать как технические, так и организационные аспекты, а также создавать удобные условия для всех участников. Вот несколько подходов:

1. Развитие инфраструктуры и технологий. Улучшение доступности интернета и технологий: важно инвестировать в улучшение интернет-соединения, особенно в удаленных и сельских регионах. Для государственной структуры это может означать расширение государственных цифровых услуг через улучшение интернет-платформ, которые обеспечивают доступность онлайн-консультирования для всех слоев населения.

2. Интеграция современных платформ: Использование удобных и безопасных онлайн-платформ для видеозвонков, чатов, форумов и других форм взаимодействия. Примером могут быть платформы с поддержкой искусственного интеллекта (чат-боты), которые могут обеспечить базовое консультирование и перенаправление к живому консультанту при необходимости.

3. Разработка мобильных приложений: в условиях высокомобильного населения и повсеместного использования смартфонов, мобильные приложения для консультаций могут стать важным элементом. Это позволит получить помощь «на ходу», не ограничиваясь компьютерами.

Приведем примеры развития инфраструктуры и технологий, улучшения доступности интернета и расширения цифровых услуг в Российской Федерации.

1. Проект "Цифровая экономика": В рамках программы «Цифровая экономика Российской Федерации» активно инвестируются средства в улучшение интернет-инфраструктуры, в том числе для улучшения доступа к интернету в удаленных и сельских районах. Реализуются проекты по расширению покрытия сети 4G и 5G, а также созданию высокоскоростного интернета для пользователей в отдаленных и труднодоступных регионах.

2. Проект "Устранение цифрового неравенства": В России проводятся работы по устранению цифрового неравенства, которое включает в себя улучшение качества связи в сельской местности. Это позволяет улучшить доступ к интернет-услугам для жителей удаленных и сельских районов, что, в свою очередь, открывает возможности для получения государственных услуг через интернет.

3. Развитие "умных" сельских территорий: В рамках реализации программы по развитию «умных» сельских территорий внедряются технологии для улучшения качества жизни в удаленных регионах. Это включает в себя установку цифровых инфраструктур и повышение доступности онлайн-услуг, таких как телемедицина, дистанционное обучение и онлайн-обслуживание государственных органов.

4. Телемедицина: Для улучшения доступа к медицинским услугам в удаленных районах развиваются платформы телемедицины, которые позволяют пациентам в сельской местности получать консультации врачей через интернет. Это решает проблему дефицита медицинских кадров в удаленных уголках страны и обеспечивает доступ к медицинской помощи без необходимости долгих поездок в город.

5. Цифровизация образования: В России активно развиваются онлайн-образовательные платформы, такие как "Российская электронная школа", которые позволяют студентам и школьникам в удаленных районах получать образование с использованием интернета. Это помогает обеспечить равный доступ к качественному образованию для всех слоев населения.

6. Государственные цифровые услуги в удаленных регионах: В рамках повышения доступности государственных услуг граждане могут обращаться в различные государственные органы через интернет, даже если они находятся в самых отдаленных уголках страны. Такие платформы, как портал Госуслуг, позволяют получать услуги, такие как оформление документов, подача заявлений и консультирование, без необходимости личного визита.

Эти примеры показывают, как развитие инфраструктуры и технологий способствует улучшению доступности государственных и социальных услуг, особенно в сельских и удаленных районах России.

Рассмотрим более подробно возможные подходы упрощения и оптимизации процессов онлайн-консультирования.

1. Автоматизация базовых консультаций: Внедрение технологий искусственного интеллекта и чат-ботов, которые могут предоставлять базовую информацию или отвечать на часто задаваемые вопросы. Это значительно снизит нагрузку на специалистов, позволяя им сосредоточиться на более сложных запросах.

2. Оптимизация рабочих процессов: Внедрение простых, понятных и эффективных интерфейсов для пользователей, которые легко интегрируются в повседневную работу государственных служащих и научных работников. Для этого можно проводить регулярные тестирования и улучшение пользовательского опыта.

Платформы онлайн-консультирования в настоящее время используются преимущественно коммерческими и в меньшей степени некоммерческими организациями для оказания помощи пользователям в поиске информации в текстовом или голосовом диалоговом режиме. Учитывая, что информирование населения по административным вопросам имеют схожие цели и задачи, можно предположить, что этот инструмент будет полезен для решения следующих задач:

- предоставление информации гражданам о государственных услугах и правилах их получения;
- обработка запросов граждан об услугах и инструкциях, предоставляемых государственными органами;
- случаи оперативного реагирования на обращения или жалобы.

Онлайн-консультирование также активно используется в контексте подачи заявок на гранты и финансирование, что способствует повышению качества научных работ и эффективному использованию финансовых ресурсов. Технологические достижения позволяют улучшить доступ к образовательным ресурсам и тренингам для научных сотрудников и студентов, что, в свою очередь, способствует развитию научной квалификации и поддержанию профессионального уровня исследовательских групп.

В сфере государственного управления онлайн-консультирование могло бы улучшить доступ граждан к различным государственным услугам, включая консультации по вопросам законодательства, налогообложения, здравоохранения и социальной защиты. Это способствует оптимизации работы государственных структур, снижению административных барьеров и повышению удовлетворенности граждан качеством предоставляемых услуг.

Несмотря на явные преимущества и растущий интерес к онлайн-консультированию, существуют факторы, которые могут препятствовать или замедлять его широкое внедрение. Вот основные проблемы, которые мешают активному использованию предложенных подходов:

1. Технические проблемы и инфраструктурные ограничения

– неравномерный доступ к интернету: в ряде регионов, особенно в сельской местности или удаленных областях, доступ к стабильному и высокоскоростному интернет-соединению ограничен. Это снижает возможность использования онлайн-консультирования для значительной части населения, что особенно важно для государственных услуг;

– старые или неэффективные технологии: в некоторых государственных, муниципальных учреждениях и научных организациях используются устаревшие технологические решения, которые не поддерживают современные платформы для онлайн-консультирования или не интегрированы с другими системами. Это затрудняет внедрение новых технологий;

– сложность интеграции с существующими системами: В государственном и научном секторах часто используются различные информационные системы и базы данных. Интеграция новых онлайн-инструментов с существующими структурами может быть технически сложной и затратной.

2. Проблемы с безопасностью и конфиденциальностью

– угрозы кибербезопасности: использование онлайн-консультирования всегда сопряжено с рисками утечек данных, хакерских атак и несанкционированного доступа. особенно это касается вопросов защиты персональных данных граждан или научной информации, которая может быть чувствительной;

– отсутствие стандартизированных мер безопасности: в некоторых странах или учреждениях отсутствуют четкие стандарты для обеспечения безопасности данных в рамках онлайн-консультирования. Это может вызывать опасения у пользователей и государственных органов по поводу защиты конфиденциальной информации;

– низкий уровень доверия к цифровым сервисам: В некоторых странах или среди отдельных групп граждан существует недоверие к государственным онлайн-сервисам из-за опасений по поводу утечек данных или некорректного использования личной информации.

Классификация, анализ структуры и функциональных возможностей современных платформ онлайн-консультирования. Классифицировать платформы онлайн-консультирования можно по нескольким основным критериям.

Выделяют активные и неактивные виды чата. Активный: чат открывается автоматически после определенных действий посетителя. Это могут быть определенные действия на данном сайте или же нахождение на странице какое-то количество времени. Это зависит от настроек самого чата. Неактивный: чат открывает сам пользователь при необходимости, нажав на иконку онлайн-консультанта. Например, при возникновении проблем или вопросов.

Также существует разделение по способам связи с пользователем: текстовые сообщения, голосовые звонки и видеосвязь.

Текстовые сообщения и голосовые звонки (с помощью IVR и голосовых ботов) позволяют достигать более высоких уровней автоматизации, чем видеосвязь:

– текстовые чаты (через чат-ботов) предлагают наиболее высокий уровень автоматизации, так как они могут работать с запросами в формате естественного языка, предоставляя ответы и выполняя действия без участия человека;

– голосовые боты и IVR (Interactive Voice Response) в голосовых звонках также могут автоматизировать многие операции, такие как сбор информации и направление запросов, но они ограничены в распознавании сложных запросов и часто требуют человеческого вмешательства для более сложных задач.

По уровню анонимности разделяют на общие: к беседам допускаются все зарегистрированные пользователи, анонимные: общение без регистрации, приватные: ограниченное число участников, задается в настройках.

Разделение онлайн-консультирования по уровню анонимности (общие, анонимные, приватные) важно, поскольку каждый тип общения подходит для разных целей и ситуаций. Выбор между этими форматами зависит от характера консультации, уровня конфиденциальности, количества участников и других факторов. Рассмотрим, когда и в каких случаях лучше использовать каждый из этих типов:

– общие беседы лучше всего использовать для массовых консультаций и публичных обсуждений, когда конфиденциальность не является важным фактором и необходимо донести информацию до широкой аудитории;

– анонимные беседы идеально подходят для ситуаций, где пользователи нуждаются в полной конфиденциальности и не хотят раскрывать свою личность (например, психологические консультации, кризисные ситуации);

– приватные беседы подходят для индивидуальных консультаций или небольших групп, где важно обеспечить высокий уровень конфиденциальности и контроля за доступом (например, юридические, медицинские консультации, закрытые группы поддержки).

В зависимости от типа хранения данных выделяют с хранением информации в облаке или на своем сервере.

Когда выбрать облачное хранение:

- когда необходимо быстрое развертывание платформы;
- если бизнес или проект малый или средний, и нет потребности в значительных инвестициях в физическую инфраструктуру;
- когда требуются высокие масштабы или гибкость (например, быстро растущий трафик);
- когда важна глобальная доступность и отсутствие привязки к физическим серверам;
- если организация не имеет возможности или желания заниматься обслуживанием серверного оборудования.

Когда выбрать собственный сервер:

- когда требуется полный контроль над данными и системами, особенно если данные чувствительны (например, персональные, медицинские или юридические данные);
- когда необходимо соответствовать строгим юридическим требованиям или локализации данных;
- когда организация обладает ресурсами для поддержания собственной инфраструктуры и готова инвестировать в ее обслуживание;
- когда важно избежать зависимости от сторонних провайдеров.

Исходя из вышесказанного:

Облачные решения — идеальны для стартапов, малого и среднего бизнеса, проектов, которым нужно быстро масштабироваться и не иметь проблем с техническим обслуживанием.

Собственные серверы — лучше подходят для организаций с высокими требованиями к безопасности и контролю за данными, а также для тех, кто работает с конфиденциальной информацией и нуждается в полной кастомизации и независимости от внешних провайдеров.

Стоимость является немаловажным критерием, так как платные версии предоставляют доступ к расширенным функциям, а бесплатные версии ограничены в количестве операторов и функциональности. Также стоит разделить по методу вопросно-ответного механизма, который использует операторов, заранее подготовленные сценарии диалогов, искусственный интеллект или комбинацию этих подходов.

Специализированные платформы онлайн-консультирования могут включать различных типов пользователей, в зависимости от того, кто и с какой целью использует ее. Вот основные категории пользователей:

1. Клиенты – это те, кто обращаются за консультацией, обычно это люди, которые ищут помощь или совет по различным вопросам: психологическим, юридическим, медицинским и т. д. Целью обращения является получение консультаций, решение личных или профессиональных проблем, улучшение качества жизни, поиск поддержки в кризисных ситуациях;

2. Консультанты/Эксперты – специалисты, предоставляющие консультации или услуги на платформе, они могут быть психологами, юристами, врачами, тренерами, педагогами и другими профессионалами в зависимости от направленности платформы. Целью их взаимодействия является обеспечение профессиональных консультаций, помощь в решении проблем клиентов, предоставление экспертных знаний.

3. Администраторы – лица, которые управляют платформой и следят за её работой, это могут быть как люди, отвечающие за техническую поддержку, так и те, кто занимается организационными вопросами. Их деятельность связана с управлением доступа пользователей, координацией работы консультантов, обеспечением функциональности и безопасности платформы. Например, системные администраторы, модераторы, менеджеры поддержки.

4. Разработчики и технические специалисты – люди, занимающиеся созданием и поддержкой самой платформы для онлайн-консультирования, это могут быть программисты, инженеры, специалисты по ux/ui-дизайну и другие. Целью их деятельности является разработка новых функций, улучшение интерфейса, решение технических проблем. К ним относятся разработчики веб-сайтов и мобильных приложений, тестировщики, специалисты по безопасности.

5. Платформенные партнеры – это организации или учреждения, которые предоставляют консультантов или поддерживают платформу, это могут быть больницы, школы, юридические фирмы, образовательные центры и т. д. Их целью является предоставление консультационных услуг своим клиентам или сотрудникам, улучшение доступности услуг. К ним относятся медицинские учреждения, юридические компании, учебные заведения.

В структуру платформ онлайн-консультирования входит три основных блока: веб-чат окно, для взаимодействия с посетителем; панель оператора, включающая базовый функционал для работы специалиста; серверная часть, которая отвечает за обработку запросов для создания и отображения списка посетителей и корректного функционирования автоматических приглашений [14, 15].

Структура окна веб-чата может содержать в себе:

- приветственное сообщение, приглашающее пользователя в чат и предоставляющее информацию о компании и услугах;
- возможность перехода из окна чата в мессенджеры или социальные сети для продолжения общения;
- возможность использования видео- и аудиочата для демонстрации преимуществ продуктов или услуг.

Панель оператора имеет следующие функции:

- омниканальность, объединение обращений из разных каналов связи в едином окне для целостной картины пути клиента;
- официальные интеграции с сервисами, законное подключение каналов связи
- интеграция с чат-ботами и CRM (Customer Relationship Management – это система управления взаимоотношениями с клиентами, предназначенная для учёта всех взаимодействий компании с клиентами и хранения важной информации о них).
- эргономичность, сокращение времени оператора и удобные подсказки для ответов;
- измерять ключевые показатели эффективности (kpi), использовать инструменты самоконтроля и анализировать результаты работы для оценки результатов.
- инструменты автоматизации, быстрая идентификация клиентов, история обращений и учёт их потребностей;
- единая статистика компании, инструменты аналитики и статистики для руководителей.

Серверная часть должна обеспечивать:

- безопасность передачи данных, контроль целостности информации на разных уровнях и защита данных;
- развёртывание на сторонних серверах, хранение данных внутри компании и контроль потоков данных.

Зависимость между различными критериями и использованием платформ в настоящее время можно рассмотреть через несколько аспектов, таких как функциональность, доступность, стоимость и потребности пользователей. Каждый из этих факторов влияет на принятие решения о выборе той или иной платформы. Рассмотрим, как различные параметры могут влиять на использование платформ [16]:

1. Бесплатный тестовый период, возможность попробовать услугу перед покупкой. Наличие бесплатного тестового периода или демонстрационной версии часто стимулирует пользователей попробовать платформу перед покупкой. Это позволяет снизить риски, оценить функциональность и принять более обоснованное решение. В настоящее время многие пользователи обращают внимание на возможность опробовать сервис без финансовых обязательств.

2. Сколько консультантов может поддерживать чат (в зависимости от стоимости). Эта характеристика напрямую зависит от бюджета компании. Чем больше количество операторов, тем более масштабной должна быть система и тем выше стоимость. Платформы с возможностью гибкой настройки количества операторов по мере роста бизнеса будут пользоваться популярностью среди небольших и средних компаний. Для крупных организаций, как правило, важна возможность поддерживать большое количество операторов.

В таблице 1 для 9 популярных платформ онлайн-консультирования проанализированы следующие критерии стоимости и условий использования: условия использования пробной версии, основание увеличения цены, тариф в месяц. Эти платформы являются платными и имеют возможность оплаты в РФ.

Таблица 1. Критерии стоимости и условий использования

№	Название	Условия использования пробной версии	Основание увеличения цены	Тариф (в месяц)
1	Chatra	10 дней	Кол-во операторов	Бесплатный \ Базовый 17 € \ Профессиональный 23 €
2	JivoSite	14 дней	Кол-во операторов	Бесплатный \ Базовый 742 Р \ Профессиональный 1 117 Р \ Корпоративный 2 617 Р
3	CarrotQuest	7 дней	Кол-во уникальных посетителей	Общий 3 192 Р \ Поддержка 5 592 Р \ Маркетинг 7 992 Р
4	Talk-me	демо-версия	Кол-во операторов или кол-во обращений	Бесплатный \ Старт 99 Р \ Про 5 280 Р
5	EnvyBox	7 дней	Кол-во операторов	Бесплатный \ 849 Р
6	Venyoo	14 дней	Кол-во заявок	по запросу минимум 1 205 Р
7	Comagic	демо-версия	Сквозная аналитика	по запросу
8	RedHelper	5 дней	Кол-во операторов	по запросу
9	Mango-chat	нет	Кол-во операторов	Бесплатный \ 588 Р

3. Поддержка чат-ботов или систем автоответов. В последние годы использование чат-ботов и автоматических систем ответов становится все более популярным. Они позволяют снизить нагрузку на операторов и ускорить обработку запросов. Платформы, которые предлагают встроенные боты или интеграции с AI-чатами, чаще выбираются компаниями, стремящимися к автоматизации и повышению эффективности работы.

4. Интеграция с мессенджерами и социальными сетями. С увеличением числа платформ для общения с клиентами (например, WhatsApp, Telegram, Facebook Messenger и др.) компании ищут решения, которые поддерживают несколько каналов связи. Возможность интеграции с мессенджерами и социальными сетями становится важным фактором при выборе платформы, особенно для компаний с разветвленной клиентской базой.

5. Наличие триггеров и их настройка. Триггерные системы позволяют автоматизировать взаимодействие с клиентами в зависимости от их действий (например, напоминания или предложения после определенных событий). Это повышает персонализацию общения и улучшает качество обслуживания. Платформы с развитой системой триггеров и их настройки будут востребованы среди тех, кто стремится к оптимизации процессов.

6. Интеграции с другими сервисами. Современные платформы требуют интеграции с другими инструментами, такими как CRM-системы, аналитика, почтовые сервисы и маркетинговые инструменты. Интеграция с внешними сервисами становится важным требованием для более сложных и масштабных организаций, стремящихся создать единую экосистему для управления взаимодействием с клиентами.

7. Наличие мониторинга статистики. Возможность мониторинга статистики и аналитики является важным инструментом для оценки эффективности работы операторов и чат-ботов, а также для оптимизации процессов. Платформы с развитыми аналитическими возможностями востребованы в компаниях, ориентированных на улучшение качества обслуживания и повышение производительности.

8. Возможность оценки диалогов или операторов. Оценка качества общения операторов с клиентами помогает выявить слабые места в обслуживании и повысить его качество. Платформы с возможностью оценивать диалоги и работу операторов важны для крупных компаний, которые стремятся к постоянному улучшению сервиса.

В таблице 2 приведено сравнение функционала платформ онлайн-консультирования по следующим отобраным критериям: поддержка чат-ботов или систем автоответов; интеграция с мессенджерами и соц. сети; наличие триггеров и их настройка; единая статистика; интеграция с другими сервисами (CRM); наличие мониторинга статистики; возможность оценки диалогов или операторов.

9. Наличие функции «обратный звонок». Функция обратного звонка может значительно улучшить взаимодействие с клиентами, предоставляя им возможность получить консультацию через голосовую связь. Это особенно важно для более сложных запросов, которые не всегда можно решить через чат. Наличие этой функции может стать важным аргументом при выборе платформы для компаний, ориентированных на предоставление качественного клиентского сервиса.

10. Другие преимущества панели оператора. Современные панели оператора могут включать множество дополнительных возможностей, таких как автоматическое распределение запросов, поддержка мультимедиа, CRM-интеграции, а также другие инструменты для упрощения работы операторов. Платформы с удобными и функциональными панелями оператора будут предпочтительны для компаний, стремящихся к оптимизации работы сотрудников.

В таблице 3 представлены преимущества панели оператора платформ онлайн-консультирования, в том числе функция «обратный звонок».

Таблица 2. Сравнительная таблица функционала платформ онлайн-консультирования

№	Название	Отобранные критерии платформ онлайн консультирования						
		Поддержка чат-ботов или систем автоответов	Интеграция с мессенджерами и соц. сетями	Наличие триггеров и их настройка	Единая статистика	Интеграция с другими сервисами (CRM)	Мониторинг статистики	Оценка диалогов или операторов
1	Chatra	+	+	+	+	+	+	+
2	JivoSite	+	+	+	-	+	+	+
3	CarrotQuest	+	+	-	+	+	+	
4	Talk-me	+	+	+	+	+	+	+
5	EnvyBox	+	+	+	-	+	+	+
6	Venyoo	+	+	-	-	+	+	-
7	Comagic	заготовленные формы	-	-	+	+	+	+
8	RedHelper	+	+	+	-	+	+	-
9	Mango-chat	+	+	+	-	+	+	-

Таблица 3. Преимущества панели оператора платформ онлайн-консультирования

№	Название	Преимущества панели оператора (уникальность)
1	Chatra	чат между операторами, мобильное приложение
2	JivoSite	функция «указка» обратный звонок
3	CarrotQuest	неограниченное количество операторов мобильное приложение
4	Talk-me	определение загруженности мобильное приложение
5	EnvyBox	-
6	Venyoo	обратный звонок
7	Comagic	анализ эффективности рекламы обратный звонок
8	RedHelper	наблюдение за посетителем, технология «Co-Browse»
9	Mango-chat	мобильное приложение

Использование искусственного интеллекта (ИИ) и чат-ботов в платформах онлайн-консультирования значительно расширяет возможности, улучшает эффективность, и помогает предоставлять персонализированную поддержку.

Автоматизация базовых консультаций:

– чат-боты и виртуальные ассистенты могут обрабатывать запросы на начальном этапе, задавая клиенту вопросы для сбора первичной информации и выявления основной проблемы. Это позволяет консультантам сосредоточиться на более сложных задачах.

– боты предоставляют ответы на часто задаваемые вопросы, инструкции по использованию платформы и общую информацию, освобождая специалистов от рутинных задач.

Персонализация взаимодействия:

– ИИ может предложить дополнительные материалы, которые могут быть полезны клиенту (статьи, видео, советы), повышая качество обслуживания.

Анализ текста и распознавание эмоций:

– ИИ способен анализировать текст на эмоциональные триггеры и выявлять сигналы тревоги, стресса или депрессии в речи клиента, что особенно важно в сфере психологического консультирования.

– на основе анализа настроения система может определять, когда необходимо вмешательство живого консультанта, тем самым предотвращая возможные кризисные ситуации.

Аналитика и прогнозирование поведения:

– ИИ анализирует поведение клиентов на платформе, выявляя закономерности, такие как частота обращений, причины отказа от услуг или предпочтительные способы связи. Это помогает платформе улучшать взаимодействие и адаптировать услуги.

– Прогнозирование позволяет консультантам заранее знать, какие ресурсы или инструменты могут понадобиться для конкретного клиента.

В таблице 4 приведено сравнение избранных продуктов по различным критериям использования ИИ в онлайн-консультировании, можно использовать шкалу оценки от 1 до 5, где: 1 — низкий уровень реализации, а 5 — максимальный уровень реализации.

Таблица 4. Сравнительная таблица использования ИИ в платформах онлайн-консультирования

№	Продукт	Автоматизация базовых консультаций	Персонализация взаимодействия	Анализ текста и распознавание эмоций	Аналитика и прогнозирование поведения
1	Chatra	4	4	3	4
2	JivoSite	5	5	4	5
3	CarrotQuest	4	5	3	4
4	Talk-me	3	3	3	3
5	EnvyBox	4	4	3	4
6	Venyoo	3	3	2	3
7	Comagic	5	5	5	5
8	RedHelper	4	4	3	4
9	Mango-chat	4	4	3	4

Рассмотрим более подробно процесс оценки использования ИИ в платформах онлайн-консультирования.

1. Автоматизация базовых консультаций:

– JivoSite и Comagic занимают высокие позиции благодаря функционалу автоматизации, включающему чат-ботов, которые могут обрабатывать запросы на начальном этапе и предоставлять ответы на часто задаваемые вопросы.

– Venyoo и Talk-me имеют более базовую функциональность по автоматизации.

2. Персонализация взаимодействия:

– JivoSite и CarrotQuest предлагают мощные инструменты персонализации, такие как рекомендации на основе предыдущих взаимодействий, персонализированные ответы и предложения, что повышает качество обслуживания.

– Venyoo и Talk-me имеют ограниченные возможности в персонализации, предлагая более базовую настройку взаимодействия с клиентом.

3. Анализ текста и распознавание эмоций:

– Comagic выделяется в анализе текста и распознавании эмоций, предлагая более точное определение эмоциональных триггеров, что важно для психологического консультирования.

– JivoSite также предлагает функции для анализа текста и выявления эмоций, но в меньшей степени, чем Comagic.

– Продукты, такие как Venyo и Talk-me, не предоставляют достаточно развитые функции в этом направлении.

4. Аналитика и прогнозирование поведения:

– Comagic и JivoSite обеспечивают аналитические функции высокого уровня, включая прогнозирование поведения пользователей на платформе и адаптацию сервисов под предпочтения клиентов.

– CarrotQuest и RedHelper также предлагают хорошие аналитические инструменты, но с меньшей глубиной, чем JivoSite и Comagic.

– Продукты вроде Venyo и Talk-me предоставляют базовую аналитику, но не включают в себя расширенные функции прогнозирования.

На основании проведенного анализа популярных платформ онлайн-консультирования можно выделить общие тенденции:

– В настоящее время автоматизация и интеграция с внешними сервисами играют ключевую роль в выборе платформ для общения с клиентами. Чем больше возможностей для автоматизации процессов, интеграций и поддержки различных каналов связи, тем выше вероятность, что такие платформы будут востребованы.

– Мобильность и омниканальность становятся важными аспектами. Компании предпочитают платформы, которые могут работать с несколькими каналами связи, включая социальные сети и мессенджеры.

– Аналитика и мониторинг становятся неотъемлемыми функциями для оценки работы и улучшения клиентского сервиса.

– Платформы, предлагающие гибкость и масштабируемость (например, возможность увеличивать количество операторов по мере роста компании или настраивать дополнительные функции), особенно ценятся.

– Таким образом, выбор платформы сильно зависит от специфики сферы применения, его потребностей в автоматизации, масштабируемости, поддержки нескольких каналов связи и возможности интеграции с другими сервисами.

Результаты исследования. Исследование показало, что онлайн-консультирование значительно повышает доступность консультаций для пользователей, позволяет сэкономить время и ресурсы, а также повышает качество обслуживания. В научной сфере это способствует улучшению сотрудничества между исследователями и экспертами, а также повышению квалификации научных сотрудников через онлайн-курсы и тренинги. В государственном секторе онлайн-консультирование помогает улучшить взаимодействие с гражданами, обеспечивая доступ к важной информации и консультациям по различным вопросам, таким как законодательство, налогообложение и здравоохранение.

Однако, несмотря на все преимущества, существуют определенные вызовы. К ним относятся технические проблемы, такие как неравномерный доступ к интернету и устаревшие технологии, а также вопросы безопасности и конфиденциальности данных. Важным фактором является недостаток доверия со стороны некоторых пользователей, что связано с опасениями по поводу утечек данных.

Для повышения эффективности онлайн-консультирования необходимо улучшать инфраструктуру, интегрировать современные технологические решения, такие как искусственный интеллект и чат-боты, а также развивать мобильные платформы для консультирования «на ходу». Важно улучшать безопасность и защиту данных пользователей, а также обеспечивать высокую квалификацию специалистов, работающих на платформах.

Выводы. Несмотря на явные преимущества онлайн-консультирования, такие как доступность, экономия времени и снижение затрат, существует ряд вызовов. К числу таковых относятся проблемы с технологической инфраструктурой в некоторых регионах, вопросы безопасности и защиты данных, а также ограничение в виде недостатка личного общения, которое иногда необходимо для более глубокого обсуждения сложных вопросов. Цифровизация различных сфер жизни приведет к тому, что средства онлайн-консультирования станут неотъемлемой частью информационного пространства. При их интеграции в цифровые платформы различных областей целесообразно учесть особенности организации средств онлайн-консультирования, рассмотренных в работе и применяемых в настоящее время в коммерческой сфере.

СПИСОК ИСТОЧНИКОВ

1. Лаврентьев, С. Ю. Тенденции развития консалтинговой деятельности в инновационной образовательной среде вуза / С. Ю. Лаврентьев, Л. Г. Ахметов, Д. А. Крылов // Вестник Марийского государственного университета. – 2023. – Т. 17, № 1(49). – С. 19-28.
2. Рыбак, К. В. Обзор современного состояния интеллектуальных вопросно-ответных систем / К. В. Рыбак, А. В. Кошкарров // Вестник науки. – 2020. – Т. 1, № 6(27). – С. 202-205.
3. Лаврентьев, С. Ю. Возможности и перспективы использования технологий искусственного интеллекта в информационно-образовательном консультировании / С. Ю. Лаврентьев, Д. А. Крылов // Вестник Марийского государственного университета. – 2024. – Т. 18, № 1(53). – С. 27-35.
4. Kozhevnikov, V. A. Integrating online counseling services with Just AI Conversational Platform / V. A. Kozhevnikov, K. V. Yaruln // Theoretical & Applied Science. – 2019. – No. 7(75). – P. 26-35.
5. Спиридонова, Ю. С. Чат-платформа как инструмент оперативного взаимодействия работников железнодорожного транспорта / Ю. С. Спиридонова // Техника и технология наземного транспорта : Материалы международной студенческой научно-практической конференции (Нижний Новгород, 18 декабря 2019 г.). – Ч. 2. Нижний Новгород: ООО "НИЦ "21 век", 2020. – С. 471-474.
6. Палкин, Н. А. Преимущества использования мессенджеров в маркетинговой стратегии на примере Whatsapp / Н. А. Палкин // Актуальные вопросы экономики: сборник статей II Международной научно-практической конференции, Пенза, 25 мая 2020 года. – Пенза: "Наука и Просвещение", 2020. – С. 61-63.
7. Крапивина, М. Ю. Оптимизация самостоятельной работы студентов при изучении иностранного языка / М. Ю. Крапивина // Science Time. – 2014. – № 3. – С. 68-72.
8. Голубев, М. В. Взаимоотношения врача и пациента в сфере электронного здравоохранения / М. В. Голубев, Н. П. Лямина, В. П. Зайцев // Российский журнал телемедицины и электронного здравоохранения. – 2024. – Т. 10, № 2. – С. 19-26.
9. Гагарина, Е. Ю. Коммуникативная компетентность врача в условиях онлайн-консультирования / Е. Ю. Гагарина, Ю. А. Сторожева // Вестник Российского нового университета. Серия: Человек в современном мире. – 2020. – № 4. – С. 68-73.
10. Белодед, Д. Р. Цифровое пространство как поле профессиональной работы психологов уголовно-исполнительной системы / Д. Р. Белодед // VI международный пенитенциарный форум «Преступление, наказание, исправление» : Сборник тезисов выступлений и докладов участников (приуроченный к 30-летию со дня принятия Конституции Российской Федерации и Закона Российской Федерации от 21 июля 1993 г. № 5473-1 «Об учреждениях и органах уголовно-исполнительной системы Российской Федерации»), Рязань, 15–17 ноября 2023 года. – Рязань: Академия ФСИН России, 2023. – С. 36-39.
11. Троян, Н. А. Влияние цифровых технологий на процессы государственного управления / Н. А. Троян // Правовая информатика. – 2024. – № 3. – С. 180-187.

12. Основные направления совершенствования использования суперсервиса "поступление в вуз онлайн" в работе приемной комиссии Казанского ГАУ / Н. А. Сафиуллин, А. Р. Валиев, Б. Г. Зиганшин, Г. Р. Фассахова // Вестник Казанского государственного аграрного университета. – 2023. – Т. 18, № 3(71). – С. 186-192.
13. Новичкова, Е. Е. Оказание государственных услуг населению. Повышение качества и доступности их предоставления / Е. Е. Новичкова, Д. В. Зинченко // Вестник Восточно-Сибирского института МВД России. – 2023. – № 2(105). – С. 64-74.
14. Главатских, К. Е. Проектирование и документирование информационной системы для онлайн-консультирования / К. Е. Главатских // Актуальные проблемы математики, механики и информатики : сборник статей по материалам студенческой конференции, Пермь, 25 мая – 10 2021 года. – Пермь: Пермский государственный национальный исследовательский университет, 2021. – С. 142-148.
15. Главатских, К. Е. Проектирование, документирование и разработка информационной системы для онлайн-консультирования / К. Е. Главатских // Актуальные проблемы математики, механики и информатики 2022 : Сборник статей по материалам студенческой конференции, Пермь, 25 мая – 10 2022 года / Под редакцией А.В. Черникова, М.М. Бузмаковой. – Пермь: Пермский государственный национальный исследовательский университет, 2022. – С. 30-35.
16. Алтынхузина, Р. Н. Онлайн-консультант как IT-решение для общения с клиентами на сайтах / Р. Н. Алтынхузина // Перспективы развития науки и образования : Сборник научных трудов по материалам XXX международной научно-практической конференции, Москва, 29 июня 2018 года / Под общей редакцией А.В. Туголукова. – Москва: ИП Туголуков Александр Валерьевич, 2018. – С. 136-140.

Поступила в редакцию 06.09.2024 г., рекомендована к печати 23.09.2024 г.

DIGITAL PLATFORMS FOR ONLINE CONSULTING: COMPARATIVE ANALYSIS AND PROSPECTS

Nizhnikova O.L., Dorokhina G.V., Bolshakova S.A.

The article discusses online consulting as an effective tool for obtaining expert recommendations through digital platforms. The focus is on the structure of such platforms and their functionality, including the ability to communicate with a consultant in real time. This improves customer service quality and enhances interaction with them. Analysis of different types of platforms helps identify best practices for optimizing consultation processes and improving service, which contributes to more efficient problem solving for users and organizations.

Keywords: online consulting, question-and-answer mechanism, web chat, consulting.

Нижникова Олеся Леонидовна

младший научный сотрудник отдела системного анализа и интеллектуальных интерфейсов ФГБНУ «Институт проблем искусственного интеллекта»,
Российская Федерация, ДНР, г. Донецк.
E-mail: lesia.niznikova@mail.ru

Nizhnikova Olesia Leonidovna

Junior Researcher at the Department of System Analysis and Intelligent Interfaces at the Institute of Artificial Intelligence Problems,
Russian Federation, DPR, Donetsk.

Дорохина Галина Владимировна

заведующий отделом системного анализа и интеллектуальных интерфейсов ФГБНУ «Институт проблем искусственного интеллекта»,
Российская Федерация, ДНР, г. Донецк.
E-mail: sgv_iai@mail.ru

Dorokhina Galina Vladimirovna

Head of the Department of System Analysis and Intelligent Interfaces at the Institute of Artificial Intelligence Problems,
Russian Federation, DPR, Donetsk.

Большакова Светлана Анатольевна

младший научный сотрудник отдела распознавания речевых образов ФГБНУ «Институт проблем искусственного интеллекта»,
Российская Федерация, ДНР, г. Донецк.
E-mail: svetlako@yandex.ru

Bolshakova Svetlana Anatolevna

Junior Researcher at the Speech Pattern Recognition Department of the Institute of Artificial Intelligence Problems,
Russian Federation, DPR, Donetsk.

ПОСТРОЕНИЕ КОМПЬЮТЕРНОЙ МОДЕЛИ РЕШЕНИЯ ЗАДАЧ ТЕОРИИ ПОТЕНЦИАЛА С ПОМОЩЬЮ МГЭ

© 2024 Глушанков Е.С., Гольцев А.С., Мельник А.-В.В., Мироненко А.Б.

Приведена методика построения компьютерной модели решения задач теории потенциала с помощью метода граничных элементов. Представлен алгоритм решения классических задач теории потенциала с помощью прямого метода граничных элементов. Описана структура программы, реализующей данный метод. Перечислены программные средства, используемые для построения компьютерной модели полученного решения. Приведен демонстрационный пример построенной компьютерной модели.

Ключевые слова: метод граничных элементов, прямой МГЭ, теория потенциала, компьютерная модель, системы компьютерной математики.

Введение. Компьютерное моделирование становится неотъемлемой частью изучения сложных систем. Компьютерные модели существенно проще и удобнее, чем натурные испытания, требующие существенных финансовых и физических затрат. Эффективность использования компьютерного моделирования подробно описана в литературе [1, 2].

Изучение сложных физических процессов требует использование численных и численно-аналитических методов. К таким методам можно отнести метод конечных элементов (МКЭ) и метод граничных элементов (МГЭ). Если развитие МКЭ привело к появлению мощных программных комплексов, таких как ANSYS, ABAQUS, то развитие МГЭ идёт по пути создания программ, учитывающих индивидуальные обстоятельства каждой решаемой задачи. Каждый из методов имеет свои достоинства и недостатки. К неоспоримому достоинству МГЭ относится адекватное моделирование физических процессов с частью границы рассматриваемой области, уходящей в бесконечность [3, 4]. Поэтому его использование для задач геомеханики является приоритетным.

Освоение алгоритмов МГЭ и методов их реализации требует постепенного усложнения. Начальным этапом в этом процессе является применение МГЭ к классическим задачам математической физики. При этом результаты решения этих задач необходимо сразу представлять в виде компьютерных моделей. Разработке такой схемы построения компьютерной модели, применительно к задачам теории потенциала, и посвящена данная статья.

Суть прямого МГЭ для решения задач теории потенциала. Для освоения и отработки алгоритма прямого МГЭ как нельзя лучше подходит вариант задач теории потенциала, для которого схема использования МГЭ достаточно хорошо известна [3-5].

Рассмотрим двумерное уравнение теории потенциала, которое имеет следующий вид:

$$\nabla^2 u(x, y) = f(x, y) \quad (x, y \in \Omega), \quad (1)$$

Исследование проводилось в рамках государственного задания (№ госрегистрации 124012400353-3).

где ∇^2 – оператор Лапласа; $u(x, y)$ – решение, которое представляет собой потенциал в точке (x, y) от функции источника $f(x, y)$, распределённой в области Ω . Если $f(x, y) = 0$, то уравнение (1) называется уравнением Лапласа.

Решение уравнения (1) ищется на плоскости в замкнутой области Ω с границей Γ . На линии границы могут быть заданы значения самой функции u или её производной по нормали $\partial u / \partial n$ к контуру Γ . В зависимости от типа граничных условий различают задачу Дирихле, когда заданы значения лишь u , задачу Неймана, когда заданы значения лишь $\partial u / \partial n$, и смешанную задачу, когда заданы обе граничные величины. Рассмотрим постановку смешанной задачи:

$$\nabla^2 u(x, y) = 0 \text{ в } \Omega; \quad u = \bar{u} \text{ на } \Gamma_1; \quad \frac{\partial u}{\partial n} = \bar{u}_n \text{ на } \Gamma_2; \quad \Gamma_1 + \Gamma_2 = \Gamma, \quad (2)$$

где \bar{u} и \bar{u}_n – известные функции, заданные на границе.

Ключевым элементом в схеме МГЭ является фундаментальное решение тех дифференциальных уравнений, которые описывают рассматриваемый физический процесс. Уравнение Лапласа описывает поведение многих физических явлений. Это задачи установившихся процессов, таких как течение жидкости, тепловой поток, передачу электричества, прогибы мембран и т. д. Оно относится к уравнениям эллиптического типа.

Фундаментальное решение имеет определённый физический смысл и в случае уравнения Лапласа поясняется следующими рассуждениями.

Пусть в бесконечной плоскости в точке $P(x, y)$ находится точечный источник единичной мощности, который порождает значения искомой функции u , удовлетворяющей уравнению (1). Функция плотности этого источника $f(x, y)$ в любой точке бесконечной области, в частности в Q , может быть выражена с помощью дельта-функции Дирака (δ) следующим образом [6, 7]:

$$f(Q) = \delta(Q - P),$$

где двумерная дельта-функция Дирака $\delta(Q - P)$ определена как:

$$\int_{\Omega} \delta(Q - P) H(Q) d\Omega = H(P), \quad P \in \Omega \quad (3)$$

для произвольной функции $H(Q)$, непрерывной в области Ω .

Тогда искомый потенциал v в точке Q должен удовлетворять уравнению:

$$\nabla^2 v = \delta(Q - P). \quad (4)$$

Частное сингулярное решение этого уравнения называется фундаментальным решением потенциального уравнения (1). Оно хорошо известно [5, 6] и даётся формулой:

$$v = \frac{1}{2\pi} \ln r, \quad (5)$$

где r – радиальная координата точки Q в полярной системе координат, начало которой совмещено с точкой P . Структура этого решения позволяет заключить, что если точки P и Q поменять местами, то решение не изменится, т.е.:

$$v(Q, P) = v(P, Q), \quad (6)$$

что будет использовано в последующих преобразованиях.

Основу прямого МГЭ составляет тождество взаимности, которое связывает два различных состояния системы внутри области и на границе. Для гармонических функций этим тождеством является вторая формула Грина [8]:

$$\int_{\Omega} (v \nabla^2 u - u \nabla^2 v) d\Omega = \int_{\Gamma} \left(v \frac{\partial u}{\partial n} - u \frac{\partial v}{\partial n} \right) ds.$$

Применяя это тождество к функциям u и v , которые удовлетворяют уравнениям (2) и (4) при нахождении источника в точке P , получим:

$$-\int_{\Omega} u(Q) \delta(Q - P) d\Omega = \int_{\Gamma} \left[v(q, P) \frac{\partial u(q)}{\partial n_q} - u(q) \frac{\partial v(q, P)}{\partial n_q} \right] ds_q, \quad q \in \Gamma,$$

здесь строчной буквой q обозначена точка на граничном контуре Γ , точки P и Q находятся внутри области Ω .

Учитывая определение дельта-функция Дирака (3) и свойство фундаментального решения (6) из предыдущего соотношения получаем интегральное представление решения уравнения Лапласа внутри области через интеграл по граничному контуру от граничных значений самой функции и её производной по нормали к контуру:

$$u(P) = -\int_{\Gamma} \left[v(P, q) \frac{\partial u(q)}{\partial n_q} - u(q) \frac{\partial v(P, q)}{\partial n_q} \right] ds_q. \quad (7)$$

В этом представлении v – фундаментальное решение уравнения Лапласа (5), а $\partial v / \partial n$ – его производная по нормали, которая определяется формулой [5]:

$$\frac{\partial v}{\partial n} = \frac{1}{2\pi} \frac{\cos \varphi}{r},$$

где $r = |q - P|$, расстояние между точками P и q ; φ – угол между вектором из точки P к точке q и внешней нормалью к контуру Γ в точке q .

Рассмотрим схему МГЭ сразу для случая смешанных граничных условий. Тогда в одной точке граничного контура q может быть задана только одна из известных граничных функций, или сама функция u , или её производная по нормали $\partial v / \partial n$. Поэтому представление (7) ещё не даёт полного решения задачи. Чтобы его определить нужно найти неизвестную граничную величину, не заданную граничными условиями. Для этого осуществляется переход на линию граничного контура Γ . Точку рассмотрения P устремляем на границу области и получаем интегральное представление для функции u в точках $P \equiv p$ на граничном контуре Γ . Для гладкой границы без угловых точек такое интегральное представление имеет следующий вид [5]:

$$\frac{1}{2} u(p) = -\int_{\Gamma} \left[v(p, q) \frac{\partial u(q)}{\partial n_q} - u(q) \frac{\partial v(p, q)}{\partial n_q} \right] ds_q. \quad (8)$$

Это уравнение и является граничным интегральным уравнением, которое позволяет найти те граничные величины, которые не заданы граничными условиями. После его решения с помощью интегральных представления (7) и (8) находится решение задачи теории потенциала (2) во всех внутренних точках области и на её границе.

Схема численной реализации. Сущность метода граничных элементов состоит в дискретизации граничного контура Γ конечным числом сегментов, которые являются граничными элементами. Рассмотрим схему реализации МГЭ с постоянными элементами, которые представляют собой отрезок прямой линии, в середине которой располагается узловая точка, где определяется значение граничной величины. Значение граничной величины принимается постоянным по всей длине данного элемента.

Граничный контур Γ разбивается на N одинаковых граничных элементов. Они нумеруются в направлении против часовой стрелки. Значения граничной величины обозначим u , а её производной по нормали $\partial u / \partial n - u_n$. Тогда для каждого i -того элемента можно записать уравнение (8), суммируя значения необходимых интегралов по каждому элементу:

$$\frac{1}{2}u^i(p_i) = -\sum_{j=1}^N \int_{\Gamma_j} v(p_i, q) \frac{\partial u(q)}{\partial n_q} ds_q + \sum_{j=1}^N \int_{\Gamma_j} u(q) \frac{\partial v(p_i, q)}{\partial n_q} ds_q. \quad (i, j=1 \dots N) \quad (9)$$

Здесь Γ_j – элемент, где расположен j -й узел; p_i – узловая точка i -го элемента.

Так как граничные значения u и u_n постоянны на каждом элементе, то их можно вынести за знак интеграла и уравнение (9) переписать следующим образом:

$$-\frac{1}{2}u^i + \sum_{j=1}^N \left(\int_{\Gamma_j} \frac{\partial v}{\partial n} ds \right) u^j = \sum_{j=1}^N \left(\int_{\Gamma_j} v ds \right) u_n^j, \quad (10)$$

где u^j и u_n^j являются значениями u и u_n на j -ом элементе.

Интегральные величины в (10) называются коэффициентами влияния. Введём для них обозначения \hat{H}_{ij} и G_{ij} и будем определять по формулам:

$$\hat{H}_{ij} = \int_{\Gamma_j} \frac{\partial v(p_i, q)}{\partial n_q} ds, \quad G_{ij} = \int_{\Gamma_j} v(p_i, q) ds, \quad (11)$$

где узловая точка p_i является фиксированной, а точка интегрирования q изменяется по j -му элементу. С учётом этих обозначений дискретная форма записи уравнения (10) примет вид:

$$-\frac{1}{2}u^i + \sum_{j=1}^N \hat{H}_{ij} u^j = \sum_{j=1}^N G_{ij} u_n^j. \quad (12)$$

Введя обозначение:

$$H_{ij} = \hat{H}_{ij} - \frac{1}{2} \delta_{ij},$$

где δ_{ij} – символ Кронекера, уравнение (12) может быть записано в виде:

$$\sum_{j=1}^N H_{ij} u^j = \sum_{j=1}^N G_{ij} u_n^j.$$

Так как индекс i связан с каждой узловой точкой всех N элементов, приходим к системе N линейных алгебраических уравнений (СЛАУ). Её матричная форма имеет вид:

$$[H]\{u\} = [G]\{u_n\}, \quad (13)$$

где $[H]$ и $[G]$ – квадратные матрицы $N \times N$, $\{u\}$ и $\{u_n\}$ – векторы размерности N .

Из уравнений (13) следует, что при задании только одной граничной величины, u или u_n , друга также может быть найдена. Осуществляется это по следующей схеме.

Пусть на части границы Γ_1 заданы N_1 граничных значений функции u , а на Γ_2 – N_2 граничных значений производной по нормали u_n ($\Gamma_1 \cup \Gamma_2 = \Gamma$, $N_1 + N_2 = N$). В этом случае неизвестные величины в системе (13) будут находиться по разные стороны равенства в зависимости от граничных условий. Поэтому до решения СЛАУ необходимо отделить известные величины от неизвестных и представить систему (13) в виде:

$$\begin{bmatrix} [H_{11}] & [H_{12}] \end{bmatrix} \begin{Bmatrix} \{\bar{u}\}_1 \\ \{u\}_2 \end{Bmatrix} = \begin{bmatrix} [G_{11}] & [G_{12}] \end{bmatrix} \begin{Bmatrix} \{u_n\}_1 \\ \{\bar{u}_n\}_2 \end{Bmatrix},$$

где известными величинами являются $\{\bar{u}\}_1$ и $\{\bar{u}_n\}_2$, а неизвестными – $\{u_n\}_1$ и $\{u\}_2$. Переносим неизвестные величины в левую часть уравнения, а известные в правую, получим:

$$[A]\{X\} = \{B\}, \quad (14)$$

где: $[A] = \begin{bmatrix} [H_{12}] - [G_{11}] \end{bmatrix}$, $\{X\} = \begin{Bmatrix} \{u\}_2 \\ \{u_n\}_1 \end{Bmatrix}$, $\{B\} = -[H_{11}]\{\bar{u}\}_1 + [G_{12}]\{\bar{u}_n\}_2$,

где $[A]$ – квадратная матрица $N \times N$; $\{X\}$, $\{B\}$ – векторы размерности N .

После решения СЛАУ (14) все граничные величины будут известны и решение внутри области можно определить в любой точке $P(x, y)$. При этом используется интегральное представление (7) и та же самая дискретизация границы, как и в уравнении (9):

$$u(P) = \sum_{j=1}^N \hat{H}_{ij} u^j - \sum_{j=1}^N G_{ij} u_n^j. \quad (15)$$

Однако коэффициенты \hat{H}_{ij} и G_{ij} в (15) необходимо вычислять снова по тем же формулам (11) в которых граничная точка p_i заменяется внутренней точкой P .

Описание программы. Описанный алгоритм решения задач теории потенциала был реализован на алгоритмическом языке C++. Укрупнённая блок-схема компьютерной программы представлена на рис. 1.

Помимо вычислений, реализованных по приведённым формулам, в программе можно выделить следующие характерные особенности.

1. В виду того, что исходные данные представляют собой набор числовых векторов, ввод которых программно требовал бы перекомпиляции проекта, а ввод с помощью клавиатуры – большого количества времени, в качестве механизма ввода данных используется чтение из файла.

Для удобства и универсальности ввода введена возможность добавления комментариев, не влияющих на считывание данных, т.е. построен некоторый простейший синтаксический анализатор.

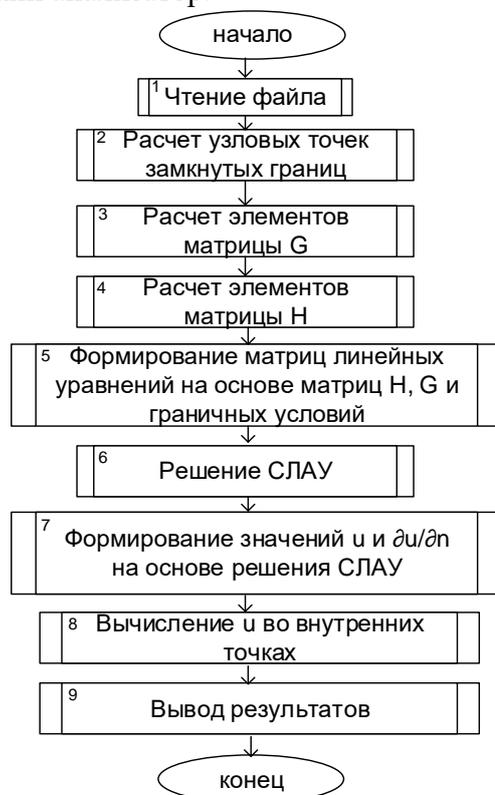


Рис. 1. Укрупненная структура программы

2. Решение СЛАУ осуществляется методом Гаусса с максимальным разрешающим элементом по столбцу. Если в результате вычислений система уравнений получится вырожденной, соответствующее сообщение будет отображено на экране, а в файл будут записаны только выполненные промежуточные вычисления.

3. Для реализации описанного алгоритма был выбран объектно-ориентированный подход. Структура используемых типов и классов представлена на рис. 2.

Блоки, указанные на рис. 1 частично перешли в функции класса `CflatTask`. Так, чтение файла и расчет узловых точек замкнутых границ реализованы в конструкторе класса и функции `calculateCoordsBounds`, которая вызывается в конструкторе. Расчетная работа программы осуществляется в функции `calculate`, которая содержит вызовы функций 3-8.

Далее приведены имена функций в соответствии с номерами в блок-схеме:

- 3 Функция `gmatr`, которая возвращает рассчитанную матрицу G .
- 4 Функция `hmatr`, которая возвращает рассчитанную матрицу H .
- 5 Функция `abMatrEq`, которая возвращает матрицу A , на основе которой реализуется решение системы линейных уравнений.
- 6 Функция `solveEq`, которая возвращает логическое значение, соответствующее `true`, если система не может быть решена из-за вырожденности. В качестве параметров функции передаются матрица A и вектор свободных коэффициентов, причем в последний сохраняются результаты решения СЛАУ. Эта функция не является членом класса, и может в дальнейшем быть заменена на вариант с параллельным вычислением решения.

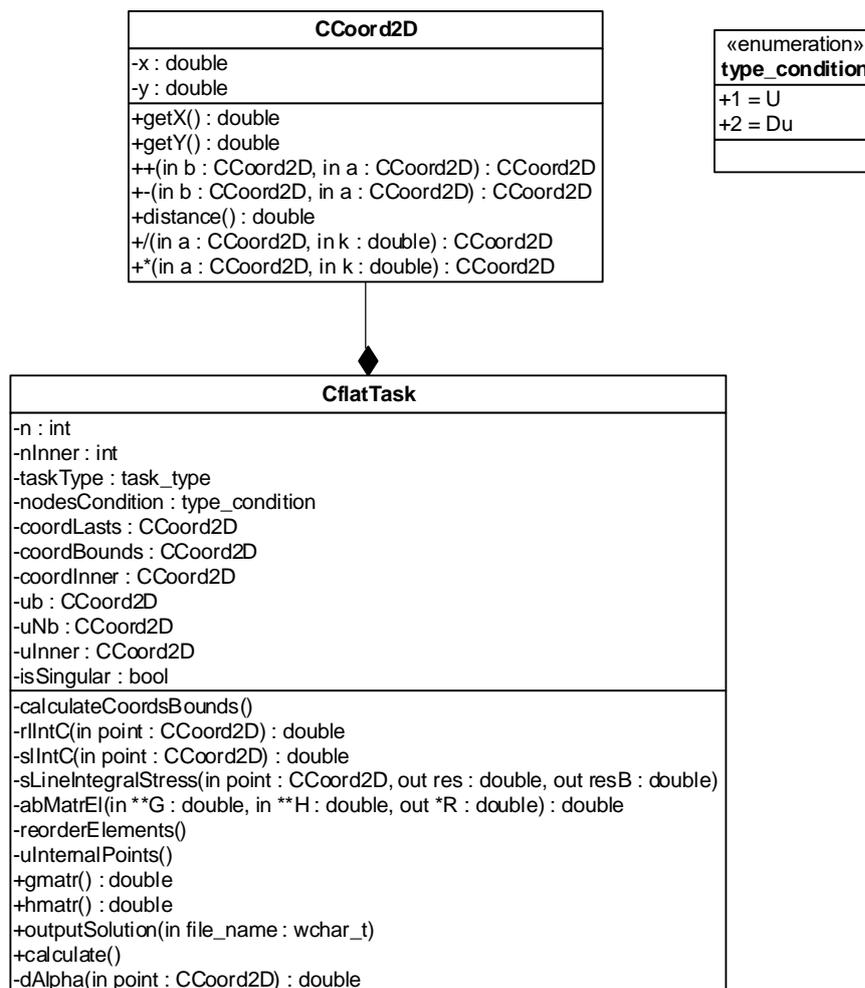


Рис. 2. Структура классов программы

- 7 Процедура `reorderElements` осуществляет сохранение значений результата решения СЛАУ из вектора r в массивы `ub` и `uNb` в зависимости от того, какой тип условий был задан изначально.
- 8 Процедура `uInternalPoints` на основе вычисленных в 7 значений определяет числовые значения вектора `uInner`, хранящего значения u во внутренних точках.

Вывод результатов расчетов реализован в функции класса `outputSolution`.

Построение компьютерной модели. Разработанная программа протестирована на простом известном примере [5]. Была рассмотрена область Ω в виде единичного квадрата (рис. 3). На границах квадрата заданы смешанные граничные условия. Расчёт проводился для 16 постоянных элементов по 4 элемента на каждой стороне.

Данная задача имеет точное решение, которое выражается формулой [5]:

$$u(x, y) = 100(1 + x). \quad (16)$$

Результаты расчетов программы сошлись с эталонными данными [5] и с аналитическим решением (16).

Для визуализации результатов тестового варианта, которая представлена на рис. 4, использована система компьютерной математики MatLab.

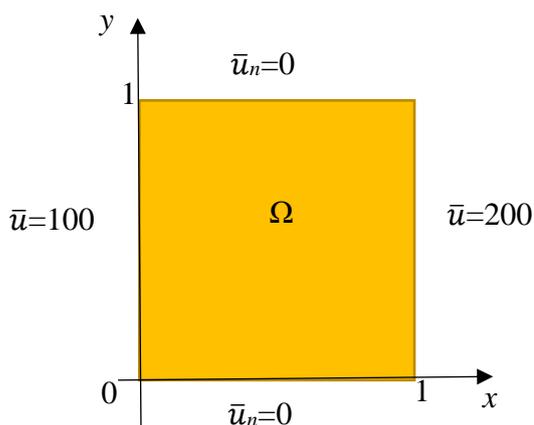


Рис. 3. Область Ω и граничные условия

Количество точек расчета во внутренней области бралось равным 25. Сформированные на выходе программы данные были считаны в MatLab с помощью функции `readtable`. Дальнейшая обработка данных осуществлялась с помощью стандартной работы с массивами. Эскиз графика построен с использованием цветовой градиентной заливки функции `surf` (рис. 4).

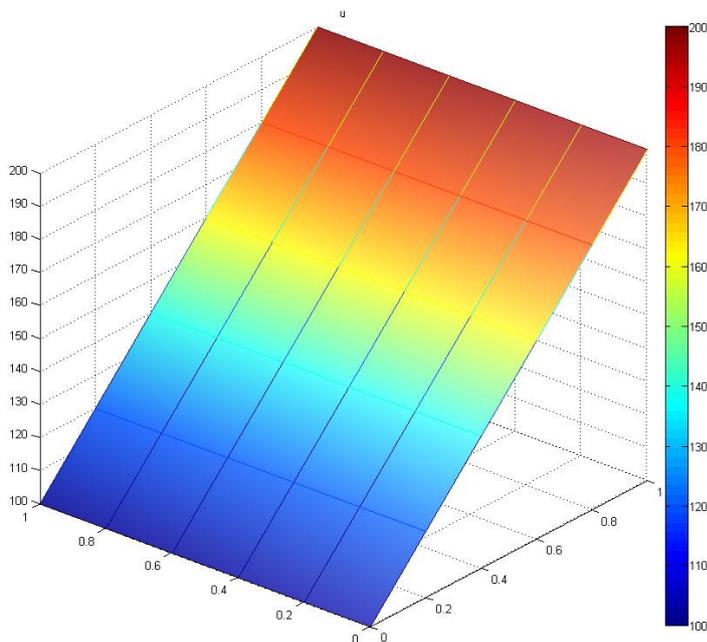


Рис. 4. Значение функции u во внутренних и граничных точках

В качестве демонстрации возможностей программы рассмотрен пример для той же области, но с более сложными граничными условиями. Вся граница разбита на четыре участка по каждой стороне квадрата. На каждом участке граничная величина F (\bar{u} или \bar{u}_n) изменяется по квадратичному закону $F = az^2 + bz + c$, где z – координата на линии участка от его начала, совмещённого с концом предыдущего участка. Схема разбиения участков, тип граничной величины и значения констант в законе изменения граничной величины взяты следующим образом:

1 участок: $x = [0,1]$, $y = 0$; \bar{u} ; $a = 10$, $b = 20$, $c = -30$;

2 участок: $x = 1$, $y = [0,1]$; \bar{u}_n ; $a = 0$, $b = 0$, $c = 25$;

3 участок: $x = [0,1]$, $y = 1$; \bar{u} ; $a = 0$, $b = 20$, $c = 10$;

4 участок: $x = 0$, $y = [0,1]$; \bar{u}_n ; $a = 0$, $b = -20$, $c = -10$.

Задача состояла в определении функции u внутри рассматриваемой области. Использован МГЭ с 80 граничными элементами и решение находилось в 81 внутренней точке. Далее использована функция PLOT3D системы компьютерной математики Maple для визуализации полученного решения. На рис. 5 показан трёхмерный график искомой функции u в рассматриваемой области.

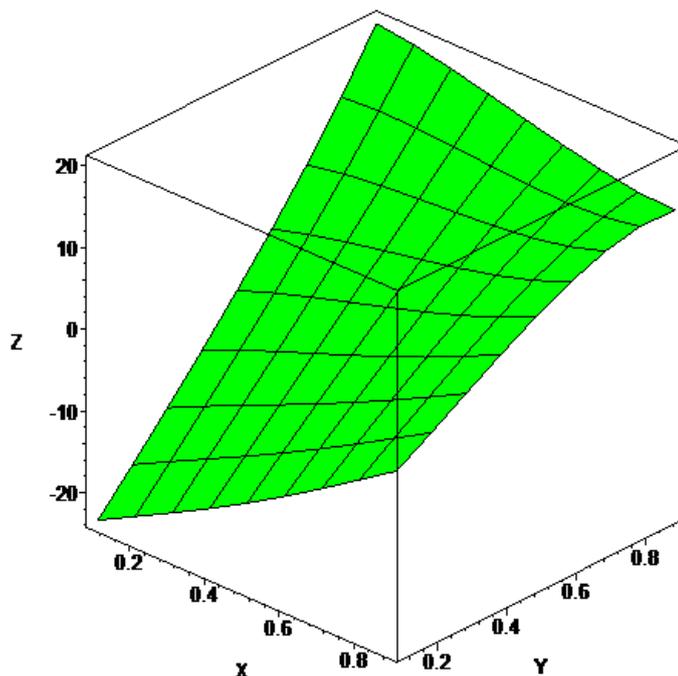


Рис. 5. Распределение функции u внутри области Ω

Выводы. Таким образом, отработана методика решения задач теории потенциала с построением компьютерной модели полученного решения. Используются стандартные средства программирования и широко распространённые системы компьютерной математики.

Дальнейшее развитие работы предполагает использование МГЭ в плоских задачах теории упругости прямым и непрямым методами.

СПИСОК ИСТОЧНИКОВ

1. Боев, В. Д. Компьютерное моделирование / В. Д. Боев, Р. П. Сыпченко. – М.: Национальный Открытый Университет “ИНТУИТ”, 2016. – 526 с
2. Компьютерное моделирование: учебник / В. М. Градов, Г. В. Овечкин, П. В. Овечкин [и др.]. – М.: КУРС: ИНФРА-М, 2018. – 264 с.
3. Бенерджи, П. Методы граничных элементов в прикладных науках / П. Бенерджи, Р. Баттерфилд; пер. с англ. – М.: Мир, 1984. – 494 с.
4. Бреббия, К. Методы граничных элементов / К. Бреббия, Ж. Теллес, Л. Вроубел; пер. с англ. – М.: Мир, 1987. – 524 с.
5. Кацикаделис, Дж. Т. Граничные элементы. Теория и приложения / Дж. Т. Кацикаделис; пер. с англ. – М.: Издательство Ассоциации строительных вузов, 2007. – 348 с.

6. Владимиров, В. С. Обобщённые функции в математической физике / В. С. Владимиров – М.: Наука, 1979. – 320 с.
7. Кеч, В. Введение в теорию обобщённых функций с приложениями в технике / В. Кеч, П. Теодореску; пер. с румын. – М.: Мир, 1978. – 518 с.
8. Голоскоков, Д. П. Уравнения математической физики. Решение задач в системе Maple / Д. П. Голоскоков – СПб.: Питер, 2004. – 539 с.

Поступила в редакцию 11.09.2024 г., рекомендована к печати 01.10.2024 г.

CONSTRUCTION OF A COMPUTER MODEL FOR SOLVING PROBLEMS OF POTENTIAL THEORY USING BEM

Glushankov E.S., Goltsev A.S., Melnik A.-V.V., Mironenko A.B.

A method for constructing a computer model for solving potential theory problems using the boundary element method is given. An algorithm for solving classical problems of potential theory using the direct boundary element method is presented. The structure of the program implementing this method is described. The software tools used to construct a computer model of the obtained solution are listed. A demonstration example of the constructed computer model is given.

Keywords: boundary element method, direct BEM, potential theory, computer model, computer mathematics systems.

Глушанков Евгений Сергеевич

кандидат физико-математических наук, доцент кафедры теории упругости и вычислительной математики ФГБОУ ВО «Донецкий государственный университет», Российская Федерация, ДНР, г. Донецк.
E-mail: ie.glushankov@yandex.ru

Glushankov Evgenii Sergeevich

Candidate of Physico-Mathematical Sciences, Associate Professor at Department of Theory of Elasticity and Computational Mathematics of Donetsk State University, Russian Federation, DPR, Donetsk.

Гольцев Аркадий Сергеевич

доктор физико-математических наук, профессор, заведующий кафедрой прикладной механики и компьютерных технологий ФГБОУ ВО «Донецкий государственный университет», Российская Федерация, ДНР, г. Донецк.
E-mail: asgoltsev@mail.ru

Goltsev Arkadii Sergeevich

Doctor of Physico-Mathematical Sciences, Full Professor, Head of the Department of Applied Mechanics and Computer Technologies of Donetsk State University, Russian Federation, DPR, Donetsk.

Мельник Анна-Валентина Валентиновна

кандидат технических наук, доцент кафедры прикладной механики и компьютерных технологий ФГБОУ ВО «Донецкий государственный университет», Российская Федерация, ДНР, г. Донецк.
E-mail: Anna-Valentina@yandex.ru

Melnik Anna-Valentina Valentinovna

Candidate of Technical Sciences, Associate Professor at Department of Applied Mechanics and Computer Technologies of Donetsk State University, Russian Federation, DPR, Donetsk.

Мироненко Андрей Борисович

кандидат физико-математических наук, доцент кафедры теории упругости и вычислительной математики ФГБОУ ВО «Донецкий государственный университет», Российская Федерация, ДНР, г. Донецк.
E-mail: andreyandrey80@mail.ru

Mironenko Andrei Borisovich

Candidate of Physico-Mathematical Sciences, Associate Professor at Department of Theory of Elasticity and Computational Mathematics of Donetsk State University, Russian Federation, DPR, Donetsk.

МЕТОДЫ ПРОЕКТИРОВАНИЯ АРХИТЕКТУРЫ МОДЕЛИ ИНТЕЛЛЕКТУАЛЬНОГО АССИСТЕНТА ДЛЯ ОБРАЗОВАТЕЛЬНОЙ ОРГАНИЗАЦИИ

© 2024 *Елисеев В.О., Бондаренко В.И.*

Большинство наиболее успешных больших генеративных языковых моделей (LLM) состоят из трех составных частей: кодировщика, декодера и механизма внимания. В статье были рассмотрены основные подходы к проектированию каждой из этих частей, а также наиболее популярные архитектуры LLM. Кроме того, был проведен сравнительный анализ методов обучения и применения больших генеративных языковых моделей на предмет оптимальности в использовании для построения интеллектуального ассистента образовательной организации. Результаты исследования показали, что при имеющихся ресурсах наиболее подходящим способом является точная настройка существующей LLM на текстовом корпусе данных об образовательной организации с последующим ее использованием напрямую либо посредством механизма RAG.

Ключевые слова: генерация текста, LLM, GPT, точная настройка, fine-tuning, чат-бот.

Введение. В настоящее время в университетах РФ наблюдается большой рост числа бюджетных мест, а соответственно и абитуриентов. Следствием этого является затратность и сложность коммуникации с абитуриентами, своевременного информирования студентов по интересующим их вопросам, приводящие к простоям и негативно сказывающиеся на качестве процесса набора и обучения студентов.

Ведущие вузы РФ решают эту проблему путем внедрения на свои платформы различных по назначению и сложности чат-ботов – от обычных коммуникаторов пользователя с администрацией вуза до более продвинутых автоматизированных версий, оперирующих знаниями о вузе или интересующих конечного пользователя его аспектах.

В связи с этим появляется задача создания инструмента, позволяющего сократить до минимума время ожидания пользователями ответа от представителей вуза и упростить работу администрации. Наилучшим решением этой задачи является создание собственного чат-бота, работающего с помощью большой языковой модели (LLM) и имеющего доступ к релевантным и актуальным данным о конкретной образовательной организации, что позволит ему давать качественные ответы на вопросы пользователей.

Для решения подобной задачи необходимо выполнить следующие шаги:

- провести исследование существующих методов разработки / обучения больших генеративных языковых моделей;
- собрать текстовый корпус данных об образовательной организации (в данном конкретном случае будет рассматриваться Донецкий государственный университет), подготовить его для тренировки LLM;
- выбрать архитектуру генеративной модели, которая будет обучена на подготовленном текстовом корпусе;

Работа выполнена при финансовой поддержке Минобрнауки России в рамках научной темы «Разработка и совершенствование интеллектуальных методов классификации и прогнозирования для задач распознавания образов и моделирования информационных процессов» FREM-2024-0001 (Регистрационный номер 1023111000141-9-1.2.1)

- обучить LLM на текстовом корпусе;
- реализовать сервис интеллектуального ассистента университета с веб-интерфейсом, либо в виде бота на одной из платформ (Telegram, VK и т. д.), интегрировать обученную модель в этот сервис, проверить ее работоспособность.

В данной статье выполним первые два шага плана: рассмотрим наиболее популярные и качественные методы построения больших генеративных моделей и возможные подходы к процессу тренировки модели на данных о конкретной образовательной организации, выберем наиболее подходящий в рамках решения задачи и соберем корпус данных, на основе которого будет производиться обучение.

Целью работы является рассмотрение и анализ методов построения и способов обучения больших генеративных языковых моделей с целью выбора их наиболее подходящей комбинации для решения задачи построения генеративной LLM, управляющей интеллектуальным ассистентом образовательной организации и строящей свои ответы на основе данных о нем.

Описание входных данных. В качестве текстового корпуса для решения поставленной задачи будут использоваться данные, полученные из официального сайта образовательной организации. Согласно методическим рекомендациям [1], сайт образовательной организации обязательно должен содержать такие разделы, как: «Основные сведения», «Структура и органы управления образовательной организацией», «Документы», «Образование», «Образовательные стандарты», «Руководство. Педагогический (научнопедагогический) состав», «Материально-техническое обеспечение и оснащённость образовательного процесса», «Стипендии и иные виды материальной поддержки», «Платные образовательные услуги», «Финансово-хозяйственная деятельность», «1 Вакантные места для приема (перевода)», «Доступная среда», «Международное сотрудничество», «Абитуриенту», «Уровни образования «Бакалавриат», «Специалитет», «Магистратура»», «Уровень образования «Аспирантура»», «Уровень образования «Ординатура»». Кроме того, допускается создание и собственных разделов, отражающих специфику конкретного образовательного учреждения.

Данные в разделах сайта образовательного учреждения могут храниться как в текстовом формате на HTML-страницах, так и в различных файлах. Согласно рекомендациям [1], допускается предоставление файлов в форматах Portable Document Files (.pdf), Microsoft Word/Microsoft Excel (.doc, .docx, .xls, .xlsx), Open Document Files (.odt, .ods), также рекомендуется делать файлы машинночитаемыми.

Исходя из регламентированности предоставления данных образовательными организациями о себе и преобладанием текстовых документов на сайтах образовательных организаций можно сделать вывод, что текстовый корпус, полученный на основе данных с сайта конкретной образовательной организации, будет обладать достаточно хорошим качеством для целей обучения LLM.

Поскольку нам важно максимальное погружение в специфику работы конкретной образовательной организации, мы будем использовать не строго унифицированные разделы целевого сайта, а полную его информацию, которую можно получить путем обхода его страниц. С этой целью был создан набор скриптов, позволяющих распарсить контент всех HTML-страниц сайта, а также загрузить все файлы, на которые будут найдены ссылки при обходе страниц. Подробнее о скриптах рассказано в тезисах [2].

Общее описание генеративных моделей. Генеративные модели искусственного интеллекта – алгоритмы, которые способны генерировать новые данные, имитируя статистические свойства исходных данных. В контексте процедурной генерации текстов, генеративные модели используются для создания текстов, которые имеют определенную структуру и логику [3].

Во время обучения, такие модели (при решении задач генерации текста) получают на вход текстовый корпус, в зависимости от архитектуры модели преобразованный или нет, изучают его структуру и особенности и пытаются посимвольно создать новый текст на основе полученных знаний. Глобально такие модели состоят из кодировщика, декодера и иногда слоев внимания. Выбор конкретной реализации каждой из частей влияет на качество генерируемого текста.

Роль кодировщика – преобразовать входные текстовые данные в векторное представление, построить так называемые эмбединги [4]. Цель такого векторного преобразования – передать компьютеру такое числовое представление слов (формат, понимаемый им) чтобы эти числа отражали суть и семантику переданного слова. Таким образом на основе векторного расстояния между закодированными словами можно определить их взаимосвязь, что является первым шагом к успешной процедуре генерации. Стоит отметить, что построение эмбедингов не обязательно привязывать к словам, вполне возможно строить их на основе целых предложений или частей слов. Еще одним из полезных свойств эмбедингов является тот факт, что мы можем поставить в конец генеративной модели обычный линейный слой, обучить веса этого слоя на своем датасете (причем он может быть сравнительно небольшим) под собственную задачу (например, классификация текста) и такая модель также будет хорошо выполнять поставленную задачу. Подобный процесс называется тонкой настройкой модели.

Декодер преимущественно представляет собой RNN (рекуррентную нейронную сеть), получающую от кодировщика векторные представления и непосредственно генерирующую символьную последовательность. Благодаря своей архитектуре и учету предыдущих символов RNN является хорошим базовым решением для подобных задач.

Возможным улучшением архитектуры генеративной нейронной сети является добавление слоев внимания. Данный механизм позволяет модели фокусироваться на определенных частях входных данных, что делает процесс обучения более гибким и в целом увеличивает качество генерации.

Для оценки качества генерируемого текста можно использовать семантическое сравнение сгенерированных предложений с эталонными, языковые модели, обученные для оценки правдоподобия сгенерированного текста, использовать такие метрики как BLEU [5], ROUGE [6], perplexity [7], а также использовать человеческую оценку результатов генерации, что лежит в основе подхода RLHF [8].

Далее рассмотрим основные подходы к построению кодировщиков и декодеров, механизм внимания, а также затронем вопрос токенизации текста и успешные существующие архитектуры генеративных текстовых моделей.

Использование word2vec для генерации эмбедингов. Word2vec – метод эффективного создания вложений, разработанный в 2013 году. Кроме работы со словами, некоторые его концепции оказались эффективны в разработке рекомендательных механизмов и придании смысла данным в коммерческих, неязыковых задачах. Эту технологию применили в своих движках рекомендаций такие компании, как Airbnb, Alibaba, Spotify и Anghami.

Подробно реализация алгоритма word2vec описана в оригинальной статье [9]. Также в статье описывается один из возможных вариантов улучшения процесса обучения word2vec – архитектура skip-gram. Она основана на угадывании слова в середине сэмпла по предшествующим и последующим, то есть модель учится смотреть в 2 стороны. Еще одним способом совершенствования алгоритма является отрицательное сэмплирование [10].

В результате обучения word2vec мы получаем векторы для каждого слова, обладающие рядом интересных свойств. Во-первых, это свойство аналогии. Если из векторного представления слова «Париж» вычесть векторное представление слова «Франция» и добавить векторное представление слова «Россия», то ближайшим вектором к результату окажется векторное представление слова «Москва», что говорит о семантической наполненности полученного векторного представления для каждого слова, построенного моделью word2vec.

Еще одним свойством получаемых векторов является их близость, что означает сходство концептов слов, которые они кодируют. Это еще раз подтверждает семантическую наполненность векторных представлений слов и уже может использоваться для простейших задач классификации.

На базовом уровне алгоритм word2vec предоставляет весьма качественные эмбединги слов и может выступать в качестве кодировщика генеративной сети для самых простых задач. Для анализа предложений используются их векторные представления, полученные по формуле:

$$v(\text{sentence}) = \sum_w P(\text{word})v(\text{word}),$$

где v – векторное представление, P – вероятность появления конкретного слова, w – слово в предложении.

Таким образом появляется возможность векторизации предложений/текстов и их дальнейшей обработки.

FastText от Facebook. FastText – библиотека, содержащая предобученные готовые векторные представления слов и классификатор, который можно настроить под собственную задачу.

Векторные представления слов создавались с помощью word2vec с отрицательным сэмплированием. Существенным улучшением подхода к тренировке являлось использование subword-модели наряду с основной. Данное улучшение решает проблему OOV (out of vocabulary), когда на вход модели поступает новое слово, ни разу не появившееся в датасете, его эмбединг не построен и соответственно обработка такого слова может привести к ошибкам в работе модели. Subword-модель является представлением слова через цепочки символов (n -граммы) с n от 3 до 6 символов (число n преимущественно зависит от языка – для языков с достаточно короткими словами и простой морфологией можно использовать небольшие n , иначе это число должно быть больше) от начала до конца слова плюс само слово целиком. Например, если взять $n = 3$ и слово «шахматы», то в модели оно будет представлено как набор n -грамм: <ша, шах, ахм, хма, мат, аты, ты> и последовательностью <шахматы>, причем модель понимает разницу между последовательностью <шах> из слова шах и n -граммой шах из слова шахматы. Результатом будет являться сумма усредненных эмбедингов по n -граммам и эмбединга всего слова-последовательности. Такой подход позволяет работать с новыми словами – если последовательность не встречалась, ее эмбедингом будет усреднение эмбедингов n -грамм, что тоже несет семантический смысл и позволяет достигать результатов лучших, чем результаты word2vec. Более подробно процесс обучения fastText описан в статье [11].

Стоит отметить, что данный подход масштабируем и при наличии достаточного корпуса текста можно обучить модель на любом языке с символами любого вида. Таким образом были получены векторы слов для 157 языков, которые находятся в открытом доступе [12]. Исходный код библиотеки также находится в открытом доступе [13].

Рекуррентные нейронные сети в роли декодеров. В отличие от большинства архитектур нейронных сетей, поддерживающих фиксированные размеры ввода и вывода (линейные, сверточные нейронные сети), рекуррентные позволяют работать с входом и выходом разных размеров.

Работа с последовательностями обеспечивается следующим внутренним механизмом RNN: каждому последующему слою рекуррентной нейронной сети в качестве входа передается не только входной вектор с данными, но и выходной вектор предыдущего слоя, которые вместе (посредством конкатенации) образуют матрицу весов для конкретного слоя RNN. Таким образом нейронная сеть помнит свое состояние на предыдущем шаге и способна учитывать его при своей дальнейшей работе, что, например, полезно для задач генерации текста, когда мы по набору предыдущих символов-состояний предсказываем следующий символ. В качестве функции активации внутри блока RNN в основном используется гиперболический тангенс, поскольку он может работать и с отрицательными значениями (в отличие от RELU), а также растет менее быстро, что позволяет адекватно обработать большее количество контекста.

RNN является базовой архитектурой построения декодера, далее рассмотрим варианты ее улучшения и другие подходы к генерации текста.

Сети LSTM как более продвинутая вариация декодеров. Одной из серьезных проблем классических RNN является их неспособность удерживать в памяти слишком большие последовательности. Это происходит потому, что когда градиент проходит через большое количество ячеек, то либо он уменьшается настолько, что теряется связь между ошибками в начале и в конце последовательности, либо разрастается до такой степени, что процесс обучения расходится (даже при использовании более подходящей и плавно растущей функции активации).

Следующим шагом в развитии декодеров стала замена ячеек RNN на ячейки LSTM [14] (long short-term memory). К обычной ячейке RNN добавляется внутреннее состояние S_t , которое преобразовывается способом, основанном на использовании skip connections, при меньшем количестве преобразований, чем у идущего через все ячейки и градиенты состояния H_t , что помогает решить проблему исчезающих или взрывающихся градиентов. Также были добавлены input и output gate, влияющие на то, что из входных данных окажет влияние на внутреннее состояние ячейки и что из внутреннего состояния должно отправиться на выход.

Одной из последующих модификаций ячейки LSTM являлось добавление forget gate, задачей которого является зануление некоторых компонент S_{t-1} перед дальнейшей передачей. Это может быть особенно полезно, например при анализе длинной последовательности-отзыва на товар на предмет тональности. Так, в самом начале отзыва могут проступить негативные слова-маркеры, но дальше он приобретет смысл близкий к нейтральному или позитивному (возможна и обратная ситуация), поэтому иногда полезно «забыть» слишком длинный предварительный контекст.

Еще одним дополнением к базовой ячейке LSTM является peephole connection, которое учитывает состояние S_{t-1} для вычисления всех ворот, что делает поведение ячейки концептуально более осмысленным.

Таким образом LSTM является более продвинутым и качественно превосходящим вариантом декодера, в связке с word2vec или fastText (на вход передаются не символы, а векторные представления слов) такая архитектура может выдавать достаточно качественные результаты для задачи генерации текста и ряда других задач NLP.

Механизм внимания (attention). Несмотря на отличия LSTM от RNN, проблемы последней все равно имеют место быть. В частности, это заметно при решении задачи машинного перевода, при которой последовательности символов одной длины соответствует последовательность символов совершенно другой, и при попытке перевести текст слишком большого размера с помощью обычной сети LSTM люди сталкивались с проблемой удержания контекста. Улучшения в архитектуре, связанные с разделением на кодировщик и декодер, имеющие разные веса, а также ряд трюков, связанных с особенностями языка (попытка научить декодер предсказывать слова в обратном порядке и др.), механизмами дополнительной передачи контекста к блокам декодера привели к созданию архитектуры seq2seq [15], которая была SOTA-решением для домена машинного перевода в 2014 году. Однако такой подход не решал всех проблем, и вопрос обработки длинных последовательностей оставался открытым.

Одним из серьезных улучшений архитектуры является использование механизма внимания (attention), призванного искать взаимосвязи между различными частями входных и выходных данных. Впервые данный механизм был описан в статье [16] и был успешно применен для улучшения архитектуры seq2seq.

При помощи механизма внимания достигается «фокусирование» декодера на определенных скрытых состояниях. В задаче машинного перевода эта возможность помогает декодеру предсказывать на какие скрытые состояния при исходных определенных словах на языке A необходимо обратить больше внимания при переводе данного слова на язык B, то есть учесть другие слова в предложении при переводе, что «оживит» результат. Стоит отметить, что вся система внимания дифференцируема, а значит мы можем полностью ее оптимизировать в процессе градиентного спуска.

Результат работы такого механизма визуализирован на рисунке 3. На нем представлено переведенное с английского на французский язык предложение с помощью seq2seq с добавленным механизмом внимания. Чем ярче закрашенный прямоугольник, тем больше вес соответствующего слова в механизме внимания.

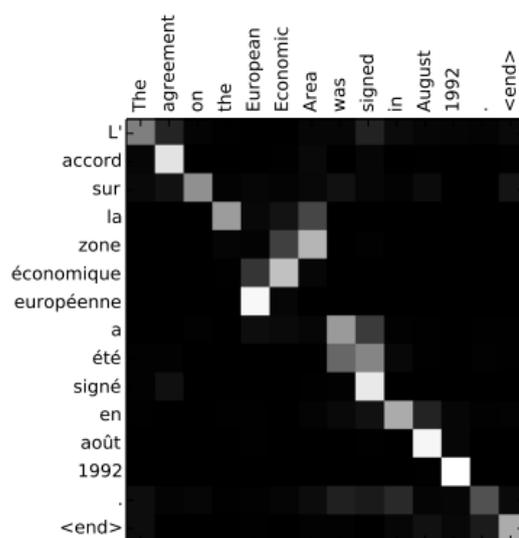


Рис. 1. Визуализация механизма внимания при переводе предложения с английского языка на французский

Архитектура Transformer. Новым шагом в развитии NLP стала архитектура Transformer, представленная Google в 2017 году в статье «Attention is all you need» [17]. Ее новизна связана с тем, что у модели отсутствуют блоки RNN, CNN и LSTM, а используется лишь механизм self-attention – внимание ко внутреннему состоянию (что и послужило причиной такого названия статьи), благодаря чему модель может сосредоточиться не на всех сразу, а на наиболее важных элементах входной последовательности, так как self-attention определяет веса важности элементов входной последовательности и не требует сохранять состояния, что решает упомянутые выше проблемы RNN. Схематически архитектура Transformer представлена на рисунке 4.

Сеть с одинаковыми весами параллельно запускается на каждом слове входной последовательности. На вход дается слово Inputs, проходящее векторизацию в блоке Input Embedding, далее результирующий вектор складывается с информацией о расположении слова в последовательности – Positional Encoding. Когда эти данные достигают слоя Multi-head attention, каждый входной вектор, посредством механизма внимания, может взаимодействовать с другими параллельно обрабатываемыми словами. Использование механизма внимания в таком подходе позволяет нам отойти от передачи внутреннего состояния как это было у RNN и LSTM.

Об устройстве блока Multi-head attention подробно можно прочитать в статье [17], можно отметить, что на входе ему поступает эмбединг обрабатываемого в данный момент слова (query), а также ряд пар key и value (которые всегда равны), и представляющие собой эмбединги слов, на которые смотрит конкретный блок посредством механизма внимания. Далее каждый из этих векторов проходит через обучаемый линейный слой, после чего вычисляется скалярное произведение query со всеми key, результат проходит через softmax и полученные веса суммируются с векторами value, после чего на выходе получается единый вектор. Таким образом мы можем использовать механизм внимания не только при декодировании, но и во время работы кодировщика, сеть может видеть, что происходит при кодировке других слов что делает кодировку конкретного слова более контекстуально осмысленной.

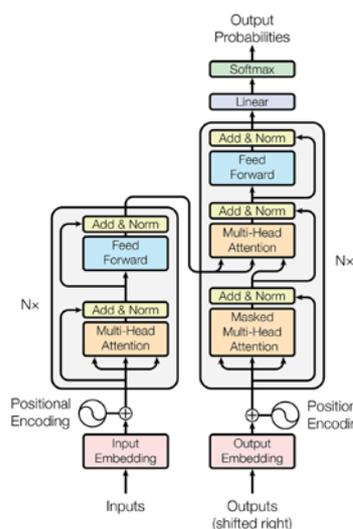


Рис. 2. Архитектура модели Transformer

Декодер работает также по одному слову, на вход он получает прошлое сгенерированное слово (или токен <start>, если процесс находится на первой итерации). В отличие от кодировщика, у декодера есть 2 блока Multi-head attention – первый способен обращаться к векторам уже декодированных слов, а второй – к выходу

кодировщика, где query представляет собой вектор входа в декодере, а пары Key/Value – финальные эмбединги кодировщика, где key равен value, но линейные преобразования внутри модуля внимания для них разные. В конце сети находится softmax, выдающий вероятности появления слов на выходе сети. По его результату выдается следующее слово в предложении, процесс повторяется пока не будет выдан токен <end of sentence>.

Данная архитектура полностью дифференцируема, что позволяет ее оптимизировать во время градиентного спуска. В предложенной в статье архитектуре глубина кодировщика составляет 8 блоков внимания, а декодера – 6. Основное различие механизмов работы кодировщика и декодера заключается в том, что во время кодирования каждый вектор взаимодействует со всеми другими, а во время декодинга каждое следующее слово взаимодействует с предыдущими и с векторами кодировщика.

Данный подход стал SOTA решением в домене машинного перевода, на рисунке 5 можно увидеть зависимость метрики BLEU [5] от архитектуры нейронной сети-переводчика (тесты проводились для перевода с английского языка на немецкий).

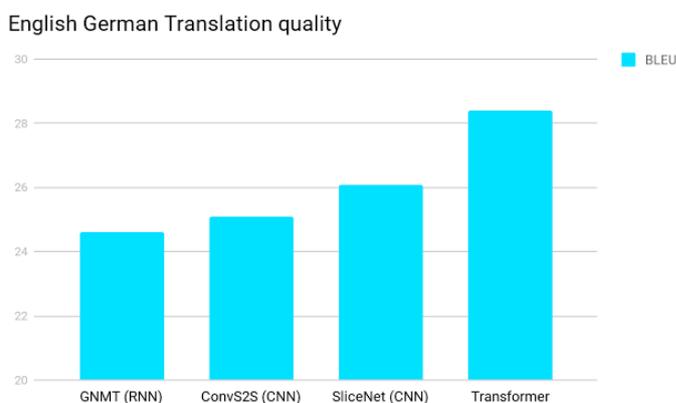


Рис. 3. Превосходство архитектуры Transformer над другими в задаче машинного перевода

Архитектура BERT. Архитектура BERT (Bidirectional Encoder Representations from Transformers – «двунаправленные презентации кодировщика для трансформеров») была представлена компанией Google в статье [18]. Она стала SOTA решением на момент выхода для ряда задач NLP, включая генерацию ответов на вопросы, формирование рассуждений и др. Модель была натренирована на большом количестве текста и была представлена в ряде конфигураций – максимальная из которых состояла из 340 млн. параметров и 24 блоков трансформеров.

Подход к тренировке выделялся тем, что если все предыдущие исследования, связанные с тренировкой языковых моделей, рассматривали текст как однонаправленную последовательность либо только слева направо, либо сочетали обучение слева направо и справа налево, то в данном случае было предложено двунаправленное обучение способное достичь более глубокого понимания языкового контекста и потока, чем модели, натренированные однонаправленно. В статье [18] была описана техника под названием MLM (Masked Language), позволяющая проводить двунаправленное обучение в моделях, для которых ранее это было невозможно.

Еще одним подходом к тренировке BERT является NSP (прогнозирование следующего предложения). При данном подходе в качестве входа модель получает пары фраз и учится предсказывать, является ли вторая фраза продолжением первой. Входные данные разбиты на пары таким образом, что половина пар действительно связана между собой, а другая половина заполняется случайным образом.

Одной из особенностей архитектуры BERT является ее способность к тонкой и точной настройке. Так, заменив последний линейный слой, можно поставить перед моделью задачу классификации, генерации текста и другие. Таким образом были натренированы вариации BERT, способные нативно работать с русским языком [19]. Также возможно натренировать BERT под задачу классификации, на официальной странице BERT на Hugging Face [20] существует специально заготовленный набор классов для гибкой настройки последнего линейного классифицирующего слоя.

Архитектура BERT была прорывом в области машинного обучения в задачах NLP, особенно благодаря возможности проводить с ней transfer learning и универсальности. Модель демонстрировала SOTA результаты во многих доменах NLP до появления архитектуры GPT, о которой пойдет речь дальше.

Generative Pretrained Transformer (GPT). Архитектура GPT была впервые представлена компанией OpenAI в 2018 году в статье [21]. Эта архитектура была также основана на трансформерах, но если до этого все модели использовали подход «обучение с учителем», то для GPT проводили обучение с частичным привлечением учителя. Обучение GPT проходило в 2 этапа: сначала проводилось генеративное обучение без учителя на большом неразмеченном текстовом корпусе, а после – тонкая настройка под конкретную задачу на сравнительно небольшом размеченном наборе данных.

OpenAI решили, что от трансформера нужно оставить только декодер, потому архитектура модели состояла из 12-слойного декодера трансформера с 12-ю слоями masked head self-attention (то есть при обучении использовался подход предсказания замаскированного слова в последовательности, или MLM). Модель имела 117 млн. параметров, еще одним из улучшений в процессе тренировки стало использование оптимизатора Adam [22] вместо применяемого ранее стохастического градиентного спуска. Также стоит отметить, что на протяжении всей тренировки (как генеративного обучения без учителя, так и при тонкой настройке с учителем) модель работала не напрямую с текстом, а с его числовым представлением в виде токенов. О процедуре токенизации речь пойдет в следующем подразделе. В качестве данных для обучения использовался набор данных BookCorpus [23], содержащий в себе текст из 7000 книг.

Первая версия модели показала себя плохо в задачах генерации текста, однако показала ряд SOTA результатов на задачах NLP, под которые была проведена процедура тонкой настройки модели, что показало высокое качество unsupervised language modeling pretraining в области построения эмбедингов, способных хорошо работать после донастройки под конкретные задачи.

В 2019 году OpenAI выпустили обновленную модель GPT-2 [24], которая имела уже 48 слоев и 1.5 млрд. параметров. Также помимо книг в качестве данных для тренировки использовались данные с 8 млн. сайтов, суммарно вышло около 40 Гб текста. Подобное количественное улучшение принесло свои плоды – модель улучшила свои результаты на задачах, под которые была произведена тонкая настройка и значительно лучше стала справляться с задачами генерации текста: модель смогла писать длинные связные тексты и решать ряд новых задач при помощи prompt engineering. Также увеличился размер запоминаемого моделью контекста – если GPT-1 запоминала только 512 токенов, то для GPT-2 это значение выросло до 1024.

Архитектура GPT показывает хорошие результаты при использовании после тонкой и точной настройки. Так, исследователи из компании Сбер провели процедуру тонкой настройки модели GPT-2, натренировав ее на текстовом корпусе на русском языке [25], и полученная модель продемонстрировала высокие результаты, наибольшая по размеру вариация модели легла в основу сервиса GigaChat.

Также настроенная под русский язык модель GPT показала превосходство над русскоязычной моделью BERT в задаче анализа тональности текстов разных размеров (покупательские отзывы и новостные статьи). Модель RuGPT_small показала значение усредненной F1-меры для отзывов, равное 0.721, в то время как RuBERT – 0.647 и 0.726 для новостей в то время, как RuBERT показала 0.582 [26]. Это говорит о более качественной работе GPT на больших текстовых последовательностях и более качественных эмбедингах, получаемых при обработке входных последовательностей. Процесс тонкой настройки в проводимом исследовании заключался в замене линейного классифицирующего слоя новым, имеющим 3 выхода (по числу классов), заморозке предыдущих слоев модели и тренировке нового слоя, что дает возможность судить о качестве эмбедингов слов, которыми оперирует модель при принятии решения.

Токенизация текста. Токенизация – процедура преобразования текста в последовательность чисел. Она необходима ввиду того, что модели машинного обучения справляются лучше с числовыми последовательностями, нежели текстовыми.

Существует множество методов токенизации – от присваивания каждому слову конкретного номера, до построения специальных токенизирующих моделей машинного обучения. Коротко рассмотрим наиболее применяемые методы токенизации на подслова: WordPiece, ULM, BPE, BBPE, BPE-Dropout. Стоит отметить, что процесс токенизации не является ресурсозатратным и может выполняться на CPU. Токенизация на подслова – попытка усреднения токенизации на уровне слов (имеющей недостаток в виде огромного размера словаря) и токенизации на уровне символов, ведущей к потере семантического значения слова. Также мы решаем проблему OOV, поскольку модель будет работать не с представлениями слов, а с представлениями их частей.

BPE-токенизация (byte pair encoding) – восходящий алгоритм токенизации подслов, который изучает словарь подслов определенного размера (размер словаря является гиперпараметром). Идея BPE-токенизации заключается в том, что более частым словам следует присваивать уникальные идентификаторы, тогда как менее частые слова следует разбивать на единицы подслов, которые лучше всего сохраняют свое значение [27]. Подробно алгоритм описан в оригинальной статье [28]. Данный вид токенизации используется в моделях GPT от OpenAI.

BBPE-токенизация (Byte-level Byte Pair Encoding) представляет собой модификацию BPE-токенизации, работающую не с текстом напрямую, а с его байтовым представлением. Подробно метод был описан в статье [29]. Данный метод используется в модели GPT-2 от OpenAI и в семействе натренированных моделей на русском языке RuGPT. Размер словаря для большой и средней модели GPT-2 составляет 50257 токенов.

WordPiece представляет собой метод, близкий к BPE-токенизации, к котором при слиянии токенов оптимизируется не частота совместного появления, а правдоподобие подобных появлений. Был предложен в статье [30] и использовался в архитектуре BERT. Популярность метода сошла после того, как авторы RoBERT показали отсутствие существенных преимуществ данного метода перед BPE.

ULM представляет собой подход к задаче токенизации, в котором предполагается независимость подслов и то что вероятность $p(x)$ слова x равна произведению вероятностей входящих слов $p(x_i)$. Далее ищется подходящая сегментация и на ее основе составляется словарь токенов. Создается большой словарь любым алгоритмом (от простого перебора символов и часто встречающихся подслов со словами до BPE словаря). После этого, до тех пор, пока размер словаря не достигнет порога, словарь фиксируется, с помощью алгоритма EM [31] максимизируется $p(x)$, для каждого

подслова вычисляется функция потерь (уменьшение правдоподобия при удалении из словаря), удаляется 20% слов с наименьшим значением функции потерь (наименее важные слова). Символы из словаря не удаляются, так как полное удаление может привести к проблеме OOV. Подробно алгоритм описан в оригинальной статье [32].

ВРЕ-Dropout является методом-комбинацией классического ВРЕ и ULM. Метод был направлен на устранение таких проблем ВРЕ, как однозначная сегментация каждого слова и слабая интерпретируемость подслов редких слов. Предлагается использовать словарь и таблицу слияний ВРЕ и на каждом шаге слияния случайно его пропускать. В таком случае вероятность пропуска слияния p является настраиваемым гиперпараметром: если $p = 0$, то получаем классический алгоритм ВРЕ, если $p = 1$, то получаем сегментацию по буквам. Подробнее алгоритм описан в оригинальной статье команды Яндекса [33]. По мнению авторов, оптимальным значением вероятности p является 0.1. В практических задачах метод продемонстрировал результаты, превосходящие результаты классических ВРЕ и ULM.

Подходы к обучению генеративных LLM. Для построения генеративных LLM обычно используется один из следующих методов: построение LLM с нуля (подход использовался компанией Bloomberg для построения модели BloombergGPT [34]), точная настройка существующей LLM на собственных данных (метод был использован компанией Google для создания модели, обрабатывающей медицинские данные, Med-PaLM2 [35]), настройка модели через расширение контекста prompt-запросы (модель остается неизменной, собственные данные подгружаются через prompt). Далее рассмотрим эти методы в целом и в контексте решаемой задачи.

Вариант с созданием LLM с нуля подходит только для тех, у кого есть огромные вычислительные ресурсы и текстовые корпуса высокого качества. Необходимы миллиарды слов и кластеры GPU чтобы качественно реализовать данный подход. Так, при создании BloombergGPT использовалась финансовая документация за 40 лет, а также большой текстовый корпус из сети Интернет. В общей сложности было задействовано 350 млрд. слов, полученная модель имела 50 млрд. параметров, а на процесс тренировки модели было потрачено 1.3 млн. часов обработки графических единиц [34]. Большой проблемой подхода является и тот факт, что необходимо проводить не только предметно-ориентированную тренировку, но и обучение модели основам языка, что также занимает множество часов обработки графических единиц. Поэтому в рамках решаемой задачи подобный подход не является оптимальным.

Подход с расширением контекста модели путем prompt-engineering не требует вычислительных мощностей, предполагая неизменность весов и архитектуры модели, что полностью исключает необходимость использования GPU. Также этот подход не требует наличия больших текстовых корпусов. Он был применен компанией Morgan Stanley, добавившей к модели GPT-4 данные со 100 000 документов на тему инвестирования. Проблема ограниченности размера контекстного окна решалась путем внедрения векторного сложения (с эмбедами имеющихся данных, полученными от другой модели). При вводе пользователем запроса в систему, алгоритм сходства определяет, какие векторы должны быть отправлены модели, тем самым экономя место в контекстном окне модели, оставляя в нем только необходимые данные. Данный подход называется RAG (Retrieval-Augmented Generation) [36]. И хотя он не является ресурсозатратным, качество ответов модели может быть менее предметно-ориентированным чем ожидается, возможны проблемы галлюцинаций модели (особенно при передаче кардинально не знакомого ей контекста, который к тому же может быть неверно представлен внешней векторизирующей моделью) и в целом

качество генераций может быть довольно низким. Поэтому данный подход также не будет использоваться в качестве основного при решении поставленной задачи, однако все еще может быть полезен для улучшения качества работы генеративной модели.

Оптимальным между двумя рассмотренными выше подходами является точная настройка существующей LLM на имеющихся данных. Данный метод не требует огромных вычислительных ресурсов и текстовых корпусов данных (достаточно иметь от сотен тысяч до миллионов слов вместо миллиардов, как это необходимо для полноценного создания LLM с нуля). Так, модель Med-PaLM2 от Google, обученная подобным образом, смогла ответить на 85% вопросов для получения медицинской лицензии США. Схожий метод был использован и при адаптации моделей BERT и GPT под русский язык, модели семейства RuGPT показывают хорошие результаты в практическом применении. Также удобство использования данного метода заключается в том факте, что ряд крупных LLM (таких как GPT-2, LLAMA и другие) являются OpenSource решениями, что позволяет оптимизировать процесс дополнительного обучения модели на собственных данных.

Последний рассмотренный подход является наиболее подходящим для решения поставленной задачи, поскольку его реализация сочетает в себе как практическую значимость – ответы полученной модели будут предметно-ориентированными, так и исследовательскую, поскольку в реализации подхода будет задействован полный цикл обучения LLM. При этом для реализации точной настройки модели не нужно располагать огромными вычислительными ресурсами и текстовыми корпусами данных, что делает процесс реализуемым без привлечения сверхмощных ПК и кластеров GPU. Также возможно совершенствование работы модели путем комбинации данного подхода и механизма RAG, что позволит актуализировать базу знаний модели без необходимости в постоянном дообучении, а также позволит улучшить качество генерации.

Выводы. По результатам рассмотрения и анализа методов построения и способов обучения больших генеративных языковых моделей можно сказать, что в качестве основы архитектуры будущей LLM наиболее подходящим выбором будет GPT, поскольку данная архитектура является SOTA решением во многих задачах как NLP в целом, так и генерации текста в частности. Кроме того, данная архитектура показывает хорошую способность к дообучению на собственных данных, что было продемонстрировано в соответствующей секции.

В качестве метода обучения LLM, оптимальным выбором является процедура точной настройки существующей LLM (в конкретном случае – основанной на архитектуре GPT), поскольку данный подход сочетает в себе как высокие понимание моделью предметной области и качество генерации ответов, так и более низкую ресурсозатратность процесса обучения, относительно полной тренировки LLM с нуля. Также возможно улучшение качества генераций полученной в результате точной настройки модели с помощью механизма RAG.

Поэтому для создания LLM, управляющей интеллектуальным ассистентом образовательной организации, будет использовано точное обучение генеративной LLM, основанной на архитектуре GPT. В качестве тренировочного набора данных будет использоваться внутренняя документация образовательного учреждения, а также данные о нем как с официальных, так и с ряда открытых источников в сети Интернет. Далее, полученная модель будет работать с обновляемой базой знаний, дополняя свои ответы ее содержимым с помощью RAG.

СПИСОК ИСТОЧНИКОВ

1. Методические рекомендации представления информации об образовательной организации высшего образования в открытых источниках с учетом соблюдения требований законодательства в сфере образования (вступают в силу с 01.09.2024). – Федеральная служба по надзору в сфере образования и науки. – 2024. – 101 с.
2. Бондаренко, В. И. Подготовка данных и выбор способа обучения модели интеллектуального ассистента для образовательной организации / В. И. Бондаренко, В. О. Елисеев // Донецкие чтения 2024: образование, наука, инновации, культура и вызовы современности: Материалы IX Международной научной конференции (Донецк, 15–17 октября 2024 г.). – Том 2: Физические, химические, технические и компьютерные науки. Часть 2. – Донецк: Изд-во ДонГУ, 2024. – С. 127-130.
3. Deep Learning и генеративные модели для создания процедурных текстов [Электронный ресурс]. – URL: https://aiplayzone.ru/tehnologii-i-metody/deep_learning_i_generativnye_modeli_dlja_sozdanija_protседurnyh_tekstov/ (дата обращения 08.02.2024).
4. Wang, L [et al.] Text embeddings by weakly-supervised contrastive pre-training. – 2022. – URL: <https://arxiv.org/abs/2212.03533>.
5. Reiter, E. A structured review of the validity of BLEU / E. A. Reiter // Computational Linguistics. – 2018. – V. 44. – № 3. – P. 393-401.
6. Lin, C. Y. Rouge: A package for automatic evaluation of summaries / C. Y. Lin // Text summarization branches out. – 2004. – P. 74-81.
7. Miaschi, A. [et al.] Is Neural Language Model Perplexity Related to Readability? / Computational Linguistics CLiC-it 2020. – 2020. – P. 303.
8. Casper, S [et al.] Open problems and fundamental limitations of reinforcement learning from human feedback. – 2023. – URL: <https://arxiv.org/abs/2307.15217>.
9. Mikolov, T. [et al.] Efficient estimation of word representations in vector space. – 2013. – URL: <https://arxiv.org/abs/1301.3781>.
10. Goldberg, Y. [et al.] word2vec Explained: deriving Mikolov et al.'s negative-sampling word-embedding method. – 2014. – URL: <https://arxiv.org/abs/1402.3722>.
11. Joulin, A. [et al.] Bag of tricks for efficient text classification. – 2016. – URL: <https://arxiv.org/abs/1607.01759>.
12. Word vectors for 157 languages [Электронный ресурс]. – URL: <https://fasttext.cc/docs/en/crawl-vectors.html> (дата обращения 11.02.2024).
13. FacebookResearch. FastText [Электронный ресурс]. – URL: <https://github.com/facebookresearch/fastText> (дата обращения 11.02.2024).
14. Staudemeyer, R. C. [et al.] Understanding LSTM – a tutorial into long short-term memory recurrent neural networks. – 2019. – URL: <https://arxiv.org/abs/1607.01759>.
15. Sutskever, I. [et al.] Sequence to sequence learning with neural networks. – 2014. – URL: <https://arxiv.org/abs/1409.3215>.
16. Bahdanau, D. [et al.] Neural machine translation by jointly learning to align and translate. – 2014. – URL: <https://arxiv.org/abs/1409.0473>.
17. Vaswani, A. [et al.] Attention is all you need. – 2017. – URL: <https://arxiv.org/abs/1706.03762>.
18. Devlin, J. [et al.] Bert: Pre-training of deep bidirectional transformers for language understanding. – 2018. – URL: <https://arxiv.org/abs/1810.04805>.
19. Kuratov, Y. [et al.] Adaptation of deep bidirectional multilingual transformers for Russian language. – 2019. – URL: <https://arxiv.org/abs/1905.07213>.
20. BERT [Электронный ресурс]. – URL: https://huggingface.co/docs/transformers/model_doc/bert (дата обращения 14.02.2024).
21. Radford, A. [et al.] Improving language understanding by generative pre-training / Computer Science, Linguistics. – 2018. – 12 p.
22. Kingma, D. P. [et al.] Adam: A method for stochastic optimization. – 2014. – URL: <https://arxiv.org/abs/1409.0473>.
23. Bookcorpus [Электронный ресурс]. – URL: <https://huggingface.co/datasets/bookcorpus> (дата обращения 15.02.2024).
24. Radford, A. [et al.] Language models are unsupervised multitask learners // OpenAI blog. – 2019. – V. 1. – № 8. – P. 9.
25. Zmitrovich, D. [et al.] A family of pretrained transformer language models for Russian. – 2023. – URL: <https://arxiv.org/abs/2309.10931>.

26. Бондаренко, В. И. Анализ эффективности глубоких языковых моделей для задачи определения тональности русскоязычных текстов / В. И. Бондаренко, В. О. Елисеев, Т. В. Ермоленко // Проблемы искусственного интеллекта. – 2024. – № 1(32). – С. 51-62. – EDN KJRRQW.
27. Пикалев, Я. С. Адаптация нейросетевой модели ALBERT для задачи языкового моделирования / Я. С. Пикалев, Т. В. Ермоленко // Проблемы искусственного интеллекта. – 2020. – № 3(18). – С. 111-122. – EDN INNNNX.
28. Sennrich, R. [et al.] Neural machine translation of rare words with subword units. – 2015. – URL: <https://arxiv.org/abs/1508.07909>.
29. Wang, C. [et al.] Neural machine translation with byte-level subwords. – 2020. – URL: <https://arxiv.org/abs/1909.03341>.
30. Schuster, M. Japanese and korean voice search / M. Schuster, K. Nakajima // 2012 IEEE international conference on acoustics, speech and signal processing (ICASSP). – IEEE, 2012. – P. 5149-5152.
31. Dellaert, F. The expectation maximization algorithm: Technical Report number GIT-GVU-02-20 / F. Dellaert // College of Computing, Georgia Institute of Technology, 2002. – 7 p.
32. Kudo, T. Subword regularization: Improving neural network translation models with multiple subword candidates. – 2018. – URL: <https://arxiv.org/abs/1804.10959>.
33. Provilkov, I. [et al.] BPE-dropout: Simple and effective subword regularization. – 2019 – URL: <https://arxiv.org/abs/1910.13267>.
34. Wu, S. [et al.] Bloomberggpt: A large language model for finance. – 2023. – URL: <https://arxiv.org/abs/2303.17564>.
35. Singhal, K. [et al.] Towards expert-level medical question answering with large language models. – 2023. – URL: <https://arxiv.org/abs/2305.09617>.
36. Lewis, P. [et al.] Retrieval-augmented generation for knowledge-intensive nlp tasks / Advances in Neural Information Processing Systems. – 2020. – V. 33. – P. 9459-9474.

Поступила в редакцию 13.09.2024 г., рекомендована к печати 04.10.2024 г.

METHODS OF DESIGNING THE ARCHITECTURE OF AN INTELLIGENT ASSISTANT MODEL FOR AN EDUCATIONAL ORGANIZATION

Eliseev V.O., Bondarenko V.I.

The vast majority of powerful modern LLM's consist of encoder, decoder and attention mechanism. This paper reviews the main approaches to designing each of these components, as well as the most popular LLM architectures. Additionally, a comparative analysis of training methods and applications of large generative language models was conducted to evaluate their optimal use in developing an intelligent assistant for an educational organization. The results of the study indicate that, given the available resources, the most suitable approach is to fine-tune an existing LLM on a textual corpus of data related to the educational organization and subsequently use it either directly or through the RAG mechanism.

Keywords: text generation, LLM, GPT, fine-tuning, chatbot.

Елисеев Вадим Олегович

стажер-исследователь
интеллектуальных систем ФГБНУ «Институт
прикладной математики и механики»,
Российская Федерация, ДНР, Донецк
E-mail: eliseevv02@mail.ru

лаборатории

Eliseev Vadim Olegovich

Intern Researcher at AI Systems Laboratory of
Institute of Applied Mathematics and Mechanics,
Russian Federation, DPR, Donetsk.

Бондаренко Виталий Иванович

кандидат технических наук, доцент, доцент
кафедры компьютерных технологий ФГБОУ ВО
«Донецкий государственный университет»,
Российская Федерация, ДНР, Донецк
E-mail: bondarenko@donnu.ru

Bondarenko Vitalii Ivanovich

Candidate of Technical Sciences, Associate
Professor, Associate Professor at Department of
Computer Technologies of Donetsk State University,
Russian Federation, DPR, Donetsk.

РАСПОЗНАВАНИЕ НЕДОСТОВЕРНОЙ ИНФОРМАЦИИ В СМИ С ПОМОЩЬЮ НЕЙРОННЫХ СЕТЕЙ

© 2024 *Криворучко К.А., Максименко И.И.*

В работе представлен процесс создания классификатора недостоверных новостей с использованием нейронной сети на основе предобученной модели BERT (bert-base-uncased). Проведен анализ набора данных новостных статей с платформы Kaggle. Осуществлена предварительная обработка данных и исследовательский анализ текстов, выявлены характерные особенности достоверных и недостоверных новостей. Модель BERT была дообучена, оптимизирована и модифицирована, что позволило достичь точности классификации 98.3%. Результаты показывают эффективность предложенного подхода для выявления недостоверной информации в СМИ.

Ключевые слова: нейронные сети, BERT, классификация текста, дезинформация, трансформеры, машинное обучение, анализ данных.

Введение. С увеличением объема доступной нам информации и ускорением темпов её распространения в современном медиапространстве возникает значительная проблема, связанная с распространением недостоверных сведений. В последние десятилетия наблюдается бурный рост влияния средств массовой информации и социальных платформ на общество, что привело к расцвету феномена "фейковых новостей" и дезинформации. Эффективное распознавание и фильтрация недостоверной информации становятся критически важными для поддержания общественного доверия к медийным источникам и обеспечения их социальной ответственности. Недостоверные сведения могут иметь серьезные последствия, начиная от манипуляции общественным мнением до угрозы национальной безопасности. Поэтому борьба с дезинформацией требует комплексных подходов, включающих как технические решения, так и междисциплинарные исследования. В этом контексте нейронные сети представляют собой перспективный инструмент, обладающий потенциалом для анализа и классификации информации на основе её содержания, структуры и других характеристик.

Основными целями данной работы является:

- 1) Анализ набора данных. Для достижения этой цели необходимо провести детальный анализ набора данных, выбранных для исследования, с целью оценки его репрезентативности и применимости в контексте задачи классификации информации.
- 2) Анализ выбранной модели для создания классификатора на основе нейронных сетей. В рамках этой цели необходимо провести выбор архитектуры нейронной сети, реализацию алгоритмов обучения и оптимизации, а также тестирование и оценку эффективности модели с использованием соответствующих метрик.

Использование нейронных сетей для распознавания недостоверной информации представляется обоснованным выбором по нескольким причинам. Во-первых, нейронные сети обладают высокой способностью к выявлению сложных паттернов в текстовых данных, что является критически важным для задач, связанных с анализом естественного языка.

Работа выполнена в рамках Государственного задания ФГБНУ ИПММ, тема FREM-2023-0003 «Разработка теоретических и практических основ создания и использования цифровых двойников технологических и информационных процессов».

Во-вторых, архитектура трансформер (включает механизм внимания) представляет собой одну из самых передовых архитектур нейронных сетей для решения задач обработки естественного языка [1], что делает их особенно подходящими для обработки информации, где необходимо учитывать контекст и нюансы языка.

Результаты работы могут найти применения в исследовании цифровых двойников систем защиты от спама поисковых систем и возможностей генерации фейковых текстов большими языковыми моделями (ChatGPT).

Анализ набора данных. Для анализа был использован набор данных с платформы Kaggle [2], включающий тексты новостных статей, помеченные как правдивые или поддельные. Набор содержит 5 полей, содержащих информацию о названии новостной статьи, авторах, тексте статьи и метку, которая характеризует статью как недостоверную или достоверную. Набор данных содержит 20800 записей.

Первоначально было проведено исследование статистических характеристик столбцов данных с информацией о заголовках и тексте новостей. В результате исследования было выявлено, что атрибут с текстом статьи имеет большее количество слов, а именно в среднем 760 слов, при этом 75% статей содержат более 1000 слов. Максимальная длина текста составляет 24 234 символа, что значительно превышает среднюю длину. Эта аномалия указывает на наличие очень длинных текстов, которые могут быть не типичными для данного набора данных, поэтому целесообразно использовать современные архитектуры, такие как BERT, способные автоматически обрабатывать текст разной длины. Атрибут с заголовком статьи является коротким утверждением, состоящим в среднем из 12 слов, при этом 75% строк содержат 15 слов. Наше исследование будет проводиться как с основным текстом статей, так и с их заголовками вместе.

Был проведен анализ распределения классов, в результате которого, было выявлено, что количество недостоверных статей равно 10413, в то время как количество достоверных равно 10387. Таким образом, количество статей, помеченных как поддельные, и количество статей, помеченных как правдивые, практически равны. Этот факт критически важен, поскольку сбалансированные классы позволяют избежать смещения модели в сторону большинства, что в свою очередь повышает её способность эффективно классифицировать как недостоверные, так и достоверные новости. Нижеприведенный график (рис. 1) демонстрирует это.

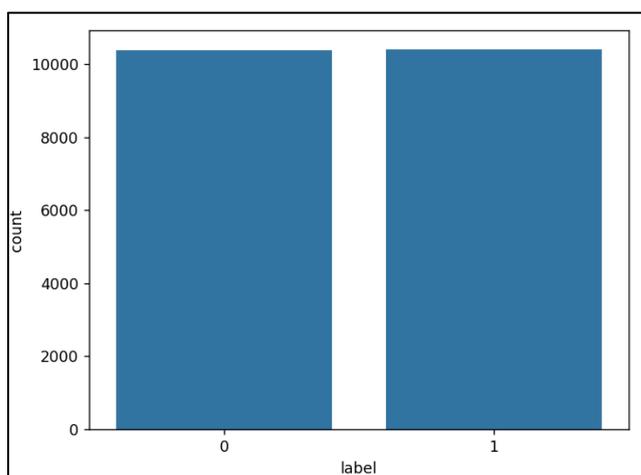


Рис. 1. График распределения новостей по классам

Очистка данных. Для обеспечения корректности и эффективности последующего анализа и моделирования была проведена предварительная очистка данных. Этот процесс включает несколько ключевых этапов:

1) Удаление неиспользуемых строк и столбцов. Были исключены строки и столбцы, которые не содержали полезной информации или были заполнены недействительными данными. Этот шаг необходим для устранения потенциального шума, негативно влияющего на качество обучения модели. Удаление лишних данных также способствует уменьшению вычислительной сложности и времени обработки.

2) Удаление нулевых значений. Все ячейки, содержащие нулевые значения, были исключены из анализа. Нулевые значения могут исказить статистические выводы и влиять на точность модели, особенно в контексте анализа текстовых данных, где полные и целостные записи имеют критическое значение для обучения.

3) Удаление специальных символов. Специальные символы, такие как пунктуация и математические знаки, были удалены из текстов. Эти символы не несут значимой смысловой нагрузки для задач классификации текста и могут вносить шум в модель, приводя к снижению её точности.

4) Удаление стоп-слов. Стоп-слова, представляющие собой часто встречающиеся, но малоинформативные слова, были исключены из текстов. Эти слова не способствуют различению содержания текстов и могут уменьшить эффективность обработки текста, усложняя выявление ключевых характеристик и паттернов.

5) Стемминг. Это упрощенная форма морфологического анализа, представляет собой эвристический процесс извлечения базовой формы слов путем отсечения концов слов [3]. Стемминг помогает стандартизировать текст, что может улучшить результаты модели путем сокращения вариативности слов.

6) Лемматизация. В дополнение к стеммингу была использована лемматизация. Это процесс группировки различных флексивных форм слова, чтобы их можно было анализировать как единый элемент. Лемматизация похожа на стемминг, но она привносит контекст в слова. Таким образом, она связывает слова со схожими значениями в одно слово [4]. Этот подход помогает сохранить смысл слов и улучшить качество классификации, обеспечивая точность и полное представление текста.

Каждый из этих шагов направлен на очистку и улучшение качества данных для последующего анализа и построения моделей машинного обучения.

Исследовательский анализ данных. На первом этапе был проведен одномерный анализ данных, в котором основное внимание было уделено выявлению наиболее частотных слов и их визуализации в виде облака слов. Результаты построения облака слов представлены на рисунках 2 и 3.

На основе двух представленных облаков слов можно провести следующий сравнительный анализ. В достоверных новостях часто упоминаются: Trump, said, one, state, people, United, work, company. В недостоверных новостях часто встречаются слова: people, one, said, country, Hillary Clinton, Trump, U.S., nation. В достоверных новостях чаще встречаются официальные названия, такие как United States (Соединенные Штаты), что говорит о более формальном стиле подачи информации. Достоверные новости сосредоточены на действиях и фактах (много слов, связанных с действием: said, work, govern). Недостоверные новости часто оперируют словами, которые создают ощущение общности или эмоционального воздействия: people, nation, country.

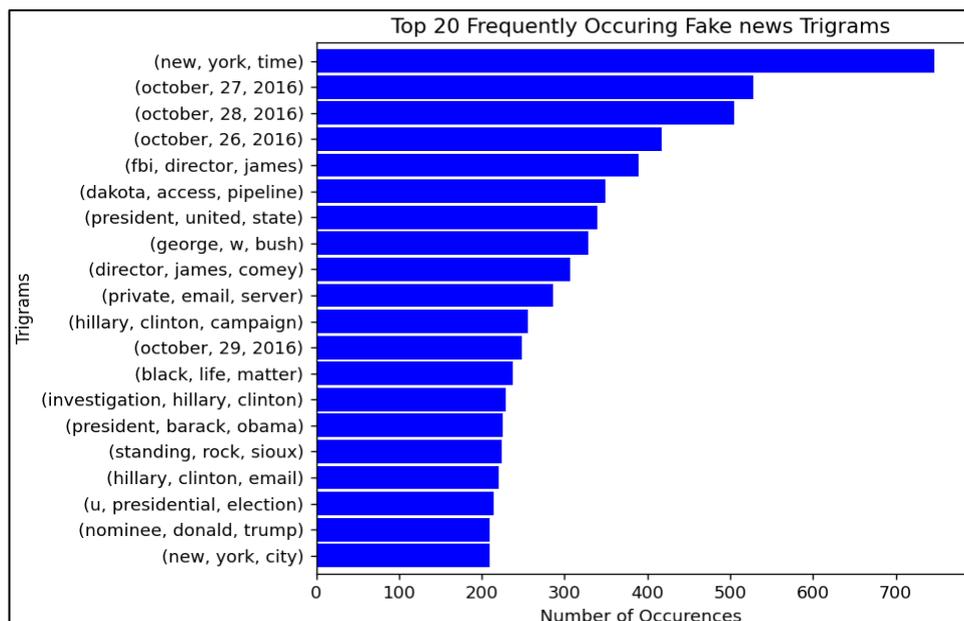


Рис. 5. График триграммы недостоверных новостей

В результате анализа данных триграмм, можно сделать следующие выводы:

1. В недостоверных новостях часто встречаются триграммы, связанные с конкретными политическими событиями и личностями, например, "fbi, director, james", "director, james, comey", "hillary, clinton, email". Наличие триграмм на тему демонстраций и протестов ("black, life, matter", "standing, rock, sioux") может свидетельствовать о специфической политической повестке или интерпретации событий, используемой в недостоверных новостях.

2. В достоверных новостях триграммы часто связаны с федеральными агентствами и законами ("department, homeland, security", "law, enforcement, official", "affordable, care, act") и указывают на наличие информации, связанной с госструктурами и законодательными актами. Триграммы на исторические и международные темы ("world, war, ii" или "george, w, bush") также могут говорить о достоверности, так как в фальшивых новостях такие темы менее распространены. Также стоит отметить, что в достоверных новостях часто используются полные имена и титулы, такие как "president, donald, trump" или "george, w, bush", что может указывать на формальный тон и достоверность. В недостоверных новостях часто встречаются фразы, направленные на эмоциональную реакцию, такие как "black, life, matter" или "standing, rock, sioux", что может указывать на попытку манипуляции мнением.

Выбор модели нейронной сети для создания классификатора. Для классификации использовалась предобученная модель BERT из библиотеки Transformers, в частности, модель bert-base-uncased. Выбор этой модели был обусловлен рядом факторов:

1) Размер. Bert-base-uncased состоит из 12 слоёв (трансформеров), 768 скрытых единиц в каждом слое и 12 голов для механизма внимания (attention heads) [5], что делает её достаточно компактной для выполнения на обычных вычислительных ресурсах, но при этом мощной для обработки сложных языковых задач.

2) Регистр. В данной модели не учитывается регистр символов, что может быть полезным для задач, связанных с текстами новостей, где использование верхнего и нижнего регистра может быть несистематическим или не иметь значимого влияния на достоверность содержания.

3) Двухнаправленность. Как уже упоминалось ранее, некоторые новости содержат достаточно длинные абзацы текста с несколькими взаимосвязанными утверждениями, и правильная оценка достоверности требует анализа сложных зависимостей между предложениями. Традиционные методы обработки текста, такие как N-граммы или подходы на основе мешка слов, не учитывают эти зависимости, тогда как BERT, благодаря своей двухнаправленной природе и механизму внимания, может строить более сложные и точные представления о тексте.

4) Трансферное обучение. Данная модель была предварительно обучена на большом количестве текстов общего характера (например, Wikipedia и BookCorpus), что позволяет адаптировать её к специализированным задачам с относительно небольшими размерами датасетов, как в нашем случае.

Настройка модели. Модель была дообучена на основе подготовленного набора данных. Процесс дообучения включал настройку параметров модели, таких как количество эпох, размер батча, скорость обучения и другие параметры. Основные параметры настройки и их значения:

1) Количество эпох обучения установлено на 1, что обусловлено тем, что предобученная модель уже обладает обширными знаниями, поэтому ей не требуется большое количество эпох для дообучения на специфичной задаче классификации.

2) Размер батча влияет на стабильность обучения и вычислительные ресурсы. Значение 20 для обучения и 40 для валидации выбрано как компромисс между использованием памяти и скоростью сходимости.

3) Warmup при обучении в глубоком обучении относится к начальной фазе, на которой скорость обучения постепенно увеличивается от небольшого значения до целевой скорости обучения, чтобы стабилизировать процесс оптимизации и предотвратить расхождения [6]. Величина `warmup_steps=100` позволяет плавно разогреть процесс обучения, особенно при использовании предобученной модели.

4) Стратегия оценки модели задана значением “step”. Данный параметр задаёт стратегию валидации модели на каждом шаге, а не по завершению эпохи. Такая настройка даёт возможность оценивать модель на валидационном наборе через фиксированные промежутки времени, что особенно полезно при ранней остановке или оптимизации модели в течение короткого времени.

Представленные параметры настройки обучения модели отражают баланс между эффективностью модели и доступными вычислительными ресурсами. Использование предобученной модели и небольшое количество эпох позволяют минимизировать переобучение и вычислительные затраты, сохраняя при этом высокую точность классификации.

Оценка модели. После дообучения была проведена оценка модели на тестовом наборе данных. Был проведен анализ полученных результатов по основным характеристикам, приведенным в таблице 1.

Таблица 1. Результаты обучения модели по эпохам

Эпоха	Loss	Eval Loss	Accuracy	Grad Norm	Learning Rate
0.27	0.4617	0.2231	0.9207	5.17	4.208e-05
0.55	0.2166	0.1389	0.9535	5.72	2.626e-05
0.82	0.124	0.0916	0.9691	1.04	1.044e-05

Исходя из полученных данных, можно сделать вывод о том, что:

1) Потери (loss) постепенно снижаются на протяжении обучения, что свидетельствует о том, что модель учится и адаптируется к данным.

2) Оценочная потеря (`eval_loss`) также уменьшается с каждой эпохой, а точность (Accuracy) классификации увеличивается. При этом модель показывает хорошую способность обобщения, так как точность классификации на валидационной выборке достигает 96.9%.

3) Значения нормы градиента (Grad Norm) показывают, насколько сильно обновляются параметры модели. Снижение нормы градиента к концу обучения говорит о том, что модель приближается к оптимуму.

4) Скорость обучения (Learning Rate) постепенно уменьшается по мере обучения модели.

Модель хорошо обучается, демонстрируя снижение потерь и увеличение точности как на обучающей, так и на валидационной выборке. Точность модели составила 96.9%, что говорит о высоком качестве классификации недостоверных новостей.

Оптимизация нейросети. В ходе настройки нейросетевой модели для задачи классификации текста было предпринято 5 этапов настройки, направленных на улучшение её производительности. На каждом этапе тестировались различные гиперпараметры и архитектурные модификации для достижения баланса между точностью и предотвращением переобучения. Рассмотрим подробно каждую попытку:

На первом этапе было принято решение изменить архитектуру модели на основе BERT. Для увеличения её способности классификации мы добавили два дополнительных линейных слоя и дропаут, а также увеличили количество эпох до 4. Dropout помогает избежать переобучения, случайным образом отключая нейроны в ходе обучения, что делает модель более устойчивой к небольшим вариациям в данных. Однако, несмотря на эти улучшения, начиная с третьей эпохи наблюдалось снижение точности, что может свидетельствовать о переобучении. Поэтому было принято решение уменьшить шаг обучения

Во втором этапе был уменьшен шаг обучения для более стабильной оптимизации, добавлен слой нормализации и увеличено количество нейронов в линейных слоях. Снижение шага обучения обычно улучшает сходимость на поздних этапах обучения. Слой нормализации может помочь улучшить обучение модели, особенно в глубоких сетях, за счет стабилизации входных данных для каждого слоя. Однако, несмотря на данные улучшения, результат оказался недостаточно хорошим, и наблюдалось снижение производительности.

Третий этап включал добавление стратегии Warmup (может помочь избежать резкого скачка градиентов в начале обучения) и регуляризации Weight Decay. Warmup позволяет модели избегать больших градиентов на ранних шагах, а Weight Decay уменьшает зависимость от отдельных весов, что должно способствовать лучшей обобщающей способности. Однако результаты ухудшились, что показало, что данный подход не подошёл для текущих данных и задачи.

На четвёртом этапе были отменены изменения с третьего этапа, а шаг обучения был увеличен. Однако это привело к ещё более значительному снижению точности. Шаг обучения напрямую влияет на сходимость. В данном случае, слишком большой шаг обучения мог привести к тому, что модель перескакивала через оптимальные точки минимизации функции потерь. Таким образом, гипотеза о том, что больший шаг обучения улучшит сходимость, оказалась неверной для данного случая.

На пятом этапе была увеличена длина обучения до 4 эпох и размер батча до 32. Также шаг обучения был увеличен, что позволило улучшить качество обучения модели. В результате точность возросла до 98.3%. В этом случае было достигнуто оптимальное соотношение между шагом обучения и количеством эпох.

Ниже (табл. 2) представлена сводная таблица произведенных изменений.

Таблица 2. Изменения и результаты во время экспериментов по оптимизации модели

Этап №	Изменения в архитектуре модели	Изменения параметров настройки	Результаты: eval_loss / eval_accuracy	Комментарии
Базовая модель	Готовая модель BERT	Epochs: 1	Eval Loss: 0.124 / Eval Accuracy: 96.9%	Модель без изменений, достигла высокой точности
Этап 1	Добавлены dropout и 2 линейных слоя, увеличено количество нейронов	Epochs: 4, LR: уменьшен до 2e-5	Eval Loss: 0.123 / Eval Accuracy: 97.8%	Улучшение точности до 98.1%, но переобучение после 3 эпох
Этап 2	Добавлен слой нормализации, увеличено количество нейронов	Epochs: 3, LR: уменьшен до 1.5e-5	Eval Loss: 0.137 / Eval Accuracy: 97%	Снижение точности, настройки не улучшили результат
Этап 3	Добавлен Warmup и Weight Decay	Epochs: 3, LR: 4e-6	Eval Loss: 0.238 / Eval Accuracy: 92.1%	Добавление weight decay ухудшило результат
Этап 4	Убраны изменения предыдущего этапа, увеличен шаг обучения	Epochs: 3, LR: 4e-5	Eval Loss: 0.286 / Eval Accuracy: 89.5%	Точность продолжила падать, LR слишком большой
Этап 5	Увеличено количество эпох до 4, увеличен batch size до 32, увеличен шаг обучения	Epochs: 4, LR: 4e-5	Epoch 1: Loss: 0.099 / Accuracy: 98.3%	Наилучший результат, максимальная точность 98.3%, стабилизация после 3-й эпохи

Выводы. В данной работе представлен процесс создания классификатора фейковых новостей на основе нейронной сети с использованием предобученной модели BERT (bert-base-uncased) и ее модификации. Анализ набора данных, состоящего из новостных статей с метками достоверности, показал сбалансированность классов и особенности текстов, что позволило применить современные архитектуры для обработки текстов различной длины. Модель была дообучена и модифицирована добавлением дополнительных слоев и экспериментальной настройкой ее параметров, что позволило достичь высокой точности классификации на валидационных данных — 98.3% вместо изначальных 96.9%. Результаты свидетельствуют о том, что предложенный подход эффективно справляется с задачей выявления недостоверной информации. Дальнейшие шаги могут включать расширение набора данных, проведение более детального анализа ошибок модели и исследование дополнительных методов улучшения классификации.

СПИСОК ИСТОЧНИКОВ

1. Узких, Г. Ю. Применение трансформеров в обработке естественного языка / Г. Ю. Узких // Вестник науки. – 2024. – Т. 4, № 8(77). – С. 186-189. – EDN GTSBKA.
2. Fake News Detection [Электронный ресурс]. – URL: <https://www.kaggle.com/competitions/fake-news/data>. (Дата обращения: 12.09.2024).
3. Учебник по обработке естественного языка [Электронный ресурс]. – URL: https://isolution.pro/ru/t/natural-language-processing?alias=ucebник-po-obrabotke-estestvennogo-azyka#google_vignette. (Дата обращения: 12.09.2024).
4. Lemmatization with NLTK [Электронный ресурс]. – URL: <https://www.geeksforgeeks.org/python-lemmatization-with-nltk/>. (Дата обращения: 12.09.2024).
5. BERT Model [Электронный ресурс]. – URL: https://huggingface.co/google/bert_uncased_L-12_H-768_A-12. (Дата обращения: 12.09.2024).
6. Training Warmup Steps in Deep Learning [Электронный ресурс]. – URL: <https://www.geeksforgeeks.org/in-the-context-of-deep-learning-what-is-training-warmup-steps/>. (Дата обращения: 12.09.2024).

Поступила в редакцию 16.09.2024 г., рекомендована к печати 08.10.2024 г.

RECOGNITION OF INCORRECT INFORMATION IN THE MEDIA USING NEURAL NETWORKS

Krivoruchko K.A., Maksimenko I.I.

The paper presents the process of creating a classifier of unreliable news using a neural network based on the pre-trained BERT (bert-base-uncased) model. The analysis of a dataset of news articles from the Kaggle platform was carried out. Preliminary data processing and exploratory text analysis were carried out, characteristic features of reliable and unreliable news were identified. The BERT model was further trained, optimized and modified, which made it possible to achieve a classification accuracy of 98.3%. The results demonstrate the effectiveness of the proposed approach for identifying unreliable information in the media.

Keywords: neural networks, BERT, text classification, disinformation, transformers, machine learning, data analysis.

Криворучко Ксения Анатольевна

магистрант кафедры компьютерных технологий
ФГБОУ ВО «Донецкий государственный
университет»,

Российская Федерация, ДНР, г. Донецк.

E-mail: alisakovalllll@gmail.com

Krivoruchko Kseniia Anatolevna

Master Student at Department of Computer
Technologies of the Donetsk State University,
Russian Federation, DPR, Donetsk.

Максименко Игорь Иванович

кандидат физико-математических наук, доцент
кафедры компьютерных технологий ФГБОУ ВО
«Донецкий государственный университет»,

Российская Федерация, ДНР, г. Донецк.

E-mail: igor.maksimenko_1967@mail.ru

Maksimenko Igor Ivanovich

Candidate of Physical and Mathematical Sciences,
Associate Professor at Department of Computer
Technologies of Donetsk State University,

Russian Federation, DPR, Donetsk.

АЛГОРИТМ КЛАССИФИКАЦИИ ИДЕОЛОГЕМ НА ОСНОВЕ НАБОРОВ ЦЕПОЧЕК РЕЛИГИОЗНО-ПОЛИТИЧЕСКИХ АССОЦИАЦИЙ

© 2024 Секерин А.В., Кудинов В.А.

В настоящей работе представлен алгоритм классификации идеологем – лексических единиц, отражающих религиозные течения и политические идеологии. Подход заключается в построении базы знаний ассоциативных обобщающих цепочек отношений между понятиями на основе векторных языковых моделей и онтологий и последующем поиске общих элементов среди цепочек, порожденных для исследуемых слов, что сводит задачу к поиску общего гиперонима. Проведенное сравнение реализации алгоритма с методом на основе запросов к большим языковым моделям показало его эффективность для классификации идеологем.

Ключевые слова: векторная языковая модель (word2vec), онтология, ассоциативная цепочка, гипероним, большая языковая модель (LLM).

Введение. Идеологемы, являясь лексическими единицами, отражающими религиозные течения и политические идеологии, характеризуют, в том числе, позиции центров влияния в рамках информационного противоборства. Вопросы процесса автоматизации идентификации и классификации таких понятий разрабатываются на стыке лингвистики, политологии и компьютерных наук. Подходы к классификации идеологем, получившие широкое распространение в отечественной лингвистике, базируются на экспертном анализе с использованием толковых и ассоциативных словарей и представлены в работах [1, 2].

В последние годы получили развитие нейросетевые алгоритмы машинного обучения с учителем. В качестве метки текста при формировании датасетов в [3, 4] выступала партийная принадлежность её авторов, что применимо для моделирования процессов, протекающих в сложившихся политических системах, где наблюдаются существенные различия между риторикой фракций, например, США и ФРГ. В данном случае классификация текстов по конечному множеству идеологий может интерпретироваться как классификация упоминаемых в тексте идеологем – токенов, повлиявших на выбор модели, что позволяет обеспечить выявление связей «идеологема – идеология» в автоматическом режиме. Другой подход основан на поиске идеологем как частного случая решения задачи извлечения именованных сущностей, при этом существующие решения обеспечивают нахождение терминов идеологических категорий без привязки к политическим идеологиям [5].

К ключевым недостаткам существующих алгоритмов относятся:

- необходимость привлечения экспертов на отдельных этапах классификации понятий, в том числе для установления их иерархии;
- сложность интерпретации результатов работы нейросетевых методов, в том числе невозможность построения логических выводов;
- невозможность автоматической классификации счетного множества идеологий, ввиду необходимости формирования набора обучающих данных, отражающих все многообразие и иерархию лексических единиц, отображающих взгляды различных социальных групп.

Целью работы является разработка алгоритма классификации понятий на основе общности их идеологических ассоциаций.

Материалы и методы. Решение обозначенных проблем может быть достигнуто путем разработки двухэтапного алгоритма, реализующего автоматическую генерацию ассоциативных идеологических обобщающих цепочек русскоязычных слов на основе векторных представлений слов и баз знаний, а также разработки метода поиска общих понятий.

Задача автоматической генерации обобщающих ассоциативных цепочек в отдельной предметной области может быть рассмотрена как частный случай решения задачи автоматического поиска пар гипонимов-гиперонимов, которая решалась в рамках соревнования RUSSE [6], при этом лучшие результаты показали алгоритм, использующий комбинацию идентификации шаблонов, анализ онтологий для иностранных языков и результаты поисковой выдачи. В работах [7, 8] предложены методы, использующие векторные представления слов и дистрибутивную семантику для выявления обобщающих понятий. В основе данного подхода адаптация лингвистической гипотезы распределения слов. Так, поскольку слова, используемые в схожих контекстах, в большинстве случаев имеют близкий смысл, следовательно расстояние между отображающими их векторами должно стремиться к нулю.

Семейство языковых нейросетевых моделей получения векторных представлений word2vec базируется на двух противоположных архитектурах, предложенных Миколовым [9]. В первой (CBOW, непрерывный мешок слов) осуществляется предсказание слова по его контексту, во второй (Skip-Gram) – предсказание окрестности по слову.

К существенным ограничениям их использования для анализа идеологием можно отнести способность предсказания модели только для ограниченного числа слов, включенных в словарь, что ограничивает обнаружение неологизмов, производных от существующих терминов идеологического характера. Этого недостатка лишена модель fastText [10], оперирующая символьными N-граммами, отражающими морфемный состав слов.

Поскольку слова, используемые в схожих контекстах, являются квазисинонимами, а между ними наблюдаются не только отношения иерархии, для выявления таких связей могут использоваться экспертные данные, включенные в онтологии. Открытая краудсорсинговая база знаний wikidata [11] содержит среди прочих следующие виды отношений, характеризующие связи между политическими и религиозными понятиями:

- X_1 also_is X_2 (наименование класса X_1 идентично наименованию X_2 , отношение синонимии, в случае с политическими идеологиями содержит также отношения «представитель – идеология»);
- X_1 subclass_of X_2 (каждый экземпляр класса X_1 также является экземпляром класса X_2);
- X_1 instance_of X_2 (X_1 является экземпляром класса X_2);
- X_1 part_of X_2 (X_1 является составной частью X_2);
- X_1 political ideology X_2 (X_2 является идеологией, которой поддерживается персон или организация X_1).

Ввиду неоднозначности экспертной трактовки принадлежности субъекта к политической идеологии, онтологическое отношение «political ideology» можно применить для построения зависимостей только между нарицательными существительными, например «коммунист – коммунизм».

В иерархический идеологический словарь *DictH* включаются понятия, отобранные с помощью нейросетевой модели выявления именованных сущностей религиозно-политического характера.

Существующие алгоритмы с использованием статистических языковых моделей могут быть адаптированы для построения ассоциативных обобщающих цепочек идеологем.

Введем следующие определения:

w – исследуемое понятие;

s_w^n – n -ый сосед понятия w в векторном пространстве слов языковой модели;

S_1 – множество соседей первого порядка понятия w ;

S_2 – множество соседей второго порядка понятия w , такие, что $S_1 \cap S_2 \neq \{\}$;

$Mprox(w_1, w_2)$ – мера сходства векторов двух понятий;

$POS(w)$ – часть речи понятия (*noun* – имя существительное, *propn* – имя собственное);

$Idea(w)$ – понятие является идеологемой по данным нейросетевой модели, реализующей выделения именованных сущностей, или словарю.

Тогда множество потенциальных идеологических гиперонимов H_w понятия w может быть определено методом голосования посредством выбора соседей второго порядка на основе меры сходства векторов (косинусное сходство) следующим образом:

$$S_1 = \{s_w^1, \dots, s_w^{50} \mid Mprox(w, s_w^n) \rightarrow 1\}, \quad (1)$$

$$S_2 = \bigcup_{k=1}^{50} \{s_{s_w^k}^1, \dots, s_{s_w^k}^{150} \mid Mprox(s_w^k, s_{s_w^k}^n) \rightarrow 1 \wedge w \neq s_{s_w^k}^n\}, \quad (2)$$

$$Vote(s_2^m) = Vote(s_{s_w^k}^n) = \sum (Mprox(s_w^k, s_{s_w^k}^n) \mid s_{s_w^k}^n \in S_2, Idea(s_w^k) = True, POS(s_w^k) \in \{propn, noun\}), \quad (3)$$

$$Trh(Vote(s_2^m), w) = \begin{cases} 1, & \text{если } (Vote(s_2^m) \geq 1 \wedge POS(w) = noun) \vee \\ & (Vote(s_2^m) > 0 \wedge POS(w) = propn) \\ 0, & \text{иначе} \end{cases}, \quad (4)$$

$$Votes = \bigcup_{m=1}^3 \left(Vote(s_2^m) \mid Vote(s_2^m) \geq Vote(s_2^{m+1}), Trh(Vote(s_2^m), w) = 1, Vote(s_2^m) \geq \frac{Vote(s_2^1)}{2} \right), \quad (5)$$

$$H_w = \{s_2^m \mid Vote(s_2^m) \in Votes\}, \quad (6)$$

где:

$Vote(s_2^m)$ – функция подсчета голосов соседей первого порядка исследуемого понятия w в пользу соседа второго порядка $s_2^m \in S_2$, с целью поиска элементов множества S_2 , ближайших для наибольшего числа элементов S_1 ;

$Trh(Vote(s_2^m), w)$ – функция порога голосования, отсеивающая наиболее удаленные элементы S_2 ;

$Votes$ – множество 3 победителей голосования, удовлетворяющим условиям: имя существительное, не являющееся именем собственным, входящее в словарный список идеологем, значение $Vote(s_2^m)$ не менее половины максимального.

Гипо-гиперонимические отношения, построенные с помощью алгоритма на основе векторных моделей, а также найденные в онтологических базах знаний, могут быть представлены в виде ориентированного графа G , вершинами которого являются термины словаря векторной модели, а ребра имеют вид:

$$edge(w_1, w_2) = \begin{cases} (w_1, h_{w_1}) & \text{если } w_1 \notin DictH \\ (w_1, h_{w_1}^{DictH}) & \text{если } w_1 \in DictH \end{cases} \quad (7)$$

где:

$DictH$ – иерархический словарь идеологем,

h_{w_1} – значение гиперонима для понятия w_1 , полученное по итогам голосования

$h_{w_1}^{DictH}$ – словарное значение гиперонима для понятия w_1 .

Тогда под ассоциативной обобщающей цепочкой идеологемы w будем понимать любой путь l графа G , начинающийся от вершины w , удовлетворяющий следующему условию:

$$l \in L(w_0) = (w_0, \dots, w_k \mid 0 < k < 5, \exists w_k \in DictH).$$

Учитывая, что наименование идеологии является гиперонимом идеологемы, классификация множества идеологем W сводится к поиску такой вершины $h_U \in H$ графа G , что количество ассоциативных путей множества L из вершин $\{w_0, \dots, w_z\}$, проходящих через вершину w_U максимально, при этом длина ассоциативной цепочки l минимальна.

При построении ассоциативных цепочек необходимо учитывать характер отношений, полученных из онтологий. Введем следующий ряд ограничений к формуле (7), используемых при построении обобщений, препятствующий созданию синонимических рядов:

- 1) если слово есть в онтологии, но это первое слово ассоциативной цепочки, то в число ассоциатов включаются слова, полученные в рамках алгоритма голосования (формулы (1)-(6));
- 2) если слово связано с другими ассоциатами отношениями «subclass_of», то добавляются слова, связанные отношениями из онтологии;
- 3) если слово получено из онтологии, то следующие ассоциаты цепочки получаются только из онтологии;
- 4) если данное слово третье и более в цепочке, то следующие ассоциаты получаются только из онтологии.

Поскольку термины, отражающие представителей и наименования политических течений, например «сталинист» и «сталинизм», являются однокоренными словами, при формировании списка обобщающих ассоциаций для слова, кандидаты должны ранжироваться исходя из близости лексем с использованием метрики Левенштейна [12].

Алгоритм создания базы знаний наборов религиозно-политических ассоциаций на основе голосования ближайших соседей представлен на рисунке 1.

Входными данными алгоритма является словарь $Dict$ языковой модели, выходными – наборы пар $w \in Dict$ и соответствующие им непустые множества ассоциативных путей $L(w)$.

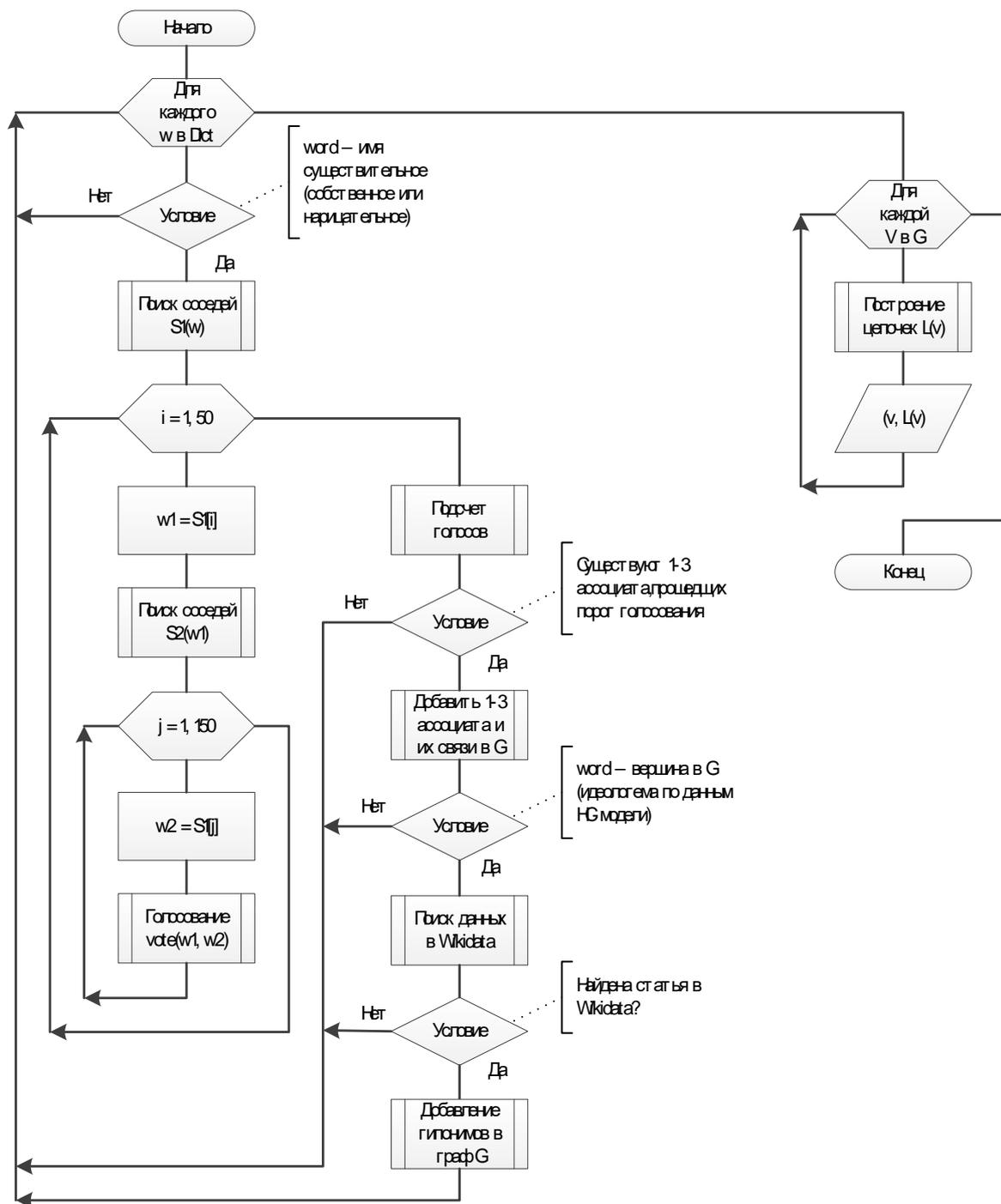


Рис. 1. Алгоритм создания базы знаний

Общий алгоритм классификации идеологем на основе наборов цепочек религиозно-политических ассоциаций имеет вид, представленный на рисунке 2.

Входными данными является множество понятий W_p , выходными – сгруппированные по соответствию обобщающих цепочек разбиения множества идеологем $W_l \subseteq W_p$ и соответствующие их элементам пути L_l .

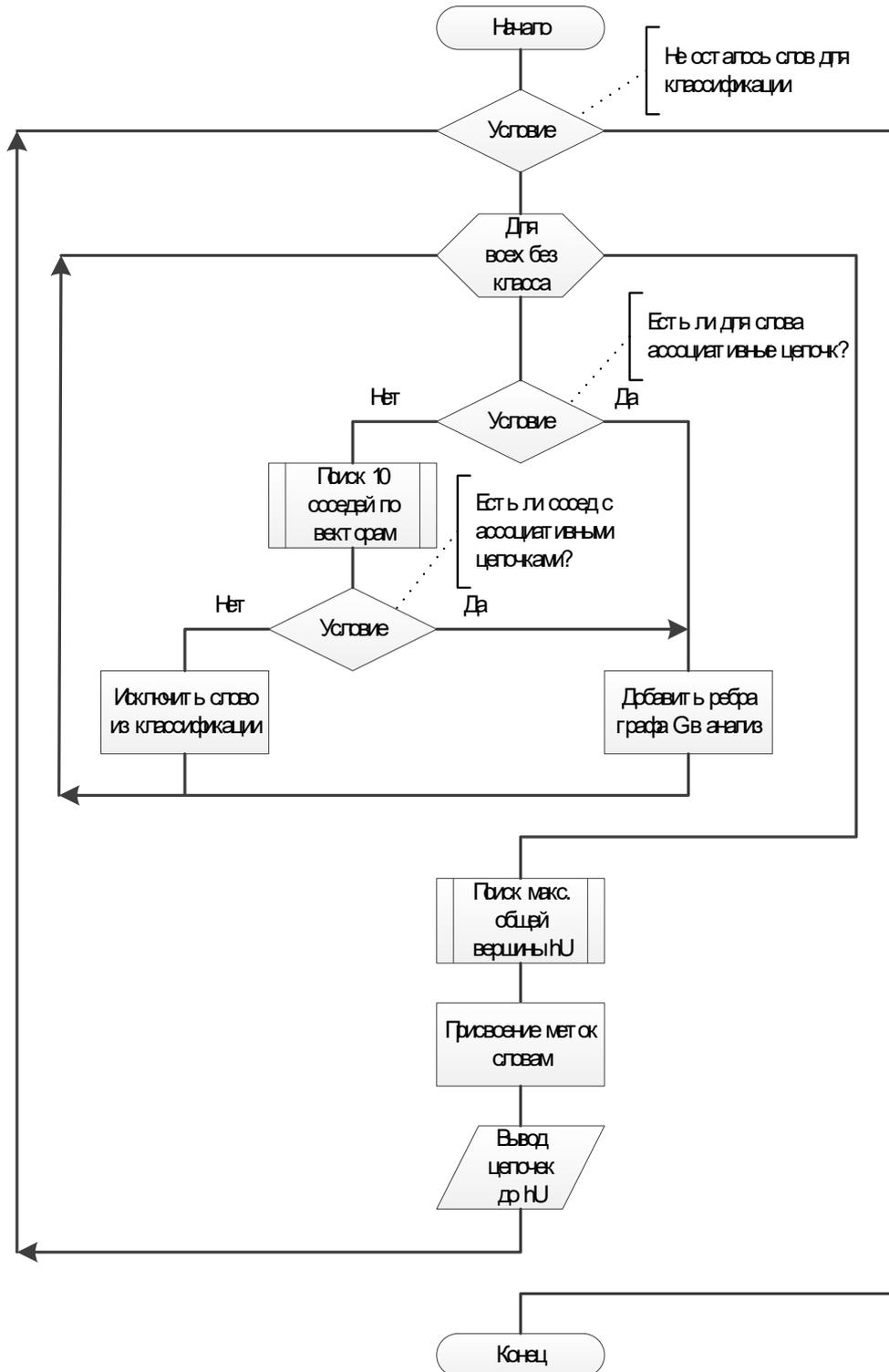


Рис. 2. Общий алгоритм классификации идеологем

Поскольку $\exists w \in W_p (w \notin G)$, необходимо с помощью статистической языковой модели на основе N-грамм установить ближайший термин $w' \approx w (w' \in G)$ такой, что соблюдается следующее условие (ближайший из десяти соседей):

$$w' = w_z (Mprox(w', w) = \max, z = \min (0, 10), Mprox(w'_z, w) \geq Mprox(w'_{z+1}, w)).$$

Для каждого отобранного на предыдущем шаге термина строятся ассоциативные обобщающие цепочки L_i на основе графа G (не более 3^5 цепочек для каждого), затем осуществляется поиск общей вершины и разбиение терминов на кластеры, где меткой класса являются вершины h_U .

Так как алгоритм стремится найти общее понятие для всех элементов подаваемого на вход множества, в ходе его работы обеспечивается разделение понятий, относящихся как минимум в одном из своих значений к разным идеологическим системам. Однако, разделение близких понятий кластера может быть достигнуто при помощи поиска общих подцепочек, при этом меткой подкластера может выступать первый элемент общей обобщающей ассоциативной подцепочки.

Результаты и их обсуждение. В ходе эксперимента применялись модели машинного обучения различных архитектур. На основе векторной языковой модели ruwikiruscorpora_upros_cbow_300_10_2021 [13] архитектуры SBOW, обученной на данных национального корпуса русского языка (НКРЯ) и Википедии по данным на конец 2021 г., производилось формирование списка слов-кандидатов для построения базы знаний, а также выявление ближайших соседей слова на этапе голосования (формулы 1)-(6). Для выявления ближайших словарных слов на этапе классификации применялась векторная модель geowac_lemmas_none_fasttextskipgram_300_5_2020 [13] архитектуры fastText, созданная в 2020 г.

Задача выявления именованных сущностей политико-религиозного характера обеспечивается моделью машинного обучения на базе многоязычного трансформера multilingual-BERT, разработанной авторами статьи [14].

Ввиду отсутствия словарей, отражающих религиозно-политические ассоциации в современном русском языке, при определении качества алгоритма применялся метод анкетирования, в рамках которого респондентам предлагалось выбрать лучшую ассоциативную обобщающую цепочку (при её наличии) между цепочкой, созданной большой языковой моделью (LLM) и алгоритмом генерации, полученной авторами.

Поскольку LLM на базе архитектуры трансформера ruGPT-3.5-13B [15] обучалась как диалоговая система, запрос на генерацию цепочки был адаптирован под формат «вопрос-ответ» и имел вид, представленный на рисунке 3.

```

prompt = f"""Ты ЭКСПЕРТ в области лингвистики и политики, готовый ответить на вопросы ЧЕЛОВЕКА.
ЧЕЛОВЕК: Построй ассоциативную обобщающую цепочку идеологического или религиозного характера для понятия "ленин".
ЭКСПЕРТ: ленин -> ленинизм -> марксизм -> социализм -> коммунизм.
ЧЕЛОВЕК: Хорошо, построй еще одну ассоциативную обобщающую цепочку для понятия "мартин лютер".
ЭКСПЕРТ: мартин лютер -> лютеранство -> протестантство -> христианство.
ЧЕЛОВЕК: Замечательно, построй ассоциативную обобщающую цепочку для понятия "{word}".
ЭКСПЕРТ: """

```

Рис. 3. Образец запроса к большой языковой модели

Пример работы разработанного авторами алгоритма классификации идеологием для множества понятий, представлены на рисунках 4-7.

Алгоритм разделил множество понятий {«сталин», «жданов», «жуков», «ворошилов», «гитлер», «холокост»}, связанные с обороной Ленинграда в годы Великой Отечественной войны на два кластера, соответствующие противоборствующим сторонам вооруженного конфликта. В то же время кластер «антикапитализм» разделяется на подкластеры «социализм» (3 понятия), «коммунизм/коммунист» (2 понятия), «марксизм» и «сталинизм».

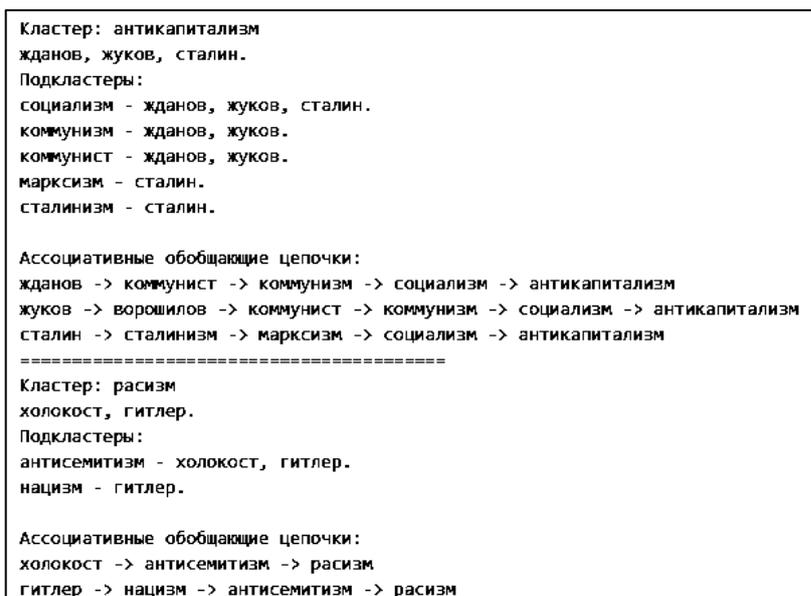


Рис. 4. Классификация понятий, связанных с противоборствующими сторонами вооруженного конфликта

На рисунке 5 представлено разбиение фамилий акторов внутренней политики СССР в 20-е годы XX века по внутрипартийным группам, например, «сталинисты» (Тухачевский, Сталин), «троцкисты» (Троцкий), «меньшевики» (Зиновьев, Рыков, Каменев). В то же время, все они могут быть объединены в понятие «марксисты».

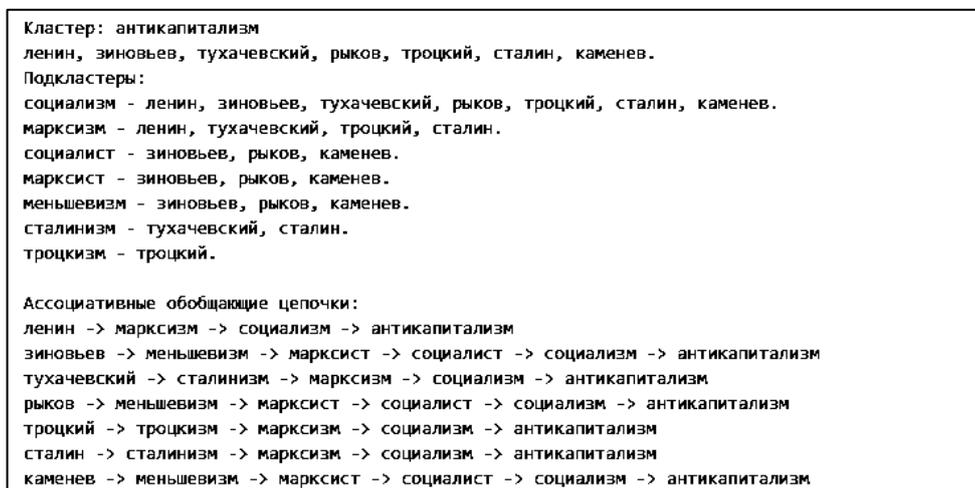


Рис. 5. Классификация понятий, связанных с внутрипартийным противоборством

Поскольку в основе алгоритма лежит база данных, созданная на основе векторного пространства понятий, отражающих современную интерпретацию терминов, классификация русских писателей XIX века характеризует оценку их творчества в настоящее время. Возникновение ассоциации «чернышевский» для понятия «белинский» в обобщающей цепочке «белинский → чернышевский → гегельянство → идеализм → монизм» свидетельствует о том, что в базе данных отсутствуют идеологические термины, связанные с личностью В. Г. Белинского, поэтому дальнейшее построение было выполнено для ближайшего термина векторного пространства, имеющего идеологические ассоциации.

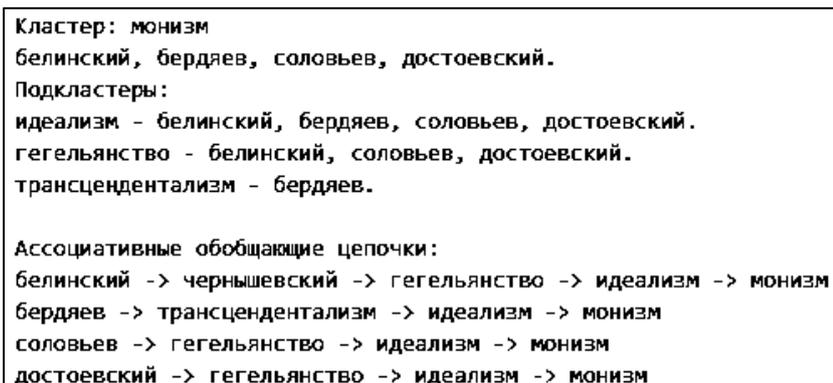


Рис. 6. Классификация русских писателей по идейному принципу

Алгоритм обеспечивает построение ассоциативных цепочек для контекстуальных идеологем. Термин «электрификация», употребляемый совместно с термином «ленин» могут быть объединены в один подкластер «социализм».

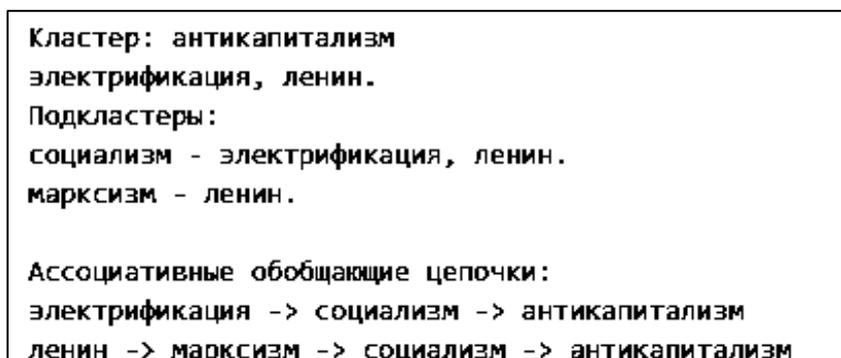


Рис. 7. Классификация терминов с учетом контекста

В ходе экспертной оценки на 100 случайных словах сферы «политика и религия», респонденты на 55.3% чаще отдавали предпочтение цепочке, полученной по алгоритму голосования, по сравнению с предсказанием модели машинного обучения. В то же время, эксперты отмечают как нерелевантные – 25.8% от предложенных цепочек, что характеризуется, в том числе, изменением трактовок политических категорий в последние годы.

Среднее время генерации ассоциативной цепочки большой языковой моделью составило 29 секунд, что более чем в 50 раз превышает время получения данных из базы знаний. Объем оперативной памяти для развертывания LLM с квантизацией (9 Гб) более чем в 8 раз превышает объем, необходимый в ОЗУ для моделей word2vec, скрипта, реализующего алгоритм, и хранения наборов цепочек.

Выводы. Таким образом, разработанный алгоритм обеспечивает классификацию идеологем русскоязычного дискурса на основе совпадения их обобщающих ассоциативных цепочек. Его преимуществом по сравнению с большими языковыми моделями является скорость работы и интерпретируемость результатов. Среди недостатков реализации алгоритма можно выделить применение устаревших статистических языковых моделей. Созданный в автоматическом режиме словарь включает ассоциативные цепочки для 25908 слов и словосочетаний русского языка и может быть использован в лингвистических и политических исследованиях при анализе трактовок идеологем второй половины 10-х годов XXI века.

СПИСОК ИСТОЧНИКОВ

1. Радина, Н. К. Методы изучения политического поля: Алгоритм идентификации контекстуальных идеологем (На примере региональной культурной политики) / Н. К. Радина, А. В. Козлова, А. А. Набокова // Политическая наука. – 2018. – № 2. – С. 252-270. – EDN XVAKIP.
2. Романова, Т. В. Идеологемы и аксиологемы русского языкового сознания как отражение констант и динамики национальной ментальности / Т. В. Романова. – Нижний Новгород: Общество с ограниченной ответственностью "Издательство "Деком", 2019. – 120 с. – ISBN 978-5-89533-444-7. – EDN GIAJRO.
3. Rao, A. Actionable and political text classification using word embeddings and LSTM / A. Rao, N. Spasojevic. – 2016. – URL: <https://arxiv.org/abs/1607.02501.X3>.
4. Biessmann, F. Automating Political Bias Prediction / F. Biessmann // Technical Report. – 2016. – URL: <http://arxiv.org/abs/1608.02195>.
5. Денисов, И. В. NEREL: Набор данных на русском языке с вложенными именованными сущностями и отношениями / И. В. Денисов, И. С. Рожков, Н. В. Лукашевич // Интеллектуальные системы. Теория и приложения. – 2021. – Т. 25, № 4. – С. 243-249. – EDN CIDPJQ.
6. Russe'2020: Findings of the first taxonomy enrichment task for the Russian language / I. Nikishina, V. Logacheva, A. Panchenko, N. Loukachevitch // Computational Linguistics and Intellectual Technologies. – 2020. – V. 19 (26). – P. 579-595. – DOI 10.28995/2075-7182-2020-19-579-595. – EDN DOMQZB.
7. Word2Vec not dead: Predicting hypernyms of co-hyponyms is better than reading definitions / N. V. Arefyev, M. V. Fedoseev, A. V. Kabanov, V. S. Zizov // Computational Linguistics and Intellectual Technologies. – 2020. – V. 19 (26). – P. 13-32. – DOI 10.28995/2075-7182-2020-19-13-32. – EDN ETWSZM.
8. Dale, D. S. A simple solution for the taxonomy enrichment task: Discovering hypernyms using nearest neighbor search / D. S. Dale // Computational Linguistics and Intellectual Technologies. – 2020. – V. 19 (26). – P. 177-186. – DOI 10.28995/2075-7182-2020-19-177-186. – EDN VITNEB.
9. Mikolov, T. Efficient estimation of word representations in vector space / T. Mikolov, K. Chen, G. Corrado, J. Dean // arXiv preprint. – 2013. – URL: <https://arxiv.org/abs/1301.3781>.
10. Fasttext. zip: Compressing text classification models / A. Joulin, E. Grave, P. Bojanowski, M. Douze, H. Jégou, T. Mikolov // arXiv preprint. – 2016. – URL: <https://arxiv.org/abs/1612.03651>.
11. Introducing Wikidata to the Linked Data Web / F. Erxleben, M. Gunther, M. Krotzsch [et al.] // The Semantic Web – ISWC 2014. – 2014. – V. 8796. – P. 50-65 – DOI 10.1007/978-3-319-11964-9_4.
12. Левенштейн, В. И. Двоичные коды с исправлением выпадений, вставок и замещений символов / В. И. Левенштейн // Доклады АН СССР – 1965. – Т. 163, №4. – С. 845-848.
13. Kutuzov, A. Webvectors: A toolkit for building web interfaces for vector semantic models / A. Kutuzov, E. Kuzmenko // Communications in Computer and Information Science. – 2017. – V. 661. – P. 155-161. – DOI 10.1007/978-3-319-52920-2_15. – EDN YVLFZN.
14. Секерин, А. В. Исследование алгоритмов машинного обучения для выявления идеологических категорий текста / А. В. Секерин, В. А. Кудинов // Интеллектуальные информационные системы: теория и практика: Сборник научных статей по материалам IV Всероссийской с международным участием конференции (Курск, 21–23 ноября 2023 г.). – Курск: Курский государственный университет, 2023. – С. 79-86. – EDN IAUWFQ.
15. Косенко, Д. Большие языковые модели для следования инструкциям на русском языке: модели и датасеты с открытой лицензией для коммерческого использования / Д. Косенко, Ю. Куратов, Д. Жарикова // Доклады Российской академии наук. Математика, информатика, процессы управления. – 2023. – Т. 514, № 2. – С. 262-269. – DOI 10.31857/S2686954323602063. – EDN GECIST.

Поступила в редакцию 16.09.2024 г., рекомендована к печати 11.10.2024 г.

CLASSIFICATION ALGORITHM OF IDEOLOGEMES BASED ON SETS OF CHAINS OF RELIGIOUS AND POLITICAL ASSOCIATIONS

Sekerin A.V., Kudinov V.A.

This paper presents an algorithm for classifying ideologemes – lexical units reflecting religious trends and political ideologies. The approach consists in building a knowledge base of associative generalizing chains of relations between concepts based on vector language models and ontologies and then searching for common elements among the chains generated for the studied words, which reduces the task to finding a common hyperonym. The comparison of the algorithm implementation with the method based on queries to large language models showed its effectiveness for classifying ideologemes.

Keywords: vector language model (word2vec), ontology, associative chain, hyperonym, large language model (LLM).

Секерин Алексей Викторович

аспирант кафедры программного обеспечения и администрирования информационных систем ФГБОУ ВО «Курский государственный университет»,
Российская Федерация, Курская область, г. Курск
E-mail: alexei.sekerin@yandex.ru

Sekerin Aleksei Viktorovich

Postgraduate Student of Department of Software and Administration of Information Systems of Kursk State University,
Russian Federation, Kursk.

Кудинов Виталий Алексеевич

доктор педагогических наук, профессор, профессор кафедры программного обеспечения и администрирования информационных систем ФГБОУ ВО «Курский государственный университет»,
Российская Федерация, Курская область, г. Курск
E-mail: kudinovva@yandex.ru

Kudinov Vitaly Alekseevich

Doctor of Pedagogy, Professor, Professor of the Department of Software and Information Systems Administration of Kursk State University,
Russian Federation, Kursk.

ИСПОЛЬЗОВАНИЕ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА В СТРОИТЕЛЬСТВЕ

© 2024 *Нестругина Е.С., Бельский Р.С.*

В данной статье проведен анализ текущего состояния и перспектив развития рынка технологий искусственного интеллекта (ИИ) в строительстве, рассмотрены наиболее перспективные области применения ИИ в строительстве, проанализированы конкретные примеры внедрения, а также выявлены ключевые преимущества и ограничения использования этих технологий строительными организациями.

Ключевые слова: искусственный интеллект, строительство, автоматизация, роботы, проектирование, прогнозирование, внедрение, эффективность.

Введение. В настоящее время строительная отрасль для повышения эффективности и качества своей деятельности активно внедряет инновационные технологии, одним из ключевых направлений которых является применение систем искусственного интеллекта (ИИ), способных изменять и оптимизировать различные этапы строительных проектов. ИИ-технологии открывают новые возможности для автоматизации, оптимизации и прогнозирования в строительстве, что особенно актуально в условиях растущей конкуренции и необходимости повышения производительности [1].

Основная часть. Использование искусственного интеллекта в строительстве является актуальной и востребованной задачей, которая активно развивается в последние годы. Искусственный интеллект находит применение в строительной отрасли на различных этапах строительных проектов – от проектирования до управления рисками. Рассмотрим применение ИИ на различных этапах строительства [1]:

1. Проектирование и инженерные расчеты.

Использование ИИ-алгоритмов позволяет автоматизировать многие расчетные задачи в процессе проектирования зданий и сооружений. Системы машинного обучения способны выполнять прочностные расчеты, оптимизировать конструктивные решения, моделировать инженерные системы. Например, компания XYZ разработала программное обеспечение, которое на основе 3D-моделей и данных о нагрузках автоматически генерирует чертежи и конструктивные схемы, значительно сокращая время проектирования.

2. Автоматизация и оптимизация строительных процессов.

Технологии искусственного интеллекта способны автоматизировать и улучшать разные стадии строительных проектов, что способствует увеличению эффективности, сокращению сроков и снижению затрат в сфере строительства. ИИ-решения могут:

– анализировать данные о ходе строительства в режиме реального времени и выявлять возможные отклонения или риски, что позволяет принимать своевременные корректирующие меры;

– моделировать и оптимизировать логистику поставок строительных материалов, уменьшая издержки и сроки доставки;

– автоматизировать составление смет, графиков работ, ведение отчетности, создание договоров и другой рутинной документации;

– прогнозировать потребность в трудовых ресурсах и эффективно распределять персонал на объектах.

3. Использование систем компьютерного зрения и роботизации.

ИИ-технологии находят применение непосредственно на строительных площадках. Роботизированные системы с компьютерным зрением способны осуществлять укладку кирпича, арматуры, бетонирование и другие повторяющиеся операции без участия человека. Это позволяет повысить производительность, качество и безопасность работ.

4. Прогнозирование и управление рисками строительных проектов.

ИИ-системы способны анализировать большие объемы данных, выявлять скрытые закономерности и на этой основе прогнозировать различные события в ходе строительных проектов. Алгоритмы машинного обучения могут предсказывать возникновение дефектов, задержки поставок материалов или погодные риски. Это позволяет строительным компаниям эффективно управлять рисками и оперативно реагировать на возникающие проблемы. Одним из ключевых направлений является применение систем искусственного интеллекта, которые способны изменять и оптимизировать различные этапы строительных проектов. По оценкам аналитиков, объем рынка ИИ-решений в строительстве превысит \$5 млрд к 2025 году [1].

Ключевыми факторами этого роста будут:

- повышение доступности и снижение стоимости ИИ-технологий;
- растущая потребность в автоматизации и оптимизации процессов;
- развитие "умных городов" и требования к экологичности строительства;
- государственная поддержка и стимулирование цифровой трансформации.

Новый этап преобразования строительной индустрии связан с внедрением искусственного интеллекта и робототехники. Роботизированные системы, усиленные компонентами ИИ, все активнее внедряются в строительные процессы. Для повышения эффективности строительных процессов внедряются новые перспективные технологии в шведской строительной компании Skanska, которая установила на своих объектах роботизированные системы для производства каркасов из арматурных стержней, помогающие удерживать бетонные конструкции [2].

Для ускорения и оптимизации процесса, компания внедряет решения на базе промышленных роботов ABB в сочетании с программным обеспечением для автономного программирования RobotStudio. Благодаря этому, удалось сократить время производства арматурных каркасов с шестнадцати часов на тонну до одного часа на тонну. Такое существенное ускорение позволяет значительно оптимизировать сроки реализации строительных проектов [2, 3].

Роботизированная система TuBOT, может связывать до 1100 пересечений арматуры в час, что эквивалентно работе 6-8 человек и существенно превосходит производительность ручного труда. Такие решения повышают эффективность, качество, безопасность, производительность и надежность строительных работ. Использование таких систем позволяет ускорить процесс вязки арматуры на 30-40% и значительно снизить риски травматизма. Внедрение роботизированных решений стало частью комплексной стратегии по цифровизации строительства. Использование роботов для производства арматурных каркасов в заводских условиях помогает оптимизировать логистику и сократить объемы транспортировки, что повышает общую экологичность строительной деятельности [3, 4].

В США появились роботы-каменщики (SAM), которые способны укладывать кирпич или вязать арматуру с высокой скоростью и точностью. Встроенный ИИ обеспечивает их адаптивность к изменениям условий. Эти автономные системы способны работать круглосуточно, без перерывов на отдых, и укладывают кирпичи в

3-4 раза быстрее, чем человек. Использование роботов-каменщиков позволило компании сократить сроки строительства на 15% и снизить количество ошибок при кладке на 80% по сравнению с ручным трудом. Кроме того, это решение помогает решить проблему нехватки квалифицированных каменщиков на рынке труда. Внедрение роботизированной кладки стало ключевым фактором повышения эффективности и рентабельности строительных проектов. Разработанные французской компанией Effidence, при сотрудничестве с Seat SA автономные транспортные роботы EffiBOT, следуют за рабочими и перевозят необходимые инструменты и материалы. Данные примеры демонстрируют, как крупные застройщики внедряют передовую робототехнику, усиленную интеллектуальными алгоритмами, для оптимизации трудоемких строительных операций. Такой подход позволяет повысить производительность, сократить сроки и издержки, а также улучшить условия труда рабочих [4].

Технологии искусственного интеллекта также интегрируются в BIM-системы. Building Information Modeling (BIM) – Технология Информационного Моделирования является передовым подходом в строительстве, обеспечивающим актуализацию данных о проекте на всех этапах его жизненного цикла для всех участников, помогающим автоматизировать проверку соответствия цифровых моделей нормативным требованиям. Это повышает эффективность и скорость проектирования, минимизируя ошибки. Кроме того, ИИ применяется для более точного прогнозирования и оценки инвестиций в строительные проекты еще на этапе планирования. Моделирование ключевых допущений и анализ затрат на жизненном цикле объекта позволяет с большей предсказуемостью оценить итоговые результаты. BIM позволяет создать виртуальную 3D-модель здания с подробной информацией о конструкции, деталях, характеристиках материалов и ценовых параметрах. Эта модель формируется на основе объектно-ориентированного проектирования, когда отдельные компоненты (стены, окна, перекрытия и пр.) объединяются в единую систему. Добавление временного параметра позволяет получить 4D-визуализацию, дающую возможность смоделировать все этапы возведения объекта в режиме реального времени. ИИ-алгоритмы способны анализировать сценарии с учетом рисков и оптимизировать длительность работ. Ключевыми характеристиками BIM-технологии являются: доступность для анализа, взаимосогласованность и взаимосвязанность данных, геометрическая привязка, интероперабельность, возможность обновления. Использование цифровых BIM-моделей становится неотъемлемой частью всего жизненного цикла строительного проекта – от начальных этапов планирования до последующей эксплуатации и техобслуживания. Архитекторы, инженеры и другие специалисты работают над созданием единой BIM-модели. Таким образом, BIM-технология позволяет существенно повысить эффективность, качество и управляемость строительных процессов [5].

Внедрение технологий искусственного интеллекта открывает широкие возможности для оптимизации строительных процессов, сокращения издержек и повышения общей эффективности.

Строительные компании активно внедряют новые технологии, в том числе беспилотные летательные аппараты (дроны) для мониторинга и обеспечения безопасности на строительных объектах. Они позволяют оперативно получать актуальную информацию со стройплощадки в режиме реального времени – проводить измерения, контролировать соблюдение техники безопасности и многое другое. Например, дрон SkyMule помогает ускорять процесс вязки арматуры в 2,4 раза [6, 7].

Использование дронов позволяет охватить большой объем территории, оперативно обнаруживать и устранять различные риски. Беспилотники способны передавать видео в режиме реального времени, обеспечивая мгновенный контроль за ситуацией на объекте. Более того, интеграция дронов с системами искусственного интеллекта расширяет возможности мониторинга. ИИ-алгоритмы могут автоматически анализировать данные с беспилотников и выявлять различные проблемы – от утечек газа до пожароопасных ситуаций. Это позволяет значительно повысить эффективность обеспечения безопасности строительных площадок. Дроны также применяются для удаленного технического обслуживания, особенно в труднодоступных зонах. Тепловизионная съемка с беспилотников помогает диагностировать нарушения в теплоизоляции и другие дефекты, что упрощает планирование ремонтных работ [6, 7].

Помимо использования беспилотников, строительные компании внедряют и другие цифровые технологии. Крупные фирмы разрабатывают собственные программные решения для автоматизации ключевых процессов – от бюджетирования до управления закупками. В то же время, малые и средние компании нередко предпочитают использовать готовые ERP-системы, такие как "Галактика", "Алтиус" или Oracle JD. Эти комплексные решения обеспечивают широкий спектр инструментов для оптимизации деятельности, включая стратегическое планирование. Таким примером являются клиентские порталы, которые повышают прозрачность строительных проектов и улучшают взаимодействие с заказчиками [8].

Таким образом, внедрение передовых технологий, включая дроны, ИИ и ERP-системы, открывает новые возможности для повышения эффективности, безопасности и качества в строительной отрасли. Отечественные компании также внедряют цифровые технологии и системы ИИ. Группа "Самолет" разработала нейросетевую систему VIJU для мониторинга стройплощадок, что позволило сократить простои и повысить производительность на 40%. ДОМ.РФ использует ИИ для оценки ликвидности строящихся объектов и оптимизации процессов [8].

Таким образом, интеграция робототехники, дронов и искусственного интеллекта становится ключевым фактором трансформации строительной отрасли, повышая её эффективность, безопасность и качество.

Использование технологий виртуальной реальности (VR) на основе искусственного интеллекта применяется для обучения рабочих техникам безопасности в строительстве [9].

Данные модели воссоздают реальные сценарии и ситуации на строительных объектах, позволяя рабочим отрабатывать навыки выявления и снижения потенциальных рисков. Искусственный интеллект анализирует реакции и действия работников в этих виртуальных условиях, предоставляя обратную связь и рекомендации по повышению безопасности [8, 9].

Таким образом, прежде чем приступить к работе на реальном объекте, сотрудники уже будут обучены распознавать и предотвращать различные опасности. Это позволяет значительно повысить уровень подготовки персонала по вопросам охраны труда и техники безопасности, снижая риски травматизма на строительных площадках.

Искусственный интеллект обладает очевидными преимуществами и существенно трансформирует систему управления в строительстве. Данные показывают рост производительности, сокращение издержек и времени проектирования, повышение эффективности [8, 9].

Применение ИИ-технологий в строительстве обеспечивает ряд преимуществ [9]:
– повышение производительности и качества работ;

- оптимизация процессов и снижение издержек;
- улучшение безопасности на строительных площадках;
- более эффективное управление проектами и рисками.

В то же время внедрение ИИ в строительстве сталкивается с определенными ограничениями и барьерами [9]:

- необходимость интеграции с существующими ИТ-системами;
- доступность и качество данных для обучения алгоритмов;
- недостаток квалифицированных кадров в области ИИ;
- опасения относительно технологических рисков и отказа от человеческого контроля.

Кроме того, строительная индустрия сталкивается с такими ключевыми проблемами, как высокие внеплановые расходы и срывы сроков строительства, ежедневно приводящие к миллионным убыткам. Часто это происходит из-за низкой эффективности взаимодействия участников и недостаточного планирования.

Использование технологий искусственного интеллекта способно повысить эффективность строительных процессов. ИИ-решения позволяют оценивать потребность в ресурсах, отслеживать использование техники, контролировать качество работ, снижать риски и обеспечивать безопасность на объекте, создавать цифровые двойники и многое другое. В частности, прогнозная аналитика на базе ИИ помогает выявлять и предотвращать многие проблемы. Система анализирует накопленные данные, обучается и дает прогнозы о влиянии различных факторов на ход строительства. Это позволяет принимать более обоснованные управленческие решения. Вместе с тем, массовому внедрению ИИ препятствует ряд факторов. Многие руководители и специалисты недооценивают или недостаточно осведомлены о возможностях этих технологий. Также существуют технические и кадровые ограничения [9].

Несмотря на барьеры, ИИ имеет высокий потенциал в строительстве. Он способен ускорять проектирование, повышать производительность, улучшать контроль над стройкой, оптимизировать ресурсы.

Однако, будучи новым инструментом, искусственный интеллект требует грамотного освоения. Это предъявляет новые требования к специалистам, их обучению, технической базе, а также к стратегическому планированию.

В качестве одной из ключевых задач выделяется развитие образовательных программ по ИИ в рамках национального проекта "Цифровая экономика". Таким образом, внедрение искусственного интеллекта в строительстве открывает значительные возможности, но также требует системного подхода к преодолению барьеров и наращиванию компетенций. Приоритетные направления развития искусственного интеллекта в строительной отрасли систематизированы в таблице 1. Для каждого из этих направлений выявлены ключевые проблемы, а также предложены возможные пути их решения. Кроме того, в таблице представлен прогноз ожидаемых результатов от внедрения технологий искусственного интеллекта в различные сферы строительства [9].

Данная структурированная информация позволяет комплексно оценить перспективы и потенциал применения ИИ-решений для повышения эффективности строительной деятельности, выявить существующие барьеры и определить необходимые меры по их преодолению. С целью ускоренного внедрения ИИ-технологий в экономике, Минстрой России совместно с ДОМ.РФ создали экспертную группу по внедрению искусственного интеллекта в строительной отрасли.

В нее входят крупные застройщики, представители госорганов, регулятора и банков. В рамках работы этой группы рассматриваются вопросы создания единой методики для оценки готовности компаний к применению ИИ. Данная методика должна стать основой для формирования реестра отечественных решений на базе искусственного интеллекта. Также на заседании экспертной группы обсуждалось развитие нормативно-правового регулирования использования ИИ в строительстве. Это предложение было озвучено на профильной сессии "Искусственный интеллект в девелопменте: от идей к инновациям" в рамках выставки-форума "Россия" на ВДНХ [9].

Таблица 1. Приоритетные направления развития ИИ в строительной отрасли

Приоритетное направление	Проблема	Решение	Результат
Мониторинг строительных объектов и рабочих	Снижение эффективности работы из-за нарушения норм рабочими, задержек и отсутствия своевременного отчета руководству	Анализ данных с носимых устройств и стационарных камер для контроля за временем активности и покоя рабочих. Использование дронов и камер для отслеживания хода строительства, поставок материалов и затрат ресурсов	Увеличение производительности на 30 %
Прогнозирование рисков и автоматизация типовых операций	Затраты времени и ошибки из-за ручного выполнения стандартных задач. Отсутствие автоматизации для учета рисков в планах	Интеграция ИИ в системы управления для разработки смет, анализа и прогнозирования на основе данных. Внедрение модулей ИИ для стандартного проектирования, включая разработку инженерных сетей и соответствие нормам	Увеличение производительности на 20 % Снижение издержек на 10 % Сокращение времени проектирования на 40 %
Применение автономной техники	Нехватка квалифицированных работников для управления сложной техникой	Внедрение автономных строительных машин, способных работать непрерывно, чтобы облегчить рабочий процесс и повысить эффективность	Снижение операционных затрат на 50 % Ускорение строительства на 50 %

Таким образом, государство, отраслевые эксперты и ключевые игроки рынка совместно прорабатывают комплексные меры по ускорению внедрения технологий искусственного интеллекта в строительной сфере.

В качестве примера применения искусственного интеллекта в строительной отрасли приведем программную реализацию модуля прогнозирования электропотребления. Распространенным подходом к прогнозированию временных рядов является использование нейросетевых моделей (ANN). Часто в таких задачах применяется прогнозирование на один шаг вперед, когда на вход нейронной сети подается вектор из предыдущих значений ряда $x(t) = (x_{t-\tau+1}, x_{t-\tau+2}, \dots, x_t)$, а на выходе получается прогноз значения x_{t+1} где параметр τ характеризует глубину погружения (ширину окна). В качестве конкретной нейросетевой архитектуры в работе используется многослойный перцептрон [10, 11]. Этот выбор обусловлен следующими факторами:

- обучение нейронной сети происходит с учителем, то есть в обучающее множество входят и выходные значения;
- используется модель прямого распространения, где связи между нейронами однонаправленные;
- наиболее распространенная архитектура – трехслойный персептрон с одним скрытым слоем.

Данный подход к прогнозированию временных рядов с применением многослойного персептрона является распространенным и хорошо зарекомендовавшим себя на практике.

Таким образом, нейронная сеть имеет следующую архитектуру, показанную на рисунке 1 [10].

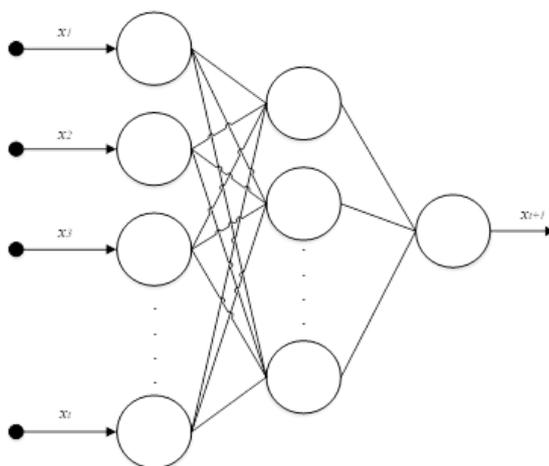


Рис. 1. Вид нейронной сети

Прогнозирование с использованием нейронной предусматривает следующие этапы [10]:

1. Разбиение временного ряда данных по электропотреблению на обучающую и тестовую выборки.
2. Обучение нейронной сети по методу обратного распространения ошибки back propagation error с использованием обучающей выборки.
3. Проведение тестового прогнозирования и определение оценки ошибки прогнозирования.

Количество нейронов скрытого слоя определяется с помощью серии опытов при обучении и тестировании нейронной сети. Для обучения нейронной сети выбран алгоритм обратного распространения ошибки, так как он обладает рядом преимуществ:

1. Достаточно высокая эффективность.
2. Высокая степень общности. Алгоритм можно легко перенастроить для произвольного количества слоев, нейронов входа и выхода, для произвольной функции активации.
3. Достаточно прост в реализации.

Ошибка обучения нейронной сети находится с использованием функции средней квадратичной ошибки [10, 11]:

$$E = \frac{1}{N} \sum_{i=1}^N (x_i - \hat{x}_i)^2, \quad (1)$$

где E – средняя квадратичная ошибка,

N – количество значений во временном ряду,

x_i – фактическое значение временного ряда,

\hat{x}_i – прогнозируемое значение временного ряда.

Для разработки нейронной сети выбран набор библиотек AForge.Neuro [11]. Данная платформа поддерживает достаточно широкий круг архитектур нейронных сетей и алгоритмов их обучения, включая многослойную нейронную сеть с обратным распространением ошибки.

Перед началом прогнозирования необходимо определить следующий набор параметров:

- learningRate – скорость обучения;
- momentum – момент обучения;
- sigmoidAlphaValue – параметр наклона сигмоидальной функции активации;
- windowSize – количество исторических значений, используемых для построения прогноза (глубина погружения или размер окна);
- iterations – количество итерация обучения.

Создание многослойной нейронной сети происходит путем создания нового объекта и задания соответствующей функции активации:

```
network := TActivationNetwork.Create(2,  
BipolarSigmoidFunction(sigmoidAlphaValue))
```

Установка алгоритма обучения нейронной сети выполняется путем создания нового объекта BackPropagationLearning:

```
teacher := BackPropagationLearning.Create((network)).
```

Само обучение происходит в отдельном потоке до тех пор, пока либо пользователь не нажмет кнопку «Стоп», либо не пройдут все заданные итерации:

```
WorkerThread := TThread.Create(SearchSolution);  
workerThread.Start.
```

Показания электропотребления для обучения сети снимались каждые пятнадцать минут в течение нескольких дней.

Количество нейронов на входном слое будет равняться установленному параметру windowSize, в качестве которого было выбрано двенадцать значений временного ряда, которые соответствуют временному промежутку равному трем часам. Согласно (1) количество нейронов скрытого слоя не должно быть меньше четырех нейронов, поэтому была проведена серия испытаний, где количество нейронов скрытого слоя нейронных сетей принимает значения в диапазоне от четырех до двенадцати (число нейронов входного слоя) и значение в два раза превышающее количество входных нейронов. Однако значительное увеличение количества нейронов скрытого слоя существенно не улучшило характеристики сети. Каждая нейронная сеть обучалась на протяжении 10000 итераций. Результаты испытаний представлены в таблице 2.

Таблица 2. Испытания нейронной сети

Номер испытания	Количество нейронов на скрытом слое	Средняя квадратичная ошибка обучения
1	4	22,600
2	5	25,040
3	6	23,320
4	7	24,400
5	8	23,880
6	9	24,740
7	10	22,030
8	11	23,540
9	12	22,070
10	24	22,910

В результате нескольких испытаний получено, что скрытый слой нейронной сети включает в себя десять нейронов. В результате средняя квадратичная ошибка обучения составила 22, что говорит об адекватности разработанной нейросетевой модели [11].

Результат работы модуля прогнозирования включает в себя формирование таблицы с реальными и прогнозируемыми значениями и построение соответствующего графика временного ряда с реальными и прогнозируемыми значениями приведен на рисунке 2.

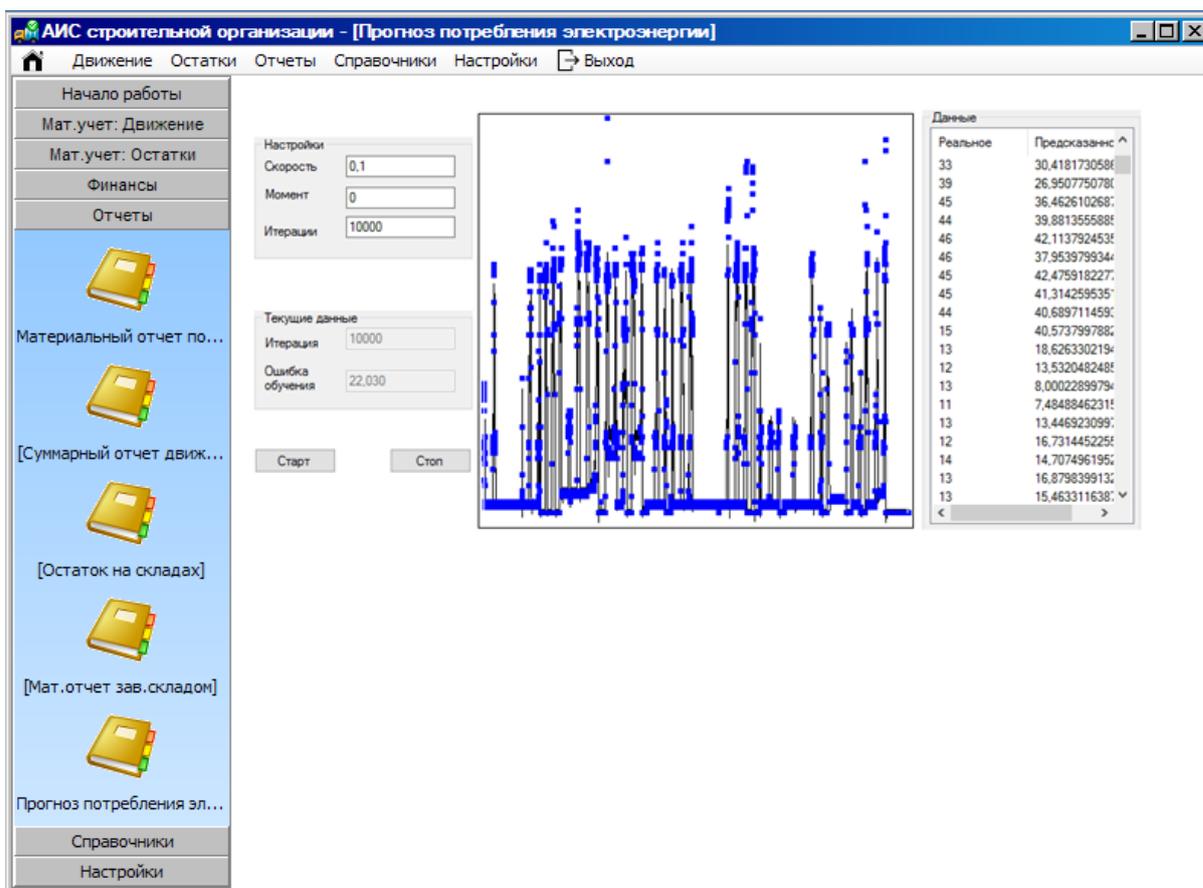


Рис.2. Прогноз потребления электроэнергии на участках

График процесса обучения приведен на рисунке 3.

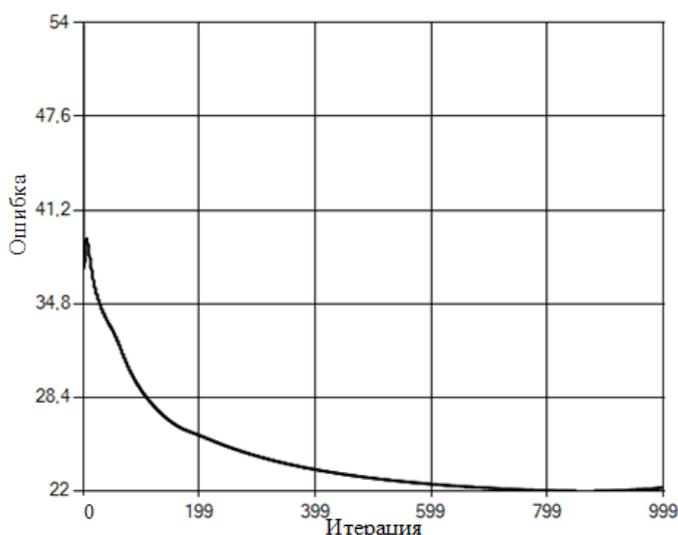


Рис. 3. График обучения нейронной сети

Выводы. Технологии искусственного интеллекта становятся неотъемлемой частью жизни человека, и строительная отрасль не является исключением. Внедрение инновационных разработок, таких как дроны, роботизированное оборудование и экзоскелеты, кардинально трансформирует строительные процессы.

Несмотря на высокие первоначальные затраты, эти технологии играют ключевую роль в формировании современного облика строительства, делая его более эффективным. Интеграция IT-решений оказывает положительное влияние как на конкретные этапы работы, так и на всю отрасль в целом.

Использование систем искусственного интеллекта открывает широкие возможности для повышения эффективности и качества строительной отрасли. ИИ-технологии находят применение на различных этапах строительных проектов, обеспечивая автоматизацию, оптимизацию и прогнозирование процессов. Несмотря на существующие ограничения, перспективы развития ИИ в строительстве выглядят весьма многообещающими. Строительным компаниям важно активно внедрять эти инновационные решения для повышения своей конкурентоспособности.

СПИСОК ИСТОЧНИКОВ

1. Васильев, А. С. Использование технологий искусственного интеллекта в управлении строительными проектами / А. С. Васильев, А. В. Зайцев // Научно-технический вестник Поволжья. – 2020. – № 5. – С. 24-28.
2. С помощью роботов Skanska устанавливает новые правила в вязке арматуры [Электронный ресурс]. – URL: <https://rcmm.ru/novosti/53529-s-pomoschju-robotov-skanska-ustanavlivaet-novye-pravila-v-vjazke-armatury.html> (дата обращения 15.08.2024).
3. Skanska устанавливает новые правила по вязке арматуры [Электронный ресурс]. – URL: <https://robogeek.ru/promyshlennye-roboty/skanska-ustanavlivaet-novye-pravila-po-vjazke-armatury> (дата обращения 15.08.2024).
4. Интеграция искусственного интеллекта с робототехникой в строительстве [Электронный ресурс]. – URL: <https://www.stroitelnye.ru/integration-of-artificial-intelligence-with-robotics-in-construction> (дата обращения 15.08.2024).
5. Моделирование и применение BIM-подхода в строительстве [Электронный ресурс]. – URL: <https://firecode.ru/blog/modelirovanie-i-primenenie-bim-podhoda-v-stroitelstve> (дата обращения 15.08.2024).

6. Сжасный Н. П. Применение беспилотных летательных аппаратов в строительной отрасли / Н. П. Сжасный // Строительство и недвижимость. – 2021. – № 3. – С. 12-17.
7. Роботы и дроны в строительстве: повышение эффективности и безопасности [Электронный ресурс]. – URL: <https://robogeek.ru/promyshlennye-roboty/skanska-ustanavlivaet-novye-pravila-po-vyazke-armatury> (дата обращения 15.08.2024).
8. Умные технологии в строительстве: искусственный интеллект и робототехника // Строительный эксперт. – 2021. – № 4. – С. 18-23.
9. Искусственный интеллект в строительстве [Электронный ресурс]. – URL: <https://stroygaz.ru/publication/technologies/iskusstvennyu-intellekt-v-stroitelstve> (дата обращения 15.08.2024).
10. Кириченко, А. А. Программное обеспечение нейросетевых исследований / А. А. Кириченко. – М.: Изд. HSE, 2016. – 152 с.
11. Садовникова, Н. П. Разработка нейронной сети с использованием AForge.NET Framework / Н. П. Садовникова, Д. С. Парыгин. – ВолгГТУ. – Волгоград, 2016. – 20 с.

Поступила в редакцию 26.09.2024 г., рекомендована к печати 14.10.2024 г.

USE OF ARTIFICIAL INTELLIGENCE IN CONSTRUCTION

Nestrugina E.S., Belskii R.S.

This article analyzes the current state and prospects for the development of the market for artificial intelligence (AI) technologies in construction, examines the most promising areas of application of AI in construction, analyzes specific examples of implementation, and identifies the key advantages and limitations of the use of these technologies by construction organizations.

Keywords: Keywords: artificial intelligence, construction, automation, robots, design, forecasting, implementation, efficiency.

Нестругина Елена Сергеевна

кандидат технических наук, доцент, доцент кафедры компьютерных технологий ФГБОУ ВО «Донецкий государственный университет», Российская Федерация, ДНР, г. Донецк.
E-mail: e.nestrugina@donnu.ru

Nestrugina Elena Sergeevna

Candidate of Technical Sciences, Associate Professor, Associate Professor at Department of Computer Technologies of Donetsk State University, Russian Federation, DPR, Donetsk.

Бельский Роман Сергеевич

магистрант кафедры компьютерных технологий ФГБОУ ВО «Донецкий государственный университет», Российская Федерация, ДНР, г. Донецк.
E-mail: belskiy_90@inbox.ru

Belskii Roman Sergeevich

Master Student at Department of Computer Technologies of Donetsk State University, Russian Federation, DPR, Donetsk.

КИБЕРНЕТИЧЕСКИЙ ПОДХОД К ПОНИМАНИЮ ЯЗЫКА

© 2024 Булыгин В.В.

Цель статьи: выразить, что такое язык и как он возникает. Подход с точки зрения кибернетики. Продемонстрировано, что поиск объектов может быть осуществлен средствами обучения с подкреплением. Такие объекты, максимизированные по коэффициенту однозначности, могут вступать в отношения. Комбинации базовых отношений по управлению и иерархии дают возможность строить предложения. Что позволяет моделировать действительность и служить средством коммуникации.

Ключевые слова: объект, иерархия, управление, коэффициент однозначности, язык, предложение.

Введение. Отдавая должное лингвистике [1] и направлению изучения знаковых структур вообще [2], базой изложенных ниже взглядов является подход Фреге [3]: значение, смысл, знак. Знак именуется смыслом, смысл частично или полностью передает значение. Рассмотрим не пример Фреге со звездой, а более простой с монетой. Монета может являться или решкой, или орлом. Тогда можно сказать (монета - решка), (монета - орел), (монета - решка, если ее повернуть из орла), (монета - орел, если ее повернуть из решки). Такие переходы более компактно можно отобразить в виде графа.

монета = {
решка: [[орел], повернуть], [[решка], не_поворачивать]],
орел: [[[решка], повернуть], [[орел], не_поворачивать]],
}

Поэтому (знак, смысл, значение) можно понимать так: монета есть знак, значение есть граф переходов, смыслом же является или сам граф, или его часть. Например, смыслом может быть: {решка: решка}, {решка: [[орел], повернуть]}.

Для наглядности, можно представить в виде рис. 1 с переходами. Протокол наблюдения за монетой может быть таким: [(р,н), (р,п), (о,п), (р,п), (о,н), (о,н), (о,н)] или, если без учета действия, таким [р, р, о, р, о, о, о]. Часть такой последовательности представлена на рис. 2

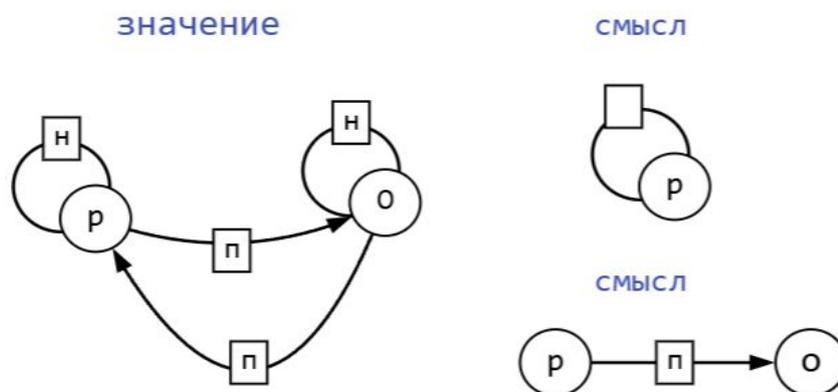


Рис. 1. Значение и смысл

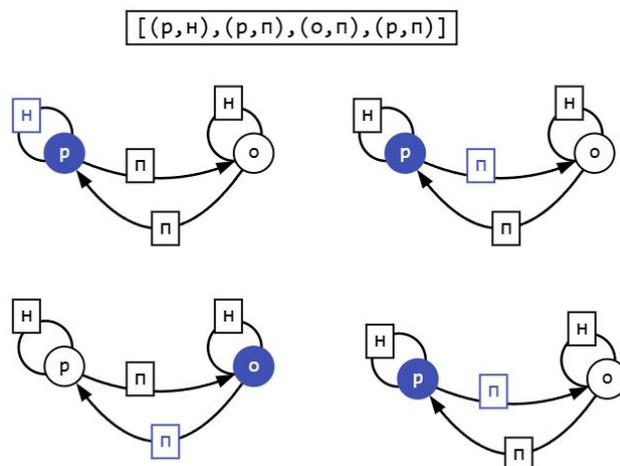


Рис. 2. Движение представляющей точки

При этом само значение должно быть, по крайней мере на каком-то периоде времени, неизменным. Иначе нет смысла хранить его отображение в памяти. Физические законы, например, можно отнести к неизменяемым законам, а биологические и тем более социальные - к изменяемым (неизменяемым на определенном промежутке времени).

Чтобы оценить предсказуемость, используется коэффициент однозначности [4] $ku = n / (n + u)$, где

n - количество различных $\langle f, x \rangle$ графа

$n + u$ - количество различных $\langle y, f, x \rangle$ графа

По графу ниже:

$b = 'b'; B = 'B'; w = 'w'; x = 'x'; y = 'y'; z = 'z'$

$g = \{$

$w: [[w, z], b], [[x, y], B], \# b: w \rightarrow w|z, B: w \rightarrow x|y$

$x: [[y, w], b], [[x, z], B], \# b: x \rightarrow y|w, B: x \rightarrow x|z$

$y: [[y, x], b], [[w, z], B], \# b: y \rightarrow y|x, B: y \rightarrow w|z$

$z: [[y, x], b], [[w, z], B], \# b: z \rightarrow y|x, B: z \rightarrow w|z$

$\}$

`import math`

`def get_ku(g):`

`n=0; u=0; h=[]`

`for x in g:`

`for y in g[x]:`

`if y[0] != []:`

`u += len(y[0]) - 1`

`if len(y[0]) > 1:`

`h.append(round(math.log(len(y[0]), 2), 4))`

`n += 1`

`if n != 0:`

`ku = round(n/(u+n), 3)`

`print("h =", h, "#биты переходов с неопределенностью")`

`print("n =", n, "#оценка познанной сложности (двойки: <f,x>)")`

`print("u+n =", u+n, "#количество переходов (тройки: <y,f,x>)")`

`print("ku = n/(u+n) =", ku)`

`print("ku == n/(sum([2**b for b in h]) - len(h) + n) is", \`

`round(ku, 3) == round(n/(sum([2**b for b in h]) - len(h) + n), 3))`

`return (n, u, h)`

`r = get_ku(g)`

Вывод:

```
h = [1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0]
#биты переходов с неопределенностью
n = 8 #оценка познанной сложности (двойки: <f,x>)
u+n = 16 #количество переходов (тройки: <у,f,x>)
ku = n/(u+n) = 0.5
ku == n/(sum([2**b for b in h]) - len(h) + n) is True
```

Основная часть. Форма записи последовательности (вход, выход) в кибернетике или, как в обучении с подкреплением, (состояние, действие (вход)) является типичной и общей

```
[(a,g), (a,j), (a,f), (a,f), (a,f), (b,f), (b,h), (b,h), (a,h), (a,j), (b,f), (a,h), (b,j), (b,f), (a,h),
(b,j), (a,f)]
```

Такая последовательность может представлять собой как психиатрическое заболевание (с вопросами {a, b} и ответами {f, g, h, j}), так и описание электрической цепи [5].

Такой протокол может быть представлен таблицей или графом:

```
obj1 = {# такой, что его переходы:
f: [[[f], a], [[h], b]], # a: f->f, b: f->h
g: [[[j], a]], # a: g->j
h: [[[j], a], [[h], b]], # a: h->j, b: h->h
j: [[[f], a], [[f], b]] # a: j->f, b: j->f
} # здесь ku=7/7
```

По сути, obj1 - это переменная, переходы состояний которой зависят от действия. Можно сказать, что (obj1 - f) истинно. Как и истинно: (obj1 - g), (obj1 - h), (obj1 - j). Но ложно (obj1 - m), т.к. такого состояния m нет. Сравнимо: (белый снег - белый), (белый снег - черный)

Такие объекты могут вступать в отношения. Базовыми являются отношения управления и иерархии. Так, объект obj2 управляет объектом obj1.

```
obj2 = {# состояния obj2 являются действиями объекта obj1.
a: [[[a], b], m]], # m: a->a|b
b: [[[a], m]] # m: b->a
} # здесь ku=2/3
```

Объект также может быть состоянием другого объекта. Ниже пример отношения иерархии.

```
obj3 = {# состоянием объекта obj3 является другой объект obj1.
p: [[[obj1], q]], # q: p->obj1
obj1: [[[obj1], q]] # q: obj1->obj1
} # здесь ku=2/2
```

Если абстрагироваться от конкретных состояний и действий, то отношения между объектами можно представить отношениями языка: obj2=>obj1, obj3.obj1.

Из таких базовых конструкций можно строить предложения. Например, таким предложением может быть: obj2.obj6 obj4.obj2 obj2=>obj1 obj1=>obj5. Прочсть можно так: шестое два четырех управляет пять единиц. Цифры в качестве примера (аналог глокой куздры) взяты намеренно, чтобы отобразить только синтаксические связи. Часть такого предложения или все предложение может выступать в качестве объекта-переменной obj7 в составе другого предложения, формируя текст.

Содержательная же часть, его семантика следует из описания переходов состояний для выбранных объектов. Касается как отношения управления, так и отношения иерархии. Конкретизировать можно, если задать какую именно последовательность выбрать. Так, obj2.[a,b,a,a] задает смену состояний, как кадры в киноплёнке, отображая смысл сказуемого в глагольной форме.

Если обозначить такую последовательность как $obj2.[[8]]$, то предложение ниже можно прочесть: (шестое два четырех восьми пять единицей) объект $obj2$ имеет состояние $obj6$, сам объект $obj2$ является состоянием объекта $obj4$, $obj2$ управляет объектом $obj5$ своим действием 8, используя $obj1$.

$obj2.obj6\ obj4.obj2\ obj2.[[8]]\ obj2 \Rightarrow obj1\ obj1 \Rightarrow obj5$

В отличие от подхода Хомского, такое абстрагирование до отношений объектов-переменных не следует относить к врожденным способностям. Оно лишь следствие процесса: механизмом самосборки с усложнением является поиск новизны с высокой предсказуемостью из существующего разнообразия.

Это разнообразие таково, что позволяет видеть такие объекты повсюду. Пнутый мячик, ушко кошки, т.е. сама среда формируют у ребенка мышление еще до того момента, как он начинает говорить звуко-раздельными словами в соответствии с принятым синтаксисом. Язык в этом, более широком смысле - это и есть мышление. Задачей которого является отображение существующих явлений. И, как следствие, возможность конструирования еще несуществующих явлений. По той простой причине, что за элементы конструктора берутся переменные, в которые можно подставить все.

Важно еще то, что рост неоднозначности является сигналом к переопределению. Пусть имеется объект obj .

$obj = \{$
 $w: [[w], b], [[x], B], \# b: w \rightarrow w, B: w \rightarrow x$
 $x: [[y], b], [[x], B], \# b: x \rightarrow y, B: x \rightarrow x$
 $y: [[y], b], [[w], B], \# b: y \rightarrow y, B: y \rightarrow w$
 $z: [[y], b], [[w], B], \# b: z \rightarrow y, B: z \rightarrow w$
 $\} \# \text{здесь } ku=8/8$

Но затем однозначность уменьшается (ku уменьшается с 1 до 0.5).

$obj = \{$
 $w: [[w, z], b], [[x, y], B], \# b: w \rightarrow w|z, B: w \rightarrow x|y$
 $x: [[y, w], b], [[x, z], B], \# b: x \rightarrow y|w, B: x \rightarrow x|z$
 $y: [[y, x], b], [[w, z], B], \# b: y \rightarrow y|x, B: y \rightarrow w|z$
 $z: [[y, x], b], [[w, z], B], \# b: z \rightarrow y|x, B: z \rightarrow w|z$
 $\} \# \text{здесь } ku=8/16$

Это могло быть следствием, что неучтенными оказались другие компоненты вектора-действия. Вместо $\{b, B\}$ могло быть $\{(a, b), (a, B), (A, b), (A, B)\}$.

$obj = \{$
 $w: [[w], (a, b)], [[x], (a, B)], [[z], (A, b)], [[y], (A, B)],$
 $x: [[y], (a, b)], [[x], (a, B)], [[w], (A, b)], [[z], (A, B)],$
 $y: [[y], (a, b)], [[w], (a, B)], [[x], (A, b)], [[z], (A, B)],$
 $z: [[y], (a, b)], [[w], (a, B)], [[x], (A, b)], [[z], (A, B)]$
 $\} \# \text{здесь } ku=16/16$

С синтаксической неоднозначностью тоже. В отношении $obj2$ и $obj5$ она устраняется, как известно, за счет порядка. В предложении ($obj2 \Rightarrow obj1\ obj1 \Rightarrow obj5$) объект $obj1$ является как управляющим, так и управляемым. Чтобы устранить неоднозначность по отношению и управляющему объекту $obj2$ и управляемому объекту $obj5$, необходимо $obj1$ маркировать особым образом. В русском языке это делается за счет творительного падежа. Той же цели можно добиться за счет использования предлога. Либо, как представлено выше, писать парами. Минусом последнего представления является ее избыточность. Аналогично с конструкциями ($obj2.obj6\ obj4.obj2$) для объекта $obj2$ относительно отношения иерархии.

Как пример, что синтаксис может переопределяться (не быть врожденным), Глухов [6] приводит лексемный синтаксис ребенка, когда грамматическая форма зачастую случайна: «в значении «вот молоко» может употребляться форма маА, усвоенная из фраз типа «хочешь молока» или «дай молока»».

Следует добавить, что «большинство наших слов, в отличие от животных, не о нас; и наши слова в основном не о здесь и сейчас» [7]. «Не абстрактные понятия, а понятия-конструкты составляют специфику человеческого мышления» [8], что и было продемонстрировано выше. Такие отношения объектов, но с учетом еще пространственных зависимостей, формируют отношения (в, на, перед, ...). Семантика таких отношений настолько частотна, что ушла в синтаксис. Аналогично и по временным.

Выводы. Абстрагирование до отношений объектов-переменных не является врожденной способностью, а является следствием процесса обучения и взаимодействия с окружающей средой. Такая трактовка позволяет не только отображать существующие явления, но и моделировать новые. Отношения управления и иерархии, вступая в различные комбинации, позволяют формировать язык. Причем, язык этот допускает изменчивость. Сигналом для переопределения служит рост неоднозначности.

СПИСОК ИСТОЧНИКОВ

1. Хомский, Н. Синтаксические структуры / Н. Хомский // Новое в лингвистике. – 1962. – №2. – С. 412-527.
2. Лотман, Ю. М. Статьи по семиотике культуры и искусства / Ю. М. Лотман. – СПб: Академический проект, 2002. – 544 с.
3. Фреге, Г. Логика и логическая семантика: сб. трудов / Г. Фреге. – М: Аспект Пресс, 2000. – 512 с.
4. Булыгин, В. В. Максимизация коэффициента однозначности для объектов, состоящих из множества признаков / В. В. Булыгин // Вестник Донецкого национального университета. Серия Г: Технические науки. – 2022. – № 4. – С. 34-38. – EDN JSVIGO.
5. Эшби, Р. Введение в кибернетику / Р. Эшби. – М: Изд-во иностр. лит, 1959. – 432 с.
6. Глухов, В. П. Основы психолингвистики: учеб. пособие / В. П. Глухов. – М: АСТ: Астрель, 2005. – 351 с.
7. Бурлак, С. Происхождение языка: Факты, исследования, гипотезы. / С. Бурлак. – М: Альпина. нон-фикшн, 2024. – 464 с.
8. Турчин, В. Ф. Феномен науки: Кибернетический подход к эволюции / В. Ф. Турчин. – М: Изд-во ЭТС, 2000. – 368 с.

Поступила в редакцию 08.10.2024 г., рекомендована к печати 22.10.2024 г.

CYBERENIC APPROACH TO UNDERSTANDING LANGUAGE

Bulygin V.V.

The purpose of the article is to express what language is and how it arises. An approach from the point of view of cybernetics. It is demonstrated that the search for objects can be carried out by means of reinforcement learning. Such objects, maximized by the coefficient of uniqueness, can enter into relationships. Combinations of basic management relationships and hierarchies make it possible to build proposals. Which allows you to model reality and serve as a means of communication.

Keywords: object, hierarchy, management, coefficient of unambiguity, language, sentence.

Булыгин Владимир Викторович
Инженер ПАО «Ростелеком»,
Российская Федерация, г. Новосибирск
E-mail: bvv2311@mail.ru

Bulygin Vladimir Viktorovich
Engineer at the Rostelecom,
Russian Federation, Novosibirsk.

СИСТЕМНЫЙ АНАЛИЗ ФУНКЦИОНАЛЬНЫХ ОСОБЕННОСТЕЙ ПРЕДПРИЯТИЯ КАК ОБЪЕКТА УПРАВЛЕНИЯ

© 2024 *Дмитрюк Т.Г., Зори С.А.*

Выполнен системный анализ производственно-логистической деятельности торгового-промышленного предприятия для разработки специального математического обеспечения системы принятия решений с целью повышения эффективности управления.

Ключевые слова: системный анализ, производство, логистика, система принятия решений.

Введение. Задача повышения экономической эффективности деятельности торгового-промышленного предприятия требует проведения системного анализа структуры производственной деятельности объекта управления и движения материальных ресурсов и готовой продукции в логистической системе предприятия. Для решения этой проблемы на основе правил системного анализа необходимо разработать инструментарий специального математического обеспечения для системы управления с целью принятия эффективных управленческих решений.

Анализ литературных источников. В [1] статье рассматриваются основные проблемы и вопросы, связанные с управлением запасами, оборотными средствами предприятия, а также анализируются и предлагаются основные методы совершенствования методов управления запасами с целью повышения эффективности использования оборотных средств предприятия.

В [2] обоснована разработка и реализация управленческих решений в области повышения эффективности использования оборотных средств предприятия, которые базируются на результатах анализа структуры и динамики оборотных средств, показателей их оборачиваемости.

В статье [3] проводится исследование системы управления уровнем запасов продукции торгового предприятия. Осуществлена характеристика модели управления запасами в рамках всего предприятия. Подробно рассмотрено формирование групп товарного ассортимента. Приведено обоснование выбора системы управления запасами в зависимости от группы.

Модели и методы управления многономенклатурными запасами с учётом производства продукции в условиях постоянного и случайного спроса рассмотрены в работе [4].

Основные аспекты имитационного моделирования и задачи имитационного моделирования в управлении производством изложены в [5].

В работе [6] исследуется применение регрессионных и нейросетевых моделей в задачах разработки производственно-сбытовых систем нового поколения. В [7] рассматриваются подходы к прогнозированию спроса на предприятии с помощью нейросетевых моделей и линейной регрессии в стратегии принятия решения при управлении запасами.

В [8] изложены основные понятия и определения в области транспортной логистики, рассмотрены задачи транспортно-экспедиционной деятельности и характеристики видов транспорта, раскрыты особенности правил перевозок грузов.

В [9] представлены современные проблемы транспортно-экспедиционного обслуживания потребителей, охарактеризованы методы и приемы организации организации перевозки и перемещения материальных потоков груза.

Авторами [10] уточнено понятие «транспортно-логистическое обслуживание потребителей» как целенаправленной деятельности по удовлетворению потребностей потребителей, связанных с перемещением грузов и (или) пассажиров в пространстве и во времени с применением транспортных средств, предоставлением необходимого комплекса услуг соответствующего качества в необходимом количестве в указанное место и в точно назначенное время по приемлемой цене. Сформирована система интегрированных процессов транспортно-логистического обслуживания потребителей, представленная в виде матрицы процессов, характеризующих материальные, информационные и финансовые логистические потоки, а также этапы обслуживания, обоснованных с позиций потребителей.

В [11] описаны новые модели и методы выбора оптимального размера партии поставок при решении задач управления запасами. Авторами изложены основные подходы к оптимизации решений о транспортном обеспечении поставок при управлении запасами одной номенклатуры товара и для многономенклатурных поставок. Представлены особенности таких решений в формате реальных для практики систем управления запасами партионных грузов. Рассматриваются проблемы эффективной организации поставок сразу несколькими транспортными компаниями, включая обоснование уровня дисконта на стоимость таких поставок.

Авторами [12] рассмотрены методы и модели системного анализа, которые обусловлены задачами выбора наилучших решений при многих критериях в цепях поставок и для систем логистики. Представлены традиционные подходы к определению наилучших решений при оптимизации систем логистики, а также так называемые обобщенные критерии выбора и возможности их использования для устранения феноменов неадекватного выбора. Приведены модификации традиционных критериев выбора, позволяющие адаптировать выбор к предпочтениям лица, принимающего решения.

В [13] авторами представлены концепции рассмотрения «больших» систем как многоуровневых (иерархических), заложена основа математической формализации теории управления в больших системах, построенных по иерархическому принципу.

Постановка задачи исследования. На основе методов системного анализа определить функциональные связи структурных элементов производственной деятельности и логистических операций торгово-промышленного предприятия для пользователей разрабатываемой системы управления. Это позволит осуществить постановку задачи оптимального планирования и представить её в виде функционала цели.

Основное содержание и результаты работы. Для реализации поставленной задачи необходимо: выделить организационную структуру предприятия; провести анализ и описать структурные уровни управления; определить специфику портфеля заказов; установить виды логистических операций распределения и сбыта произведенной продукции.

Основной целью системного анализа производственной деятельности предприятия является определение уровня его обеспеченности ресурсами и выявление резервов повышения эффективности их использования. Пользователями разрабатываемой системы являются менеджеры по производству и финансовые менеджеры в задаче выбора модели покрытия потребности в оборотных средствах.

В качестве эталона предприятия третьего уровня производственной иерархии в настоящей работе рассматривается торгово-промышленное предприятие ООО «Донецкий пивоваренный завод» (ДПЗ) как объект управления, который является крупнейшим производителем пивоваренной продукции в регионе.

Согласно методике системного анализа, для торгового предприятия поставлены и решены следующие задачи [14]:

- провести анализ характеристик номенклатуры сырья и готовой продукции, используемой в производственной деятельности, условия её поставок на предприятие, установить взаимосвязь между характеристиками заказов и сырья;
- определить перечень поставщиков сырья и материалов;
- провести анализ характеристик заказов на произведённую продукцию торгового предприятия как объекта управления;
- сформировать раздел базы данных, характеризующий потребителей готовой продукции;
- исследовать структуру производственной деятельности как основу функциональных связей;
- установить совокупность логистических операций производственного цикла торгового предприятия;
- оценить эффективность существующих методов управления предприятием.

Анализ номенклатуры продукции, которую производит завод с 2019 года (условно постоянная), позволил составить перечень из 25-ти наименований. При этом предприятием производится и разливается пять сортов пива, суммарные годовые объёмы производства которого представлены в табл. 1.

Таблица 1. Суммарные объёмы выпуска продукции ООО «ДПЗ» по видам пива

№	Наименование вида пива	Объём реализации, дал (2019 г.)
1	Пиво «Жигулёвское»	497329
2	Пиво «Легенда Донбасса»	596540
3	Пиво «Медведь»	317400
4	Пиво «Кружка Свежего»	61687
5	Пиво «Добрый Шубин»	155718
	Суммарный выпуск	1628674

Торговое предприятие производит продукцию, спрос на которую не является постоянным в течение года. Наибольшее потребление пива приходится на май-сентябрь, а также на праздничные дни. Следовательно, и в выпуске готовой продукции отмечаются сезонные колебания и амплитудные скачки производства и отгрузки продукции со складов в праздники. Это обуславливает большее потребление сырья и материалов (рис. 1).

Оценка положительной динамики развития производственного потенциала ООО «ДПЗ» осуществляется с помощью системного анализа структуры производственной деятельности и взаимодействия элементов системы предприятия в целом.

Колебания показателей экономической деятельности определяют необходимость оценки логистических операций на предприятии и разработки стратегии повышения их эффективности. Исследование структуры управления организацией и выделение пользователей разрабатываемой системы определяет необходимость создания контура системы управления логистическими операциями.

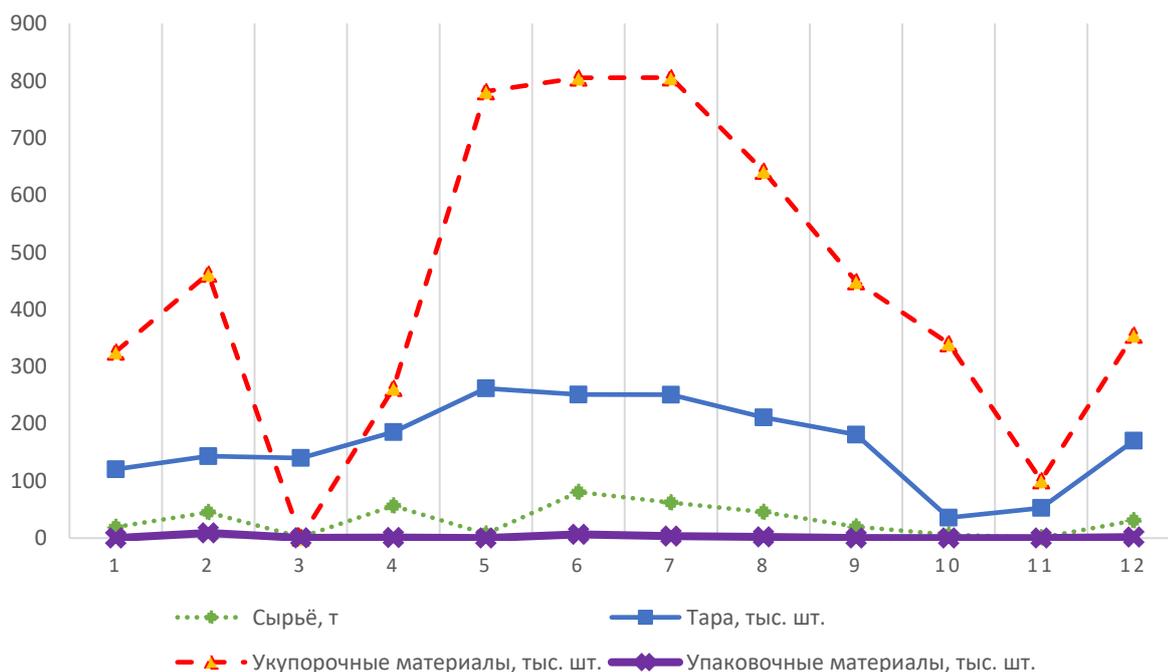


Рис. 1. Сезонные потребности предприятия в сырье и материалах (2019 г.)

Организационная структура предприятия и схема основного производства приведены на рис. 2 и рис. 3. Организационная структура предприятия состоит из трёх контуров управления (рис. 3). Согласно результатам проводимого системного анализа хозяйственно-экономической деятельности предприятия, пользователями разрабатываемой системы управления являются следующие структурные подразделения:

- службы: закупок и планирования (2.6), качества (2.12), пивоварения (2.13), розлива (2.14), продаж (2.15);

- отделы: планирования (3.8), материально-технического обеспечения (3.9) и сбыта (3.19), парокотельный цех (3.13), а также склад товаро-материальных ценностей (3.10), производственная лаборатория (3.17) и участок обслуживания службы розлива (3.18).

Основное производство продукции завода контролируется в следующих службах, подлежащих анализу (рис. 2):

- качества,
- пивоварения,
- розлива.

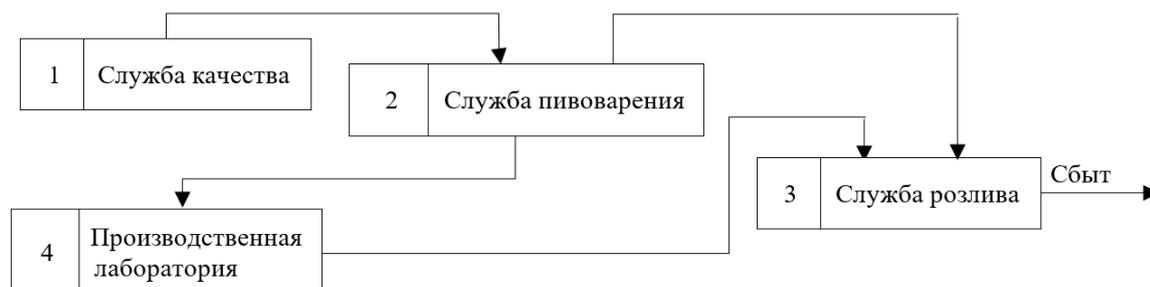


Рис. 2. Схема основного производства ООО «ДПЗ»

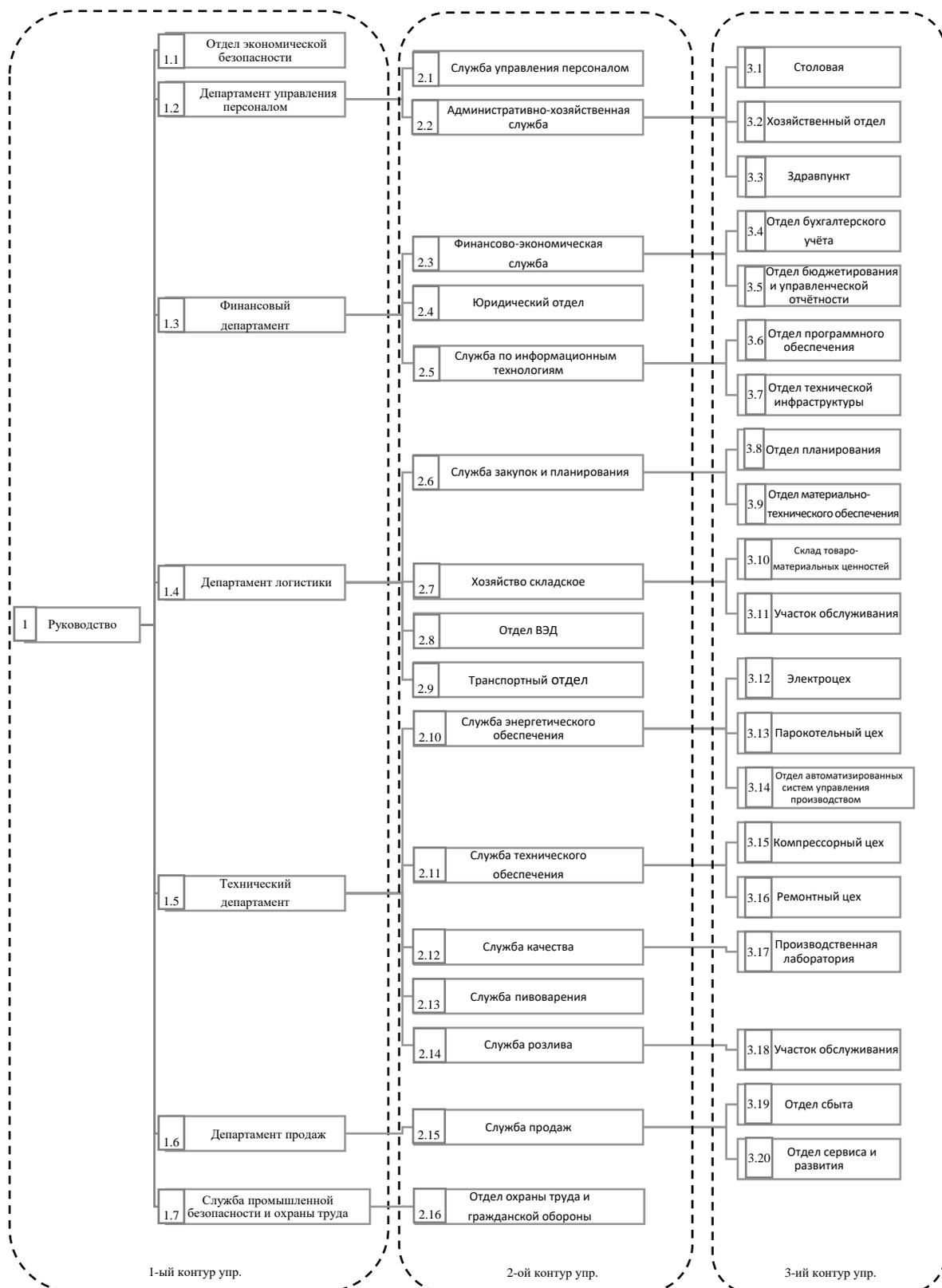


Рис. 2. Организационная структура ООО «ДПЗ»

Служба качества (рис. 2, блок 1) производит оценку качества на всех этапах производства продукции в соответствии с ГОСТами по общепринятым международным методикам. Исследование качества используемого сырья, компонентов пива на стадии производства и готовой продукции осуществляется в соответствии с показаниями производственной лаборатории (блок 4). Лаборатория получает продукты из службы пивоварения (блок 2), анализирует их и передаёт обратно в службу (блок 2) результаты исследований с рекомендациями. Далее сваренное пиво поступает в службу розлива (блок 3). Разлитая, укупоренная, промаркированная продукция завода подлежит последующему сбыту.

В службе пивоварения (рис. 2, блок 2) осуществляются следующие процессы.

1. Подготовка сырья к использованию и непосредственно технологические процессы приготовления хмельных напитков низового брожения, т.е. процесс отбора солода, хмеля, ячменя и дрожжей для разных сортов пива и смешивание компонентов по заводской рецептуре.

2. В парокотельном цеху (рис. 3, блок 3.13) производится фильтрация и умягчение питьевой воды, подготавливаемой для пивоварения, мощностью очистки 175 м³/сут.

3. Далее службой пивоварения осуществляется варка суслу (занимает около 8 ч). Сваренное пиво проходит процесс созревания в бродильных резервуарах (около 12 сут.).

4. Готовый продукт пивоварения поступает в фильтрационные резервуары и проходит фильтрацию (занимает 1 ч), которая выполняется кизельгурами трёх типов зернистости (крупный, средний, мелкий). Сначала пиво продавливают через слой крупного кизельгура, задерживающего наиболее крупные частицы, далее в процессе фильтрации в поток пива добавляется средний, а затем и мелкий кизельгур. Такая технология делает фильтрацию особенно качественной и, одновременно, сохраняет при этом как можно дольше работоспособность слоя. Обновление фильтрующего слоя производят без повторного использования. Такая технология позволяет повысить прозрачность пива и увеличить срок годности.

После фильтрации напитки поступают в цех службы розлива (рис. 3, блок 2.14). Во время спадов потребления (см. рис. 1) пиво хранится в цилиндро-конических танках (ЦКТ), время хранения определяет температурный режим. В технологии розлива используется тара объёмом 0,5-50 л и укупорочно-маркировочные материалы различного назначения.

На этапе розлива на участке обслуживания (рис. 3, блок 3.18) тара распаковывается, моется и проходит химическую обработку. Чистая тара устанавливается на конвейерные ленты, где производится автоматический розлив пива, укупорка тары (на стеклянную - кроненпробка, на пэт - колпачок, на кеги - крышка-пломба), отбраковывается некачественный товар, наклеиваются этикетки, контрэтикетки и кольеретки.

После маркировки на конвейере счётчик бутылок подсчитывает количество единиц готовой продукции. Среднемесячные производственные мощности по розливу готовой продукции приведены по видам пива на рис. 4.

После подсчёта количества единиц изготовленной продукции производится её расфасовка в гофролотки стрейч-плёнкой и в гофроподдоны термоусадочной плёнкой соответственно.

По завершении всех операций на основном производстве завода (рис. 3, блок 2.14) продукция отгружается на склад товаро-материальных ценностей (блок 3.10) с последующей реализацией дистрибьютору. Сбыт продукции осуществляется отделом сбыта (блок 3.19). Структура и функционирование отделов и служб основного производства ООО ДПЗ подчинены производственной программе и обеспечивают эффективную её реализацию.

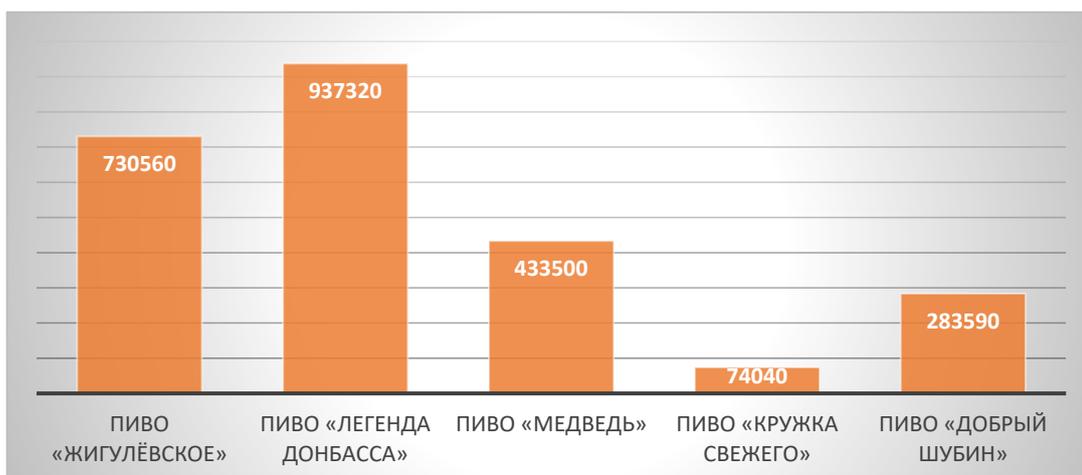


Рис. 4. Мощность розлива ООО «ДПЗ» по сортам пива за 2019 г., шт./мес.

В реальных условиях производственной программой осуществляется планирование производства на ООО ДПЗ. В соответствии с перечисленными пользователями структурная схема представлена на рис. 5.

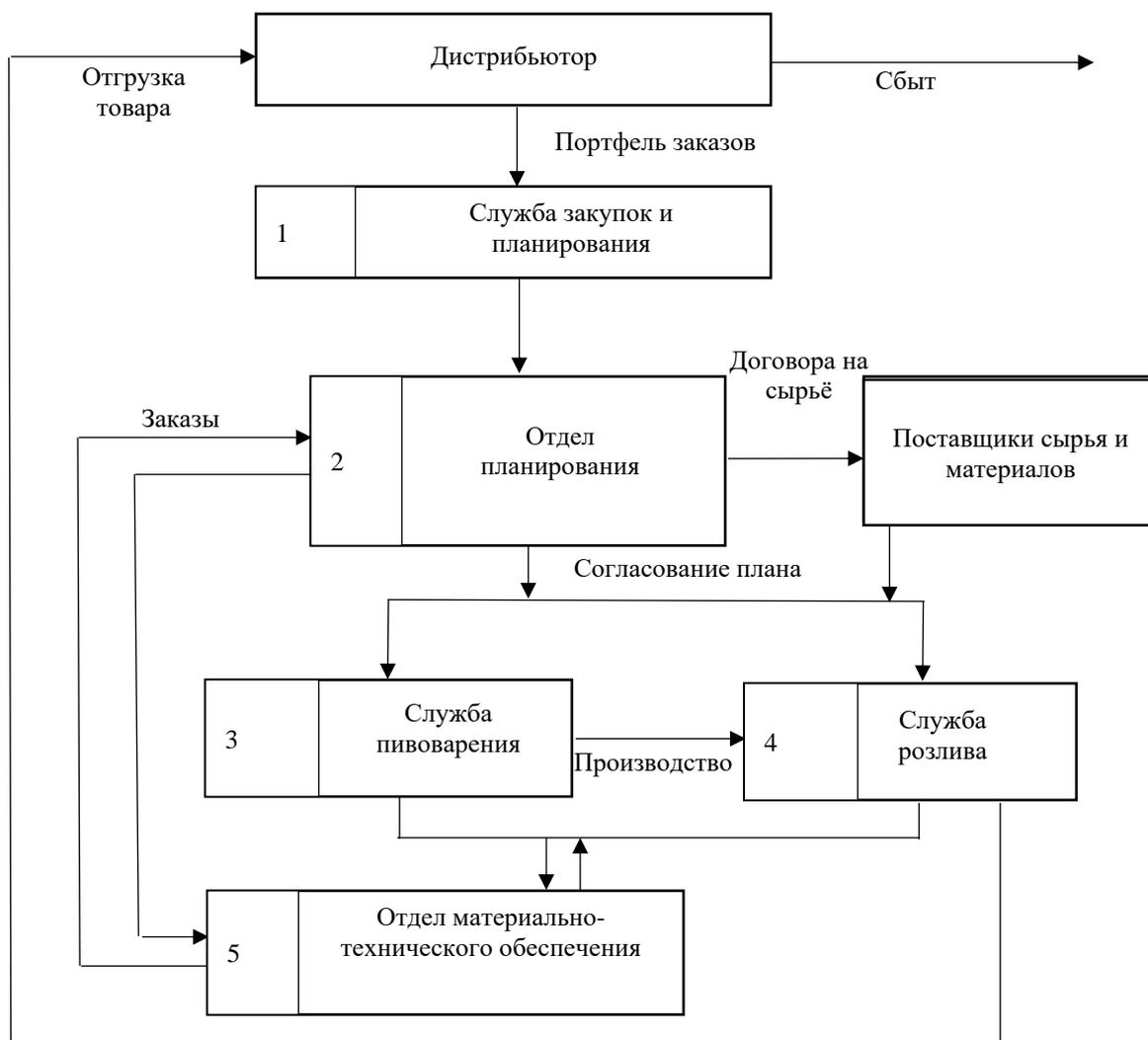


Рис. 5. Структура планирования производства на ООО «ДПЗ»

Исходная информация для составления производственной программы определена совокупностью долгосрочных договоров с дистрибьютором ООО «Фуд Лайн Трейдинг-2014». Заключённые договора и их копии службой закупок и планирования (рис. 5, блок 1) передаются дистрибьютору. В конце текущего года договора поступают в данную службу, которая в дальнейшем занимается планированием поставок продукции в соответствии с заключёнными договорами дистрибьютору, а он, в свою очередь, занимается отгрузкой товара потребителям. В отделе планирования (рис. 5, блок 2) по совокупности договоров на следующий месяц формируется годовой портфель заказов торгового предприятия с разбивкой по месяцам (ежемесячно уточняется в зависимости от условий сезонности - наибольшие объёмы спроса приходятся на май-сентябрь, а также на праздничные дни).

В свою очередь из блока 1 (рис. 5) службой подаётся запрос на отдел материально-технического обеспечения (блок 5) о возможности выполнения планов по выпуску продукции. Для портфеля заказов отдел составляет плановую калькуляцию себестоимости продукции, что даёт возможность сформировать оптовые цены. Годовое планирование включает в себя составление годовой производственной программы, которая будет передана в службы пивоварения и розлива (блоки 3 и 4), а также в отдел планирования (блок 2). Между последними тремя указанными блоками происходит механизм согласования планов. Пересчёт калькуляции производится в блоке 2 только в случае изменения региональной ценовой политики на законодательном уровне.

В службах 3 и 4 (рис. 5) составляются суточные программы производства (согласованные с блоком 2) согласно транзитным нормам (минимального количества продукции, одновременно отгружаемого поставщиком одному грузополучателю для обеспечения полной загрузки транспортных средств) и составляются графики текущих плановых ремонтов. По этим программам в блоке 5 рассчитываются графики потребления сырья, упаковочных и укупорочных материалов. Задания ежемесячно планируются для каждой из четырёх смен работников с учётом суточных программ и невязок план-факт.

Кроме того, потребности в сырье, упаковочных и укупорочных материалах службы 3 и 4 согласовывают с отделом 2, которые также поступают в блок 5 (см. рис. 5). Между блоком 2, по согласованию с блоком 5, и Поставщиками заключаются долгосрочные договора с годовыми объёмами поставок. В блоке 5 ежемесячно формируется заказ по поставкам сырья, материалов и комплектующих и передаётся в соответствующий блок.

Необходимое сырьевое и материально-техническое обеспечение от блока Поставщиков поступает на производство в блоки 3 и 4 (рис. 5). Осуществляются циклы пивоварения, укупорки, розлива и упаковки готовой продукции. Одновременно блоком 4 формируются наряды на транспортировку и складирование товара и последующее согласование с блоком «Потребитель» отгрузки и реализации произведённой продукции.

Выявленные функциональные связи элементов логистической системы в производственной деятельности ООО «Донецкий пивоваренный завод» и анализ задач планирования позволяют сделать следующие выводы.

1. Планирование производства и будущих затрат опирается на показатели предыдущих периодов и не учитывает сезонные колебания спроса на продукцию.

2. Надёжность гарантированных поставок продукции потребителям в условиях нестабильной экономической ситуации без создания такого количества запасов сырья и произведённой продукции, которая будет необходима для непрерывного бесперебойного функционирования производственного процесса, невозможно обеспечить без решений системы управления, создаваемые на основе системного анализа экономико-производственной ситуации.

Повышение эффективности управления движением материальных ресурсов и готовой продукции в логистической системе предприятия является приоритетной задачей объекта управления при осуществлении торговой деятельности предприятий подобного типа. Следовательно, необходимо определить взаимосвязи элементов в системе управления логистическими процессами завода для дальнейшего осуществления постановки задач планирования доставки и сбыта готовой продукции. Пользователями системы управления логистикой являются менеджеры по планированию производства и сбыта продукции.

Процесс сбыта произведённой продукции, осуществляемый между предприятием и конечным потребителем за единицу квантования (месяц), сопровождается включением транспортно-экспедиционных компонент. Транспортное обслуживание перевозок является элементом многоуровневой системы взаимосвязанных звеньев структуры объекта управления [9].

Логистическая система ООО «ДПЗ» как объекта управления характеризуется как развитая транспортная структура, используемая для сбыта пивоваренной продукции в регионе. Анализ характеристик логистической системы объекта управления предусматривает [15]:

- анализ взаимосвязей предприятия с контрагентами — поставщиками сырья и материалов и потребителями продукции завода;
- классификацию контрагентов по типам перевозимого груза;
- анализ маршрутов для планирования грузоперевозок ООО «ДПЗ»;
- анализ характеристик автотранспортного парка, обслуживающего грузоперевозки предприятия.

Цель системного анализа логистики — определение механизма обслуживания заказов на продукцию пивзавода транспортными средствами. Исследования показали, что существуют договора потребителей (заказчиков) с дистрибьюторами пивзавода сроком на год и более. В конце предыдущего планового периода (месяц) заказчики (потребители) присылают дистрибьютору спецификацию, где указаны виды продукции, её количество, цена и общая стоимость заказа. В соответствии со спецификацией, не позднее 25 числа предыдущего месяца, заказчик осуществляет предоплату продукции дистрибьютору, который переводит положенные суммы на расчётный счёт завода. Схема взаимодействия контрагентов с объектом управления представлена на рис. 6.

Следует учитывать тот факт, что пункты назначения для сбыта произведённой продукции на предприятие передаёт дистрибьютор. Отдел планирования (рис. 6, блок 2) разрабатывает производственную программу по видам, количеству и расфасовке пива и определяет задания на транспортировку грузов заказчикам. Плановое задание поступает менеджерам отдела сбыта и транспортного отдела (блок 4), где определяется вид транспорта, его количество, грузоподъёмность и длины маршрутов. Заказчик, получивший продукцию на основании акта приёмки, сообщает в службу продаж (блок 8) и дистрибьютору о количестве поставленной продукции.

По сформированным в службе закупок и планирования (блок 2) заказам, согласованным с поставщиками (блок 5) на поставки сырья, тары, укупорочных и упаковочных материалов для пивоварения и розлива готовой продукции, используемых на участках основного производства предприятия (блок 1), транспортными средствами поставщиков осуществляется доставка грузов на склад товаро-материальных ценностей (блок 7).

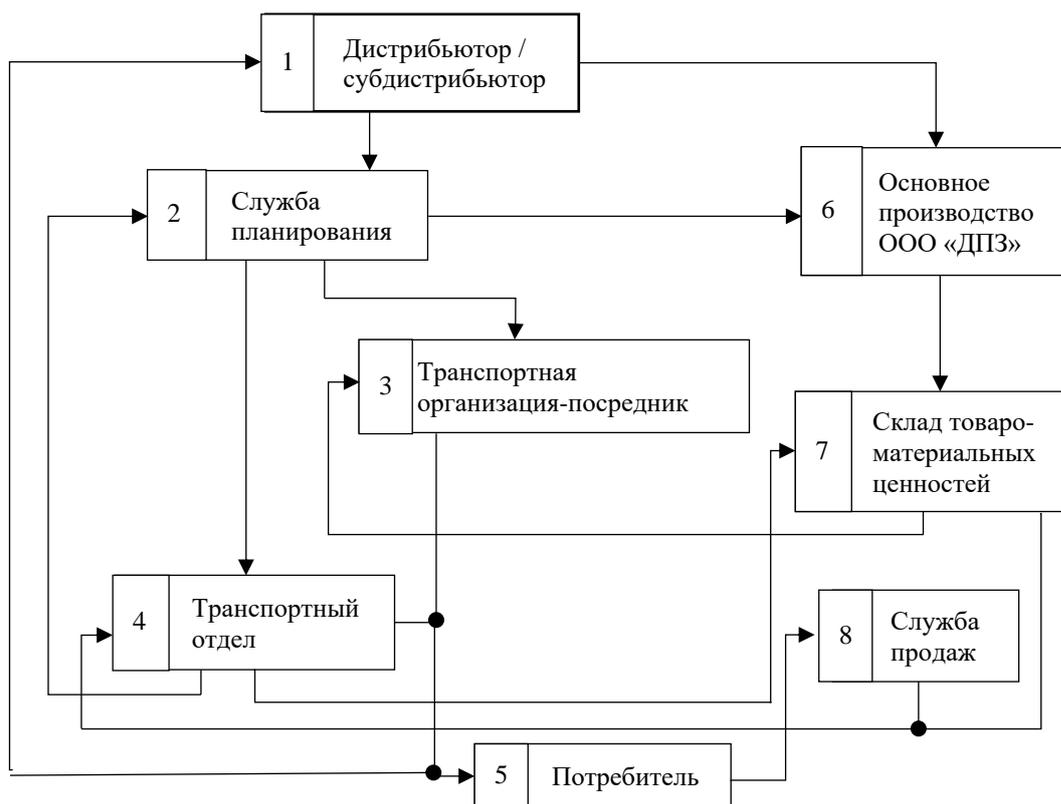


Рис. 6. Схема взаимодействия элементов логистической системы ООО «ДПЗ»

Анализ организации логистики грузоперевозок показал, что доставка пива потребителям транспортными средствами в соответствующие пункты назначения на объекте управления осуществляется в следующем порядке.

Служба закупок и планирования (рис. 6, блок 2) департамента логистики составляет транспортные характеристики груза, исходя из его физико-механических и физико-химических свойств, проявляемых в процессе его транспортировки. Однако, в управлении логистикой пивоваренного завода значение имеют только физико-механические свойства пивной тары. К последним относят хрупкость, которая определяет условия хранения пива на складе, погрузочно-разгрузочные работы и транспортировку произведённой продукции. Материал тары определяет безопасность (токсичность) материала по влиянию на продукт. Следует отметить, что в таре тёмно-зелёного и красно-коричневого цвета коэффициент поглощения ультрафиолетовых лучей выше, чем в других видах, это позволяет продлить сроки хранения пива и защитить продукт от фотохимического разрушения его биохимических компонентов [16].

После этого служба закупок и планирования (блок 2) составляет сетевую модель транспортной сети, учитывая портфель заказов на реализацию продукции. В качестве показателей — переменных, характеризующих сетевую модель, указываются номера вершин, в узлах которых располагаются пункты отправления и пункты назначения. Для них проводятся вычисления длины звеньев между соседними вершинами, определяются объёмы продукции, отгружаемой заказчику, а затем — осуществляется выбор транспортных средств, выполняющих перевозки. В соответствии с этим осуществляется группировка пунктов назначения по одному маршруту исходя из требования: «один маршрут — один день», с полной загрузкой грузоподъёмности транспорта. Далее обработанная информация передаётся в транспортный отдел (блок 6).

Диспетчерами транспортного отдела (блок 4) на основании сетевой модели выбираются тип и модель собственного подвижного состава и, если есть необходимость, то производится заказ в нужном количестве соответствующего наёмного автотранспорта, по согласованию с транспортной организацией-посредником (блок 3), применительно к дорожно-транспортным условиям и вершинам пунктов назначения для выбранного типа груза. Согласно заданию на перевозку грузов выбирается тип погрузочно-разгрузочного средства. Погрузка и транспортировка готовой продукции, её рациональная укладка в выбранный тип транспорта осуществляется из склада товаро-материальных ценностей (блок 7) в текущем суточном режиме.

Произведённая продукция, согласно заданию дистрибьютора, доставляется со склада товаро-материальных ценностей (блок 7) завода потребителю (блок 5) собственными и наёмными транспортными средствами по рассчитанным маршрутам в пункты назначения, предварительно поставив в известность посредников об отгрузке. Оплата транспортной организации-посреднику производится по факту, с учётом объёмов перевезённого груза, длины пробега транспорта, а также своевременности и надёжности доставки. Комплексным обслуживанием и сопровождением клиентского портфеля занимается служба продаж (блок 8), которая в рамках установленных предприятием задач взаимодействует непосредственно с потребителями продукции, а также с транспортным отделом (блок 4) и службой планирования предприятия (блок 2).

Системный анализ перечисленных особенностей логистической структуры объекта управления позволяет осуществить детализацию характеристик движения сырья и материалов. Эффективное управление перевозками грузов осуществляется посредством рационального планирования использования транспортных средств для своевременных, ритмичных поставок конечным потребителям. Для анализа структуры логистической деятельности пивзавода определяются пункты географического расположения потребителей, требующих обобщения в соответствии с длиной маршрутов. Задача маршрутизации грузоперевозок решается при условии, что в отдел планирования поступает информация о транспортной сети сбыта продукции: количество и удалённость складов, наличие дополнительных точек сбыта по маршруту следования. Для решения задачи маршрутизации введём обозначение переменных и опишем физический смысл, необходимых для расчётов моделей доставки пива (таблица 2).

Таблица 2. Физический смысл и обозначение переменных маршрутизации

Обозначение переменной	Наименование переменной маршрутизации
a_m	Пункт транспортной сети (потребитель), $m = \overline{1,28}$
l_m	Удалённость потребителя a_m от предприятия
A_k	Конечный пункт назначения по маршруту k , $k = \overline{1,15}$
M_k	Маршрут, соединяющий предприятие с конечным потребителем A_k
SM_k	Длина маршрута как совокупная длина многозвенных расстояний между точками транспортной сети, т. е. километраж за сутки для одного автомобиля за один рейс

Взаимосвязь автомобилей собственного и наёмного транспортных парков предприятия, доставляющих готовую продукцию потребителям, с маршрутами определяет привязку автомобилей к маршрутам и загрузке. Маршруты рассматриваются как комплекс точек назначения, содержащий промежуточные точки.

Груз развозится согласно путевым листам грузового транспорта по маршрутам, составленным службой закупок и планирования и выданным транспортным отделом. Водители осуществляют доставку пива самостоятельно по всем направлениям, кроме a_{28} (г. Луганск), куда груз доставляется в сопровождении экспедитора. Несколько точек, связанных последовательно, образуют маршруты, определённые следующим образом (табл. 3).

Таблица 3. Множество маршрутов перевозок грузов ООО «ДПЗ»

Обозначение переменной маршрута	Множество пунктов транспортной сети, включённых в маршрут	Наименование переменной маршрута	Длина маршрута, км
$M_1(A_1)$	$\{a_1, a_2\}$	Донецк-1	2,2
$M_2(A_2)$	$\{a_3, a_4\}$	Донецк-2	2,9
$M_3(A_3)$	$\{a_5\}$	Донецк-3	4,7
$M_4(A_4)$	$\{a_6\}$	Донецк-4	10,3
$M_5(A_5)$	$\{a_7, a_8\}$	Донецк-5	16,1
$M_6(A_6)$	$\{a_9, a_{10}\}$	Донецк – Макеевка-1	7,5
$M_7(A_7)$	$\{a_{11}, a_{12}\}$	Донецк – Макеевка-2	18,3
$M_8(A_8)$	$\{a_{13}, a_{14}\}$	Донецк – Макеевка-3	10,0
$M_9(A_9)$	$\{a_{15}\}$	Донецк – Докучаевск	38,0
$M_{10}(A_{10})$	$\{a_{16}, a_{17}, a_{18}\}$	Донецк – Новоазовск	120,0
$M_{11}(A_{11})$	$\{a_{19}, a_{20}, a_{21}\}$	Донецк – Амвросиевка	77,4
$M_{12}(A_{12})$	$\{a_{22}, a_{23}, a_{24}\}$	Донецк – Снежное	78,2
$M_{13}(A_{13})$	$\{a_{25}\}$	Донецк – Горловка	78,0
$M_{14}(A_{14})$	$\{a_{26}, a_{27}\}$	Донецк – Дебальцево	85,0
$M_{15}(A_{15})$	$\{a_{28}\}$	Донецк – Луганск	165,0

В таблице 3 представлен перечень маршрутов $M = \{M_1, \dots, M_{15}\}$ в соответствии с приведёнными конечными пунктами назначения. Порядок пунктов в маршруте указан по порядку следования грузового автомобиля. Маршруты составлены менеджерами таким образом, чтобы транспортировка груза осуществлялась одним видом транспортного средства ежедневно. Груз доставляется в конечные пункты по маршруту (в обратном направлении автомобиль движется порожним). Произведённая продукция ООО «ДПЗ» доставляется автотранспортом согласно договорам с дистрибьютором (транспортировка пива в г. Луганск осуществляется в соответствии с договорами с субдистрибьютором «Каменный дом») по адресам $a_m, m = \overline{1,28}$. На условии требования: «один маршрут — один день» рассматривается укрупнение по пунктам доставки с указанием среднемесячных показателей объёмов поставок грузов потребителям одним транспортным средством в конечные пункты назначения.

Для выполнения плана перевозок (доставки всего объёма произведённой продукции по всем точкам потребления на каждом маршруте) требуется распределение транспортных средств по выделенным маршрутам. Для этой цели используется информация о технико-эксплуатационных характеристиках транспортных средств, являющаяся исходной в работе менеджеров транспортного отдела при выборе типа автомобиля, следующего по маршруту. Однако, учёт движения и расходная составляющая осуществляются этими менеджерами в отдельной системе.

Для решения поставленной задачи был исследован парк собственных и наёмных транспортных средств, используемых для транспортировки грузов, классифицировать их по функциональному назначению, а также установлены технико-экономические показатели их работы. Транспортное обслуживание логистической деятельности на ООО «ДПЗ», которое должно удовлетворять спрос потребителей, основано на существующем подвижном составе, представленном собственным автопарком и средствами организации-посредника ФЛП «Пичахчи» (наёмными). Технические характеристики транспорта, обслуживающего предприятие, представлены в табл. 4.

Таблица 4. Техничко-эксплуатационные характеристики транспортного парка ООО «ДПЗ»

Принадлежность	Функциональное назначение транспортного средства	Наименование транспортного средства	Кол-во ед.	Грузоподъёмность, т	Средний пробег за мес.
Собственный	Транспортировка готовой продукции с участка по розливу на склад товаро-материальных ценностей	Автопогрузчик JUNGHEINRICH	2	1,6	50 моточасов
		Автопогрузчик JUNGHEINRICH	2	2,5	50 моточасов
		Автопогрузчик JUNGHEINRICH	2	3	50 моточасов
	Подъём или спуск грузов на складе товаро-материальных ценностей	Электроштабелёр JUNGHEINRICH	1	1,2	50 моточасов
	Доставка сырья и материалов, отгрузка готовой продукции потребителю	Автомобиль Scania R730	1	10	3000 км
		Автомобиль Scania R420	4	14,6	4000 км
Автомобиль Scania A702		4	20	7000 км	
Наёмный	Отгрузка готовой продукции потребителю	Автомобиль Scania R420	1	5	2000 км
		Автомобиль Scania R730	1	10	3000 км
		Автомобиль Scania R420	1	14,6	4000 км
		Автомобиль Scania A702	1	20	7000 км

При анализе подвижного состава было установлено, что перевозки грузов осуществляются как на территории завода — автопогрузчиками и средствами подъёма/спуска, так и за пределами предприятия — доставка пива потребителям грузовыми автомобилями по правилу: «один маршрут — один автомобиль».

Эксплуатация подвижного состава производится транспортным отделом (рис. 6, блок 4) по планам, разработанным департаментом логистики. Распределение автомобилей осуществляется по маршрутам, характеризующим расположение конечных потребителей перевозок с подбором транспортных средств по необходимым характеристикам. Планирование, организацию и координацию логистики перевозок осуществляет диспетчер транспортного отдела (рис. 6, блок 4) посредством составления маршрутных листов с учётом фактора сезонности работ согласно заказам. При расчёте видов средств погрузки и транспортировки грузов учитываются особенности грузовых и складских работ. Данные таблицы 4 служат менеджерам департамента логистики основанием для распределения автомобилей по маршрутам в соответствии с объёмами заказов (или доставляемых грузов).

Для решения задачи формализации показателей доставки грузов введём обозначение переменных и опишем физический смысл, необходимых для составления моделей, учитывающих технические характеристики транспорта для доставки пива (табл. 5).

Таблица 5. Физический смысл и обозначение переменных технических характеристик транспорта предприятия

Обозначение переменной	Наименование переменной маршрутизации
g	Грузоподъёмность одного автомобиля (т), которая имеет дискретные значения: $g = \{g_1, g_2, g_3, g_4\}$
p	Среднемесячный пробег одной единицы транспорта (тыс. км), который имеет дискретные значения: $p = \{p_1, p_2, p_3, p_4\}$
TSO_{ngp}	Грузовой автомобиль собственного транспортного парка предприятия грузоподъёмностью g , осуществляющий средний пробег за месяц p , количество автомобилей с соответствующими характеристиками $n = 1; 4$
TSH_{ngp}	Грузовой автомобиль наёмного транспортного парка предприятия грузоподъёмностью g , осуществляющий средний пробег за месяц p , $n = 1$
NRM_k	Количество выполненных рейсов по маршрутам за месяц

Согласно технико-эксплуатационных характеристик собственных и наёмных транспортных средств (табл. 5), формализация задачи распределения транспортных средств по маршрутам с плановыми объёмами перевозок продукции пивзавода одной единицей транспорта собственного и наёмного автопарков (табл. 4) будет выполнена по следующим условиям.

1. Грузовой автомобиль собственного транспортного парка предприятия грузоподъёмностью 10 т, осуществляющий средний пробег за месяц 3 тыс. км, выполняет перевозку груза по маршрутам: M_7 (Донецк – Макеевка-2) — 596,34 т и M_8 (Донецк – Макеевка-3) — 477,85 т, в соответствии с условием вида:

$$TSO_{1g_2 p_2} \in \{M_7, M_8\}.$$

2. Два грузовых автомобиля собственного транспортного парка предприятия грузоподъёмностью 14,6 т, осуществляющие средний пробег за месяц 4 тыс. км, выполняют перевозку груза по маршруту M_{10} (Донецк – Новоазовск) с дополнительными выгрузками в г. Старобешево и г. Тельманово — 74,59 т, в соответствии с условием вида:

$$\{TSO_{1g_3 p_3}, TSO_{2g_3 p_3}\} \in M_{10}.$$

3. Грузовой автомобиль собственного транспортного парка предприятия грузоподъёмностью 14,6 т, осуществляющий средний пробег за месяц 4 тыс. км, выполняет перевозку груза по маршруту M_{11} (Донецк – Амвросиевка) с дополнительными выгрузками в г. Харцызск и г. Иловайск — 38,18 т, в соответствии с условием вида:

$$TSO_{3g_3 p_3} \in M_{11}.$$

4. Грузовой автомобиль собственного транспортного парка предприятия грузоподъёмностью 14,6 т, осуществляющий средний пробег за месяц 4 тыс. км, выполняет перевозку груза по маршруту M_{12} (Донецк – Снежное) с дополнительными выгрузками в г. Шахтёрск и г. Торез — 39,02 т, в соответствии с условием вида:

$$TSO_{4g_3 p_3} \in M_{12}.$$

5. Четыре грузовых автомобиля собственного транспортного парка предприятия грузоподъёмностью 20 т, осуществляющие средний пробег за месяц 7 тыс. км, выполняют перевозку груза по маршруту M_{15} (Донецк – Луганск) — 220,99 т, в соответствии с условием вида:

$$\{TSH_{1g4p4}, \dots, TSH_{4g4p4}\} \in M_{15}.$$

6. Грузовой автомобиль наёмного транспортного парка предприятия грузоподъёмностью 5 т, осуществляющий средний пробег за месяц 2 тыс. км, выполняет перевозку груза по маршрутам: М₁ (Донецк-1) — 218,55 т, М₂ (Донецк-2) 195,66 т, М₃ (Донецк-3) — 368,32 т, М₄ (Донецк-4) — 249,35 т и М₅ (Донецк-5) — 263,38 т, в соответствии с условием вида:

$$TSH_{1g1p1} \in \{M_1, M_2, M_3, M_4, M_5\}.$$

7. Грузовой автомобиль наёмного транспортного парка предприятия грузоподъёмностью 10 т, осуществляющий средний пробег за месяц 3 тыс. км, выполняет перевозку груза по маршрутам: М₆ (Донецк – Макеевка-1) — 388,65 т и М₉ (Донецк – Докучаевск) — 42,50 т, в соответствии с условием вида:

$$TSH_{1g2p2} \in \{M_6, M_9\}.$$

8. Грузовой автомобиль наёмного транспортного парка предприятия грузоподъёмностью 14,6 т, осуществляющий средний пробег за месяц 4 тыс. км, выполняет перевозку груза по маршруту М₁₃ (Донецк – Горловка) — 173,58 т, в соответствии с условием вида:

$$TSH_{1g3p3} \in M_{13}.$$

9. Грузовой автомобиль наёмного транспортного парка предприятия грузоподъёмностью 20 т, осуществляющий средний пробег за месяц 7 тыс. км, выполняет перевозку груза по маршруту М₁₄ (Донецк – Дебальцево) с дополнительной выгрузкой в г. Енакиево — 115,28 т, в соответствии с условием вида:

$$TSH_{1g4p4} \in M_{14}.$$

Перечисленные особенности планирования грузоперевозок по разработанным маршрутам согласно заявкам потребителей в пункты назначения позволяют разработать модели прогноза доставки пива всех сортов в каждый пункт назначения по всем маршрутам. Изложенные методические соображения служат основанием для формализации показателей, характеризующих процесс доставки грузов, а технические характеристики делают возможным разработку модели, уравнение которой позволяет осуществлять прогноз загрузки логистической системы на календарный период – месяц.

Выводы. Научная новизна исследования заключается в том, что на основании системного анализа производственной деятельности предприятия выявлены взаимосвязи её элементов, определены пользователи разрабатываемой системы управления. На основе этого возможно сформировать функциональную структуру системы управления предприятиями аналогичного типа. В свою очередь системный анализ технико-экономических показателей логистической системы позволил исследовать механизмы движения запасов в процессе производства и распределения готовой продукции. На основе анализа возможна формализация этих показателей деятельности предприятия в виде математических моделей прогноза и формальной постановки задач управления, т.е. является основанием для создания специального математического и программного обеспечения автоматизированной системы управления предприятием. Это определяет практическую значимость работы.

СПИСОК ИСТОЧНИКОВ

1. Варфоломеев, А. И. Управление запасами как способ повышения эффективности использования оборотных средств / А. И. Варфоломеев // Молодой ученый. – 2019. – № 22(260). – С. 496-499. – EDN AJEQQI.
2. Паймулина, К. А. Анализ оборачиваемости оборотных средств предприятия и разработка мероприятий по оптимизации материальных запасов / К. А. Паймулина, В. Е. Павлович // Экономика и бизнес: теория и практика. – 2020. – № 4-2(62). – С. 175-180. – DOI 10.24411/2411-0450-2020-10321. – EDN HQKFMM.
3. Зиновьева, О. Г. Модель управления запасами торгового предприятия / О. Г. Зиновьева, П. В. Сухина // Экономика и бизнес: теория и практика. – 2019. – № 4-1. – С. 94-97. – DOI 10.24411/2411-0450-2019-10483. – EDN TNUGJR.
4. Новосельцев, В. И. Модель управления производственными запасами в условиях неопределенности спроса на выпускаемую продукцию / В. И. Новосельцев, О. Е. Шугай, Н. Н. Попов // Вестник Воронежского института высоких технологий. – 2019. – № 1(28). – С. 51-52. – EDN TTAGGT.
5. Милехин, И. А. Методы имитационного моделирования и алгоритмы решения задач в управлении производством / И. А. Милехин, М. А. Слепцова, О. Н. Юркова // Вестник современных исследований. – 2021. – № 1-3(39). – С. 9-12. – EDN KUWHZQ.
6. Использование регрессионного и нейросетевого моделирования в производственном мониторинге промышленного предприятия / А. Р. Холова, Ю. С. Вождаева, И. А. Мельницкий [и др.] // Экология и промышленность России. – 2021. – Т. 25, № 5. – С. 58-64. – DOI 10.18412/1816-0395-2021-5-58-64. – EDN ISUZQW.
7. Александров, А. А. Нейросетевое моделирование производственно-сбытовых систем нового поколения / А. А. Александров, Е. Н. Горлачева // Наука и бизнес: пути развития. – 2021. – № 11(125). – С. 64-70. – EDN MTQIQH.
8. Колочева, В. В. Транспортная логистика : учебное пособие / В. В. Колочева, С. А. Максимов, В.А. Назаркина. – Новосибирск: Новосибирский государственный технический университет, 2021. – 88 с.
9. Транспортно-экспедиционная деятельность: Учебник и практикум / М. Г. Григорян, Е. В. Будрина, Е. В. Табачникова [и др.]. – 1-е изд.. – Москва : Издательство Юрайт, 2020. – 369 с. – (Высшее образование). – ISBN 978-5-534-04168-2. – EDN JWCMSV.
10. Криворучко, О. Н. Интегрированные процессы транспортно-логистического обслуживания потребителей / О. Н. Криворучко // Экономика транспортного комплекса. – 2017. – № 29. – С. 86-99. – EDN ZQONTX.
11. Управление запасами: многофакторная оптимизация процесса поставок: учебник для СПО / Г. Л. Бродецкий, В. Д. Герами, А. В. Колик, И. Г. Шидловский. – М.: Издательство Юрайт, 2019. – 322 с.
12. Бродецкий, Г. Л. Системный анализ в логистике. Выбор при многих критериях: учебник / Г. Л. Бродецкий, Д. А. Гусев, О. А. Мазунина. – М.: Издательский центр «Академия», 2018. – 224 с.
13. Месарович, М. Теория иерархических многоуровневых систем : пер. с англ. / М. Месарович, Д. Мако, И. Такахага. – М.: Мир, 1973. – 344 с.
14. Дмитриук, Т. Г. Анализ структуры производственной деятельности предприятия как объекта управления / Т. Г. Дмитриук, С. А. Зори // Проблемы искусственного интеллекта. – 2020. – № 1(16). – С. 37-53. – EDN QHRUGP.
15. Зори, С. А. Анализ характеристик транспортно-логистической системы предприятия / С. А. Зори, Т. Г. Дмитриук // Программная инженерия: методы и технологии разработки информационно-вычислительных систем (ПИИВС-2020) : сборник научных трудов III Международной научно-практической конференции, Донецк, 25–26 ноября 2020 года. Том 1. – Донецк: Донецкий национальный технический университет, 2020. – С. 93-100. – EDN TYFUUT.
16. Стеклотара [Электронный ресурс]. – URL: <https://ru.wikipedia.org/wiki/Стеклотара> (дата обращения 11.09.2024).

Поступила в редакцию 22.10.2024 г., рекомендована к печати 06.11.2024 г.

SYSTEM ANALYSIS OF THE ENTERPRISE FUNCTIONAL FEATURES AS A CONTROL OBJECT

Dmitriuk T.G., Zori S.A.

A system analysis of the production and logistics activities of a commercial and industrial enterprise has been performed to develop special mathematical support for a decision-making system in order to improve management efficiency.

Keywords: system analysis, production, logistics, decision-making system.

Дмитрюк Татьяна Григорьевна

ассистент кафедры программной инженерии
им. Л.П. Фельдмана ФГБОУ ВО «Донецкий
национальный технический университет»,
Российская Федерация, ДНР, г. Донецк.
E-mail: tнаука@lenta.ru

Dmitriuk Tatiana Grigorevna

Assistant at Department of Software Engineering of
Donetsk National Technical University,
Russian Federation, DPR, Donetsk.

Зори Сергей Анатольевич

доктор технических наук, доцент, заведующий
кафедрой программной инженерии
им. Л.П. Фельдмана ФГБОУ ВО «Донецкий
национальный технический университет»,
Российская Федерация, ДНР, г. Донецк.
E-mail: sa.zori1968@gmail.com

Zori Sergei Anatolevich

Doctor of Technical Sciences, Associate Professor,
Head of the Department of Software Engineering of
Donetsk National Technical University,
Russian Federation, DPR, Donetsk.

ДЕТЕКТИРОВАНИЕ ТРЕНДОВ МОДНОЙ ОДЕЖДЫ С ОБУЧЕНИЕМ НЕЙРОСЕТИ

© 2024 *Евдачёва Е.Д., Максименко И.И.*

Прогнозирование модных тенденций является важнейшей задачей как для научных кругов, так и для промышленности. Ранее для решения этой сложной задачи, изучали лишь ограниченные элементы моды с сезонными или простыми узорами, которые вряд ли могли выявить реальные тенденции моды. Стремясь к прогнозированию, эта статья предназначена для исследования четких тенденций в мире моды для определенных групп пользователей. В работу был взят крупномасштабный набор данных о модных тенденциях (FIT). Так же для более эффективного моделирования временных рядов данных была взята модель рекуррентной сети, которая использует преимущество от нейронных сетей при моделировании временных рядов. Обширные эксперименты демонстрируют, что предложенная модель КЕРНА может эффективно улавливать сложные закономерности объективных элементов моды, следовательно, делая предпочтительный прогноз модных тенденций.

Ключевые слова: прогнозирование модных тенденций, анализ моды, прогнозирование временных рядов.

Введение. Карл Лагерфельд утверждал, что мода заключается в умении меняться. Поэтому для специалистов, которые хотят следовать этим изменениям, важно прогнозировать модные тенденции. Это помогает модным компаниям создавать продукты, быть в списках лидеров и разрабатывать маркетинговые стратегии более эффективно. Италия помогает потребителям моды делать лучший выбор. Специалисты по прогнозированию тенденций путешествуют по миру, чтобы изучить различные культурные факторы, влияющие на моду. Они обращают внимание на музыку, искусство, образ жизни и поведение потребителей. Собранная информация помогает им предсказать будущие тренды.

Прогнозирование модных тенденций основано на субъективных выводах специалистов, не несет высокую точность таких прогнозов. В последнее десятилетие из-за быстрого развития моды прогнозировать тенденции стало ещё сложнее. С другой стороны, настал век цифровых технологий. Современные технологии позволили собрать огромное количество данных о моде. Эти данные можно использовать для прогнозирования модных тенденций на основе объективного анализа.

Цель этой статьи — изучить большие объёмы исторических данных о моде и предсказать её возможное развитие в будущем.

Использование нейронных сетей для детектирования трендов модной одежды представляется обоснованным выбором по нескольким причинам. Нейронные сети способны обрабатывать большие объёмы данных, включая изображения, видео и текстовые данные. Это позволяет им выявлять закономерности и тенденции в модных трендах. Нейросети могут автоматически анализировать данные о модных тенденциях, что упрощает и ускоряет процесс анализа. Это особенно полезно для компаний, занимающихся модой, которые хотят быстро реагировать на изменения в индустрии. Так же нейронные сети могут предоставлять более точные и объективные результаты, чем человеческий анализ. Они не подвержены влиянию личных предпочтений или предубеждений.

Однако стоит отметить, что использование нейронных сетей требует значительных ресурсов и времени на обучение и настройку. Также необходимо учитывать этические аспекты использования таких технологий в сфере моды.

Анализ существующих решений. Существует несколько известных решений, которые используют глубокое обучение нейросети для прогнозирования и детектирования трендов в модной одежде. В данном разделе мы рассмотрим наиболее известные из них, а также проанализируем их преимущества и недостатки [1].

1. DeepFashion

Решение для прогнозирования модных трендов, которое использует сверхточную нейронную сеть для анализа изображений из Интернета, собранное Лабораторией Мультимедиа Китайского университета. Оно было обучено на большом количестве изображений одежды и аксессуаров и может предсказывать модные тренды на основе этих данных. DeepFashion содержит 800 тыс. изображений, которые помечены 50 категориями, 1000 атрибутами, ориентирами одежды и более 300 тыс. пар изображений. Это самый большой и полный набор данных по одежде на сегодняшний день [2].

Плюсы:

- Автоматизация процесса анализа и генерации модных образов.
- Возможность создания новых, уникальных образов на основе существующих.
- Улучшение качества и разнообразия ассортимента для ритейлеров.

Минусы:

- Необходимость большого количества данных для обучения модели.
- Риск неточных результатов в работе модели.
- Зависимость от качества и актуальности исходных данных.

2. Fashion-Gen

Решение представляет собой генеративно-сопоставительную сеть (GAN), которая может генерировать новые модные образы на основе существующих данных. Сеть обучается на большом наборе фотографий одежды и аксессуаров, а затем использует полученные знания для создания новых образов [3]. Каждая модная вещь сопровождается описательными подписями длиной в абзац, полученными от экспертов (профессиональных дизайнеров). Распределение длины описаний представлено на рисунке 1.

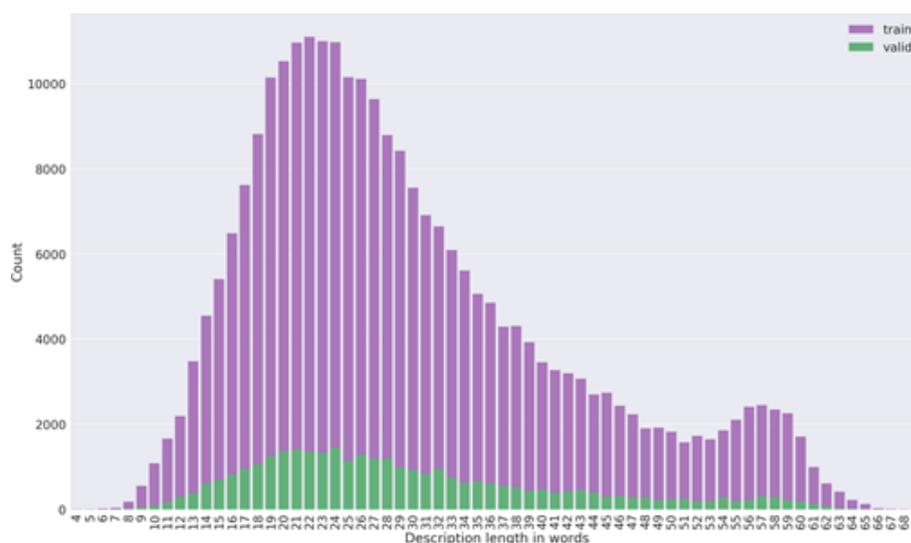


Рис. 1. График распределения длины описаний

Плюсы:

- Возможность создавать уникальные и интересные образы.
- Помощь в создании виртуальных примерочных.
- Экономия времени и денег на создание коллекций.

Минусы:

- Необходимость обучения модели на больших объемах данных.
- Риск генерации некачественных или нереалистичных изображений.

Как эффективно моделировать соответствующие данные для получения точных прогнозов? Прежде всего, исходные данные должны быть подробными и масштабными, чтобы охватывать достаточно длительный период времени и отражать изменения в моде с течением времени. По сравнению с началом модных показов, социальные сети являются более актуальным источником данных, поскольку они оперативно отслеживают тенденции моды, ежедневно публикуя фотографии и комментарии о моде из множества источников: от пользователей, модных блогеров и брендов. С помощью компьютерного зрения или других методов машинного обучения можно получить много информации о пользователях и элементах моды из изображений, метаданных и других источников. Однако наборы данных, основанные на социальных сетях, содержат ограниченное количество информации о моде и не подходят для прогнозирования значимых тенденций в индустрии моды.

В этой статье мы используем набор данных, который содержит множество подробных элементов моды, включая категорию, атрибут и стиль. Он также охватывает более длительный период времени и содержит больше информации о пользователях. Для точного прогнозирования модных тенденций необходимо эффективно фиксировать данные, на основе которых можно выявить закономерности во временных рядах. Традиционные модели, такие как статистические или матричная факторизация, хорошо работают с простыми данными временных рядов, но не позволяют делать обоснованные прогнозы для сложных тенденций моды.

Недавние достижения в области глубокого обучения позволили найти эффективные решения для многих задач. Однако этот подход ещё не применялся для анализа модных тенденций. Большинство существующих работ прогнозируют тренд для каждого элемента моды отдельно. Однако элементы моды не являются независимыми, а тесно связаны друг с другом. Например, свитер похож на модель с горловиной «черепаша», но отличается от платья. Чтобы сделать более точный прогноз тренда свитера, можно использовать результаты прогноза для «черепашьей» горловины и платья. Это позволит уточнить прогноз для свитера на основе наблюдаемых корреляций между этими элементами [4].

В мире моды существуют таксономические отношения между элементами, например, между одеждой и всеми атрибутами принадлежности (рисунок 2). Эти таксономические соотношения необходимо учитывать при моделировании модных тенденций. В этой статье используется подход под названием Knowledge Enhanced Recurrentnetwork (KERN) для прогнозирования модных тенденций среди разных групп людей. Этот подход позволяет эффективно моделировать элементы данных временных рядов, используя сложные шаблоны с помощью фреймворка кодирования-декодирования долгосрочной и краткосрочной памяти (LSTM).

Более того, KERN объединяет два типа знаний: внутренние и внешние. Внутренние знания он получает через анализ сходства временных рядов в наборе данных и вводит тройную регуляризацию потерь на основе шаблонов. Внешние знания позволяют связать элементы моды с таксономией, а также обновлять их с помощью обмена сообщениями между корпорациями.

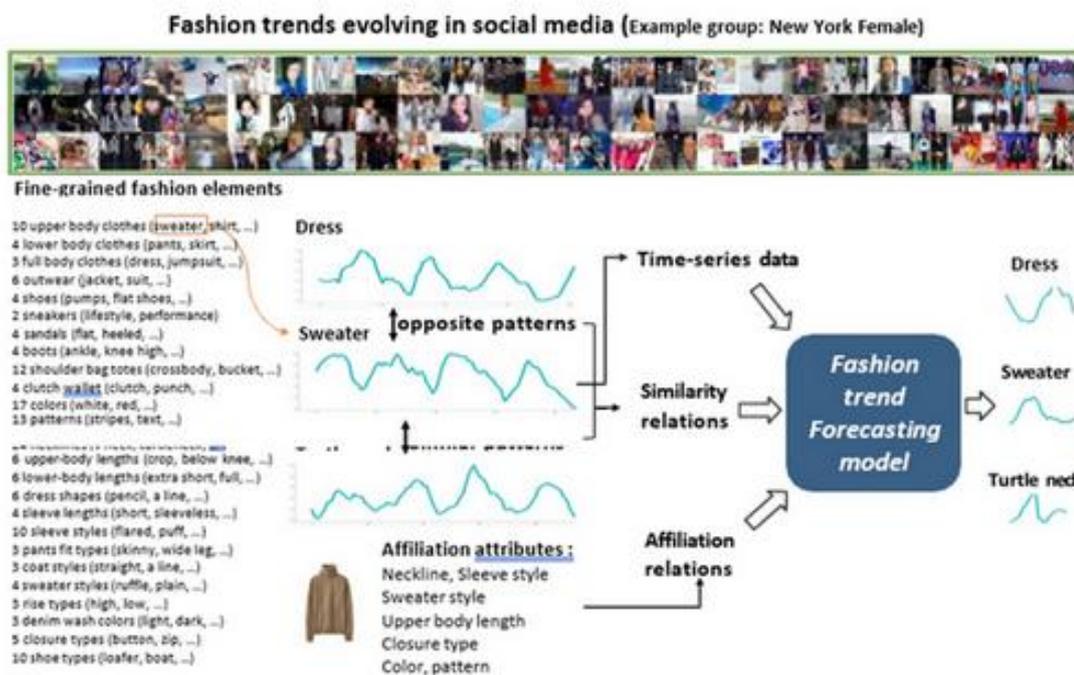


Рис. 2. Задача прогнозирования модных тенденций

Модель KERN включает в себя данные о временных рядах отдельного элемента моды и его связи с другими элементами. Мы используем информацию о пользователях, чтобы лучше моделировать модные тенденции для разных групп пользователей.

Результаты можно разделить на три группы:

1. Для определения актуальных модных тенденций мы используем большой набор данных о тенденциях в моде, основанный на данных из Instagram (Fashion Instagram Trending, или FIT).

2. Для прогнозирования модных тенденций предлагается использовать усовершенствованную модель на основе STM (KERN), которая эффективно моделирует временные ряды данных о модных тенденциях.

3. Мы показываем, что наша модель может собирать данные о модных тенденциях во временных сериях и эффективно прогнозировать их развитие.

Анализ модных тенденций. В последнее время исследователи проявляют интерес к различным задачам в области моды, таким как распознавание, поиск, смешивание, визуализация и т. д. Однако анализ модных тенденций и прогнозирование их развития всегда оставались важными классическими исследованиями в сфере моды. Они играют ключевую роль в управлении всей модной индустрией. Хидаяти и другие учёные разработали фреймворк для автоматического определения модных тенденций, которые часто встречаются на неделях моды. Виттая Корнет и другие исследователи расширили эту задачу, увеличив набор данных и изучив современную мировую моду. Их цель — провести количественный анализ моды и тенденций, а не делать прогнозы.

Matzen исследовали тенденции в одежде с помощью статистического анализа, используя масштабный набор данных об изображениях из социальных сетей.

В свою очередь, Mall сосредоточились на специфических элементах моды и попытались выявить детализированные модные тенденции. Они моделировали сигналы модных тенденций для каждого элемента с помощью комбинации линейных и циклических компонентов.

Эти компоненты, согласно исследованию, способны улавливать как общие уровни тенденций, так и тонкие изменения. Однако их исследование ограничено тем, что они рассматривали только определённые элементы моды, которые демонстрировали простые шаблоны в своих трендовых сигналах (например, наличие или отсутствие шляпы), и не включали элементы с более сложной структурой. Таким образом, существующие исследования по-прежнему ограничены статистическим анализом или прогнозированием тенденций развития определённых элементов моды на основе простых шаблонов [5].

Подход. Цель данной статьи — разработать модель для прогнозирования модных тенденций на основе исторических данных.

Мы предлагаем использовать базовую платформу LSTM encoder decoder framework, которая объединяет временные ряды и связанную с ними информацию о последовательности в единую модель и позволяет создавать многогоризонтный прогноз. Чтобы улучшить модель, мы добавляем расширенный термин «регуляризация», который явно включает внутренние знания [6]. Также мы вводим модуль передачи сообщений, чтобы использовать внешние знания, полученные из таксономии.

Предложенную систему фреймворков мы называем улучшенной рекуррентной сетью (KERN), как показано на рисунке 3



Рис. 3. Два примера из набора данных FIT, где красные кривые взяты из набора данных FIT, а синие кривые взяты из Google Trends (оба примера относятся к группе [Нью-Йорк, женщины])

Эксперимент. Чтобы проверить эффективность предложенного нами подхода, были проведены эксперименты с двумя наборами данных. В частности, нас интересуют главный вопрос для исследования:

- 1) Соответствует ли наша модель современным методам прогнозирования будущего модного тренда?

Экспериментальные настройки. Мы используем два набора данных для прогнозирования модных тенденций: предложенный нами набор данных о соответствии и набор данных о геостиле. Например, мы используем данные за один год (52 точки данных), чтобы предсказать результаты следующего полугодия (26 точек данных). Мы разрабатываем две настройки, которые позволяют:

1) использовать данные за два года (48 точек данных) в качестве входных данных и прогнозировать результат следующего полугодия (12 точек данных);

2) использовать данные за два года (48 точек данных) в качестве входных данных и прогнозировать результат следующего полугодия (24 точки данных).

Поскольку набор данных Geostyle содержит только один атрибут (city) и не содержит пользовательских атрибутов fage и gender, то упрощаем объединение групповых атрибутов и напрямую используем привязку к городу в качестве привязки к группе. Стратегия скользящих окон была применена к обоим наборам данных для создания обучающей и тестовой выборок.

Детали реализации. Мы устанавливаем размер привязки к пользователям (включая привязку к возрасту, полу и городу), размер привязки к элементам одежды и времени до 10, а также размер каждого кода и декодера в нашей сети to50. Гиперпараметр λ для регуляции равен 2×10^{-4} . Во время обучения случайным образом отбираем набор из 400 различных временных рядов для каждой итерации. Для каждой экспериментальной установки настраиваем одну модель ядра для всех модных элементов. Для каждого тестового примера вычисляем эффективность текущих точек данных в соответствии с результатами проверки, а также эффективность последующих точек данных в соответствии с результатами тестирования.

Базовые методы. Выбираем несколько современных методов прогнозирования временных рядов для сравнения с предложенной базовой моделью: среднее и последнее. В качестве прогнозного значения используется среднее значение или значение последней точки в исходных исторических данных.

Авторегрессия (AR) - линейный регрессор, который использует линейную комбинацию последних наблюдаемых значений с предшествующим значением.

Векторная авторегрессия (VAR) - обобщение ARby, допускающее более одной изменяющейся переменной.

Экспоненциальное сглаживание (ES) - объединяет все исторические значения с показательно уменьшенным весом, более поздние значения оказывают большее влияние на прогноз будущего [7].

Линейная и циклическая - линейная или циклическая параметрическая модель, историческое значение которой соответствует конкретной заданной модели.

Геостиль - параметрическая модель, сочетающая линейный компонент и циклический компонент. Это современный метод прогнозирования модных тенденций в наборе геостилированных данных

Мы оцениваем эффективность нашей модели для прогнозирования модных тенденций, сравнивая её с несколькими классическими и современными базовыми моделями. Результаты представлены в таблице 1.

На основе полученных данных мы можем сделать следующие выводы:

1. Модель KERN показывает наилучшие результаты на обоих наборах данных при любых экспериментальных настройках. В частности, на Geostyle только KERN позволяет получить MAE ниже 0,014 и MAPE ниже 15%.

2. На Offitdataset все базовые модели показали MAE и MAPE около 0,11 и 40% соответственно. Наша модель достигает значений менее 0,1, а MAPE — около 30%, что говорит о её высокой производительности.

Набор данных Fitdataset содержит больше детализированных элементов моды, информации для пользователей и более реалистичные и сложные шаблоны временных рядов. Поэтому его сложнее моделировать.

Таблица 1. Показатели KERN и базовые показатели для прогнозирования модных тенденций

Метод	Полгода		Полгода		Один год	
	MAE	MAPE	MAE	MAPE	MAE	MAPE
Mean	0.0292	25.79	0.132	65.31	0.135	63.21
Last	0.0226	21.04	0.125	46.45	0.147	54.04
AR	0.0211	20.69	0.114	54.36	0.119	51.96
VAR	0.0150	17.95	0.157	62.97	0.126	47.35
ES	0.0228	20.59	0.133	55.29	0.150	57.42
Linear	0.0365	24.40	0.112	43.30	0.133	45.89
Cyclic	0.0165	16.64	0.129	49.92	0.143	51.66
GeoStyle	0.0149	16.03	0.136	52.40	0.149	53.14
KERN	0.0134	14.24	0.083	30.02	0.094	33.45

Базовые методы не могут эффективно работать с таким набором данных. Однако наш основной метод справляется со сложными шаблонами лучше других базовых моделей. Это связано с тем, что он обеспечивает лучшую нелинейность в данных, а также позволяет использовать обширные знания о домене. Именно поэтому наша модель показывает улучшение стиля.

Большинство методов дают лучшие результаты при полугодовом прогнозе, чем при годовом, включая нашу основную модель.

Обсуждение эффективности внедрения знаний (RQ2). В этой работе проводим эксперименты и анализируем, насколько эффективно можно внедрить корпоративные знания в модель прогнозирования модных тенденций.

Используем два типа знаний:

1. внутреннее знание отношений подобия, которое применяется путём введения члена триплетной регуляризации в функцию потерь;
2. внешнее знание отношений принадлежности, построенное на основе таксономии, которое используется путём обновления встроенных элементов fashion через передачу сообщений в соответствии с деревом принадлежности.

Чтобы оценить эффективность каждого типа знаний, мы тестируем модель, удаляя каждый тип по очереди, а также оба одновременно. Модель без использования внутренних знаний (KERN-I) игнорирует потерю триплетной регуляризации, а модель без использования внешних знаний (KERN-E) не оснащена модулем включения внешних знаний.

Модель KERN-IE не содержит ни внутренних, ни внешних знаний. Из показателей в таблице 2 видно, что удаление отношений принадлежности или сходства может ухудшить производительность нашей модели. Разница, получаемая за счёт использования дополнительных знаний, более существенна для долгосрочного прогнозирования на основе подходящего набора данных. Это показывает, что использование знаний может быть особенно полезным для получения сложных и требующих проверки данных.

Таблица 2. Вклад различных знаний в модель KERN (результаты MAE)

Датасет	GeoStyle	FIT	
		Полгода	Один год
Предсказание	Полгода	Полгода	Один год
KERN-IE	0.0137	0.0840	0.0966
KERN-E	-	0.0835	0.0953
KERN-I	0.0134	0.0831	0.0942
KERN	0.0134	0.0836	0.0939

На рисунке 4 показаны два примера, в которых сравниваются результаты прогнозирования тенденций с помощью KERN и удалённого KERN-IE. Результаты прогнозирования с помощью KERN явно лучше, чем с помощью KERN-IE, что ещё раз показывает преимущества использования внутренних и внешних знаний в модели KERN [8].

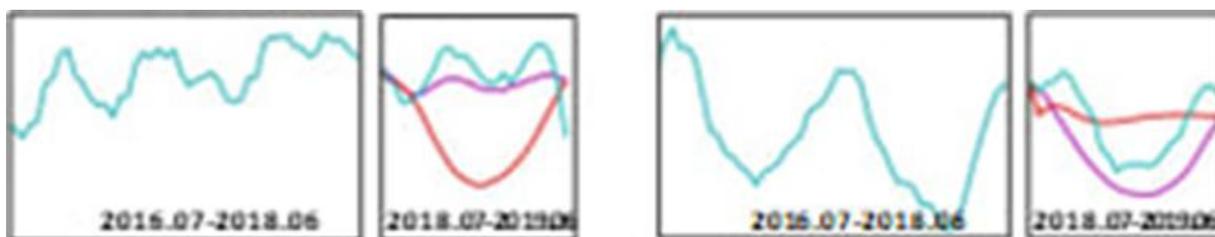
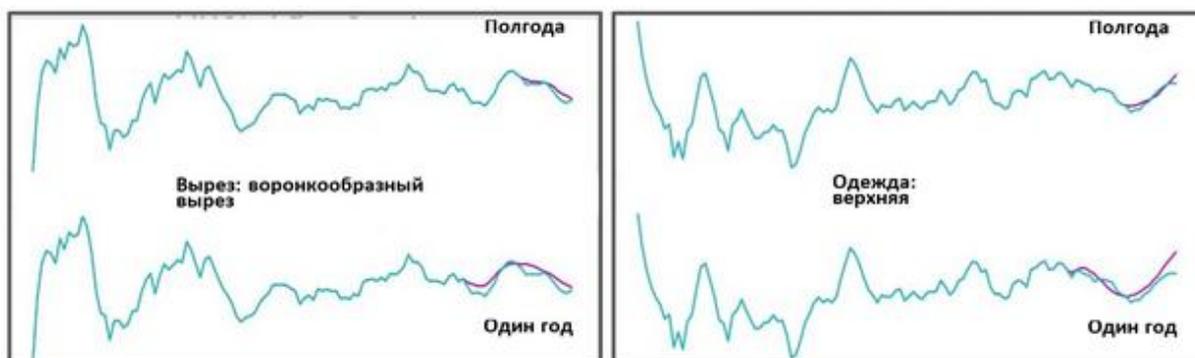


Рис. 4. Два примера прогнозирования тенденций, выполненных с помощью моделей KERN и KERN-IE:

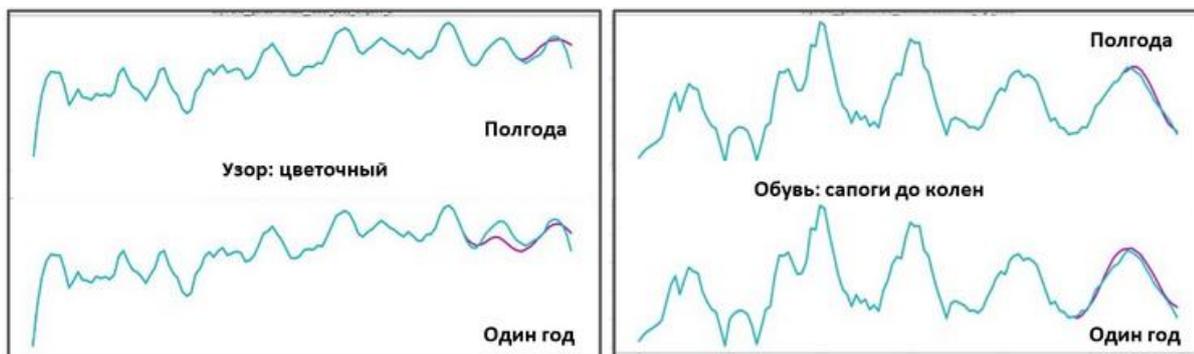
- Основная истина
- Прогнозирование с использованием пользовательской информации
- Прогнозирование без информации о пользователе

Анализ модных тенденций (RQ3). Чтобы ещё раз проиллюстрировать эффективность нашей модели KERN для прогнозирования моды, мы представляем дополнительные результаты визуализации.

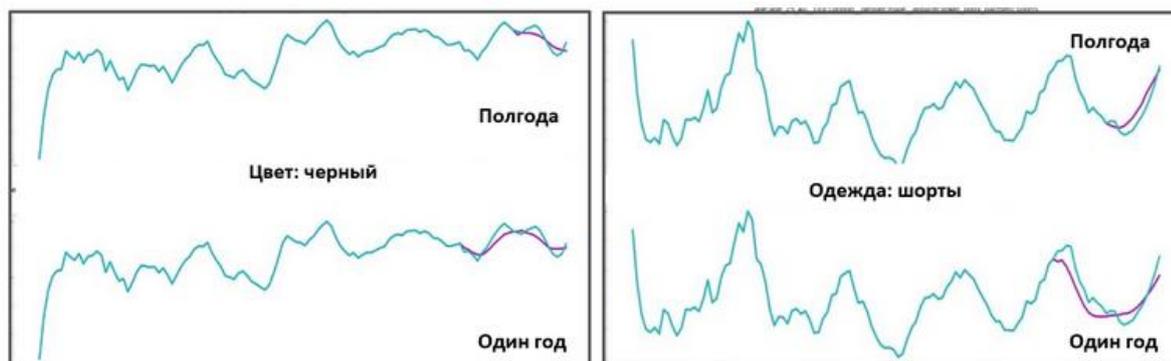
На рисунке 5 показан прогноз на год и полугодие для шести репрезентативных элементов моды (включая категорию одежды, обуви, цвет, рисунок и атрибуты) для трёх групп пользователей. Мы видим, что в целом KERN может очень хорошо прогнозировать тенденции в модных элементах как на год, так и на полугодие, даже для тех, у кого довольно сложные модели.



а) Лос-Анджелес, женщина, 18-25 лет



б) Париж, женщина



в) Лондон, мужчина, 25-40 лет

Рис. 5. Примеры прогнозирования тенденций для различных элементов моды и групп пользователей: Голубые кривые соответствуют действительности, а фиолетовые - результатам прогнозирования КЕРНА

Результаты на рисунке б показывают, что один и тот же элемент моды для разных групп пользователей может отличаться. Это также демонстрирует, что метод KERN способен эффективно предсказывать различные тенденции, используя информацию о пользователях. Результаты прогнозирования с использованием информации о пользователях намного лучше, чем без неё.

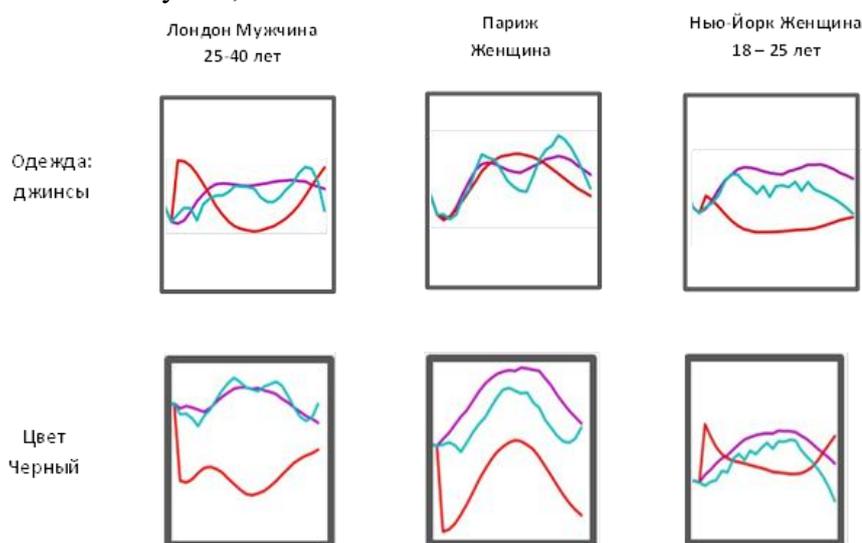


Рис. 6. Результаты прогнозирования одного модного элемента для разных групп пользователей:
 — Основная истина — Прогнозирование с использованием пользовательской информации
 — Прогнозирование без информации о пользователе

На основе прогнозирования различных типов модных элементов мы можем составить подробный и профессиональный отчёт о тенденциях моды. Он охватывает множество модных тенденций, включая категорию, цвет, рисунок, стиль и даже такие подробные характеристики, как цвет.

Заключение. В этой статье рассматривается проблема прогнозирования модных тенденций на основе данных из социальных сетей. Цель исследования — выявить сложные закономерности в исторических записях временных рядов элементов моды и спрогнозировать будущие тенденции.

Для этого была протестирована эффективная модель Knowledge Enhanced Recurrent Network (KERN), которая позволяет фиксировать сложные закономерности во временных рядах данных и прогнозировать модные тенденции.

Хотя были приложены значительные усилия и получены желаемые результаты, есть некоторые аспекты, которые можно улучшить в будущем. Во-первых, необходимо изучить больше информации о пользователях, например, об их профессиях или хобби. Во-вторых, следует учитывать многочисленные источники знаний, такие как анализ моды из профессиональных модных журналов, блогов и брендов.

СПИСОК ИСТОЧНИКОВ

1. Zhu, S. A Be Your Own Prada: Fashion Synthesis with Structural Coherence / S. Zhu. – 2017. – URL: <https://arxiv.org/abs/1710.07346>.
2. Large-scale Fashion (DeepFashion) Database [Электронный ресурс]. – URL: <https://mmlab.ie.cuhk.edu.hk/projects/DeepFashion.html> (дата обращения 10.09.2024).
3. Fashion-Gen: The Generative Fashion Dataset and Challenge [Электронный ресурс]. – URL: <https://deepai.org/publication/fashion-gen-the-generative-fashion-dataset-and-challenge> (дата обращения 10.09.2024).
4. Власенко, А. И. Анализ методологии прогнозирования модных трендов / А. И. Власенко. – СПб: Издательство СПбГУ, 2020. – 120 с.
5. Использование нейросетей в индустрии моды [Электронный ресурс]. – URL: https://www.researchgate.net/publication/376596939_Ispolzovanie_nejrosetej_v_industrii_mody (дата обращения 12.09.2024).
6. Details of the LSTM [Электронный ресурс]. – URL: https://www.researchgate.net/figure/Details-of-the-LSTM-and-Bi-LSTM-models_fig5 (дата обращения 12.09.2024).
7. Построение моделей временных рядов [Электронный ресурс]. – URL: <https://www.iprbookshop.ru/68826.html> (дата обращения 12.09.2024).
8. Предсказание временных рядов с помощью нейросети Keras [Электронный ресурс]. – URL: <https://newtechaudit.ru/predskazanie-vremennyh-ryadov-s-pomoshhyu-keras/> (дата обращения 12.09.2024).

Поступила в редакцию 28.10.2024 г., рекомендована к печати 14.11.2024 г.

DETECTING FASHION TRENDS WITH NEURAL NETWORK TRAINING

Evdacheva E.D., Maksimenko I.I.

Forecasting fashion trends is a crucial task for both academia and industry. Previously, to solve this difficult task, only limited fashion elements with seasonal or simple patterns were studied, which could hardly reveal real fashion trends. Aiming at forecasting, this article is intended to explore clear trends in the fashion world for certain user groups. A large-scale data set on fashion trends (FIT) was used in the work. Also, for more efficient modeling of time series of data, a recurrent network model was taken, which takes advantage of neural networks when modeling time series. Extensive experiments demonstrate that the proposed CORE model can effectively capture complex patterns of objective fashion elements, therefore making a preferable forecast of fashion trends.

Keywords: forecasting fashion trends, fashion analysis, forecasting time series.

Евдачёва Екатерина Дмитриевна

магистрант кафедры компьютерных технологий
ФГБОУ ВО «Донецкий государственный
университет»,
Российская Федерация, ДНР, г. Донецк.
E-mail: alisakovalllll@gmail.com

Evdacheva Ekaterina Dmitrievna

Master Student at Department of Computer
Technologies of Donetsk State University,
Russian Federation, DPR, Donetsk.

Максименко Игорь Иванович

кандидат физико-математических наук, доцент
кафедры компьютерных технологий ФГБОУ ВО
«Донецкий государственный университет»,
Российская Федерация, ДНР, г. Донецк.
E-mail: igor.maksimenko_1967@mail.ru

Maksimenko Igor Ivanovich

Candidate of Physical and Mathematical Sciences,
Associate professor at Department of Computer
Technologies of Donetsk State University,
Russian Federation, DPR, Donetsk.

ФИЛЬТРАЦИЯ СПАМА МЕТОДАМИ ГЛУБОКОГО ОБУЧЕНИЯ

© 2024 Ермоленко Т.В., Шалун Н.А.

Фильтрация спама относится к задачам классификации естественно-языковых текстов, для решения которой эффективно используются модели глубокого обучения. В статье анализируется возможность применения основных базовых архитектур глубоких сетей к фильтрации спама, таких как сверточные и рекуррентные нейронные сети, а также архитектура Трансформер. Рассмотрены также векторные представления текста на уровне слов и с помощью языковых моделей. Современные языковые модели имеют архитектуру Трансформер, обеспечивают высокую точность классификации, но имеют огромное количество параметров и вычислительную мощность. В связи с этим в работе для фильтрации спама в качестве модели классификации использовался многослойный перцептрон, а для получения векторного представления текста – TF-IDF. Такой выбор продиктован необходимостью соблюсти баланс между вычислительной мощностью и точностью. Оценка точности модели осуществлялась кросс-валидацией с разбиением данных методом k-Fold и составила в среднем 99.962%.

Ключевые слова: фильтрация спама, глубокое обучение, word embedding, TF-IDF, языковая модель, модель Трансформер, многослойный перцептрон, метод k-Fold.

Введение. Электронная почта – ключевой канал коммуникаций, однако она подвержена спаму, который несет угрозы безопасности, такие как вредоносные ссылки.

Условно методы фильтрации спама можно разделить на три основные категории:

1. Методы на основе правил: эти методы используют ключевые слова и черные списки для идентификации спама [1-4]. Отличаются простотой реализации и малой ресурсоемкостью.

2. Классические методы машинного обучения – методы на основе статистики: наивный байесовский классификатор и метод опорных векторов (Support Vector Machine, SVM), которые определяют вероятность того, что сообщение – спам.

3. Методы глубокого обучения, использующие глубокие нейронные сети как для представления текста в виде вектора (word embedding), так и для его бинарной классификации (спам/не спам).

Главной проблемой методов, относящихся к первым двум категориям, является то, что они подвержены ошибкам при встрече с новыми шаблонами спама, поэтому требуют ручной настройки [1-4], что приводит к потере эффективности этих методов фильтрации спама из-за улучшения способов обхода фильтров.

Глубокие модели – эффективные классификаторы, обладающие адаптивностью к изменениям в спам-тактике, высокой точностью, благодаря анализу больших объемов данных, отсутствием ручной настройки.

Сложность их применения состоит в огромном числе настраиваемых параметров, что требует как мощных вычислительных ресурсов, так и большого количества данных для их обучения. Кроме того, модели могут терять обобщающую способность из-за переобучения (overfitting) [5].

Для решения этих проблем используются технологии трансферного обучения (transfer learning) и тонкой настройки модели или дообучения (fine tuning).

При transfer learning происходит перенос знаний с одной модели на другую, т.е. обученная модель переиспользуется для другой, связанной, задачи. Суть дообучения заключается в размораживании последних слоев нейронной сети, которые имеют наиболее абстрактные представления, и их обучении. Дообучение только нескольких слоев уменьшает риск переобучения, и, самое главное, позволяет сделать текущую модель ещё более подходящей к решаемой задаче. Техники трансферного обучения и тонкой настройки модели требуют гораздо меньших вычислительных затрат и наборов данных по сравнению с обучением модели с нуля.

Вышеозначенное делает глубокие модели неоспоримым лидером в решении задач фильтрации спама.

Типовая структура системы фильтрации спама. Решение любой задачи классификации естественно-языковых текстов, к которым принадлежит и фильтрация спама, методами глубокого обучения состоит из четырех последовательных этапов:

- предварительная обработка текста;
- получение векторного представления текста;
- построение модели бинарной классификации и ее обучение;
- оценка точности классификации и корректировка параметров модели.

Таким образом, типовая структура системы фильтрации спама, использующая глубокое обучение, включает элементы, соответствующие этим этапам: модуль предобработки текста, в котором текст выделяется из всего сообщения и приводится к единому формату; модуль преобразования обработанного текста в вектор признаков; модуль бинарной классификации почтовых сообщений, в котором в режиме обучения осуществляется настройка модели, а в режиме классификации письма автоматически относятся к той или иной категории; модуль корректировки параметров модели.

Помимо основных модулей, для увеличения скорости работы и эффективности классификации система может включать в себя следующие блоки:

- архив электронной почты, что исключает потерю важной информации при фильтрации, а также позволяет осуществлять первичное обучение модуля классификации почтовых сообщений;
- анализ репутации IP-адреса и домена, который заключается в обращении спам-фильтров к чёрным спискам IP-адресов и доменов, сообщение блокируется или помечается как спам, если отправитель находится в блок-листе;
- проверка протоколов аутентификации, т.е. подлинности отправителя, письмо, которое не прошло аутентификацию, классифицируется как спам.

Современные системы фильтрации спама в качестве модели бинарной классификации используют глубокие нейросети.

Глубокие модели для фильтрации спама. Глубокой нейронной сетью принято называть нейросеть с двумя или более скрытыми слоями [6].

При обучении модели глубокого обучения нам необходимо изменить веса – коэффициенты, которые присваиваются каждому признаку-столбцу и передают важность этого соответствующего признака при классификации. Более того, веса позволяют минимизировать функцию потерь (loss function). Чем меньше ее значение, тем ближе точнее классификация.

Можно выделить следующие категории для основных нейросетевых архитектур:

- 1) прямонаправленные нейронные сети (Feed-forward networks, FFN) [7];
- 2) рекуррентные нейронные сети (Recurrent Neural Networks, RNN) [8];
- 3) сверточные нейронные сети (Convolutional Neural Networks, CNN) [9];
- 4) модель Трансформер.

FFN являются одной из простейших моделей глубокого обучения для представления текста. Эти модели рассматривают текст как мешок слов. Для каждого слова они обучают векторное представление, используя модели векторного представления слов, а затем вычисляется сумма или среднее значение этих векторных представлений как представление текста, которые передают его через один или несколько слоев прямонаправленных нейронных сетей, известных как многослойные перцептроны (Multi-Layer Perceptron, MLP). Затем выполняется классификация на основе представления финального слоя, используя классификатор.

RNN широко используются для захвата долгосрочной зависимости через рекуррентные вычисления. RNN изучает информацию, учитывая расположение всех слов, подходящих для задач классификации текста. Каждое входное слово представляется определенным вектором с использованием векторного представления, которые подаются в ячейки RNN по одному токenu. Выход ячеек RNN имеет ту же размерность, что и входной вектор, и подается в следующий скрытый слой. RNN использует общие параметры для разных частей модели и имеет одинаковые веса для каждого входного токена. Метка входного текста может быть предсказана последним выходом скрытого слоя.

В процессе обратного распространения ошибки в RNN веса регулируются градиентами, вычисленными через непрерывные произведения производных, и если производные крайне малы, то это может вызвать проблему исчезновения градиента из-за непрерывных умножений. Для устранения этой проблемы используется модель с долгой кратковременной памятью (Long Short-Term Memory, LSTM) [10]. Метод классификации LSTM лучше улавливает связь между словами контекста, использует вентиль забывания для фильтрации бесполезной информации, что способствует улучшению общей способности классификатора к захвату информации.

CNN разработаны для задач обработки изображений. При работе с текстом вместо изображений, как правило, в качестве входных данных используется одномерный массив, представляющий текст. Поэтому для задач такого типа используется особая архитектура CNN, где выполняются операции одномерной свертки и их объединения в единый вектор. Одной из самых распространенных задач классификации текста, в которой используются CNN, является классификация предложений, т.е. классификация предложения по заранее определенным категориям, учитывая n -граммы, т.е. последовательность токенов.

В CNN для классификации текста, как правило, используется одномерный (1-D) сверточный слой с m фильтрами, производящий m -мерный вектор для каждой n -граммы токена в документе. Полученная матрица подается на сверточный слой, в котором перемещается скользящее окно по тексту.

Одномерная свертка представляет собой результат перемещения скользящего окна по входному вектору и применение того же сверточного фильтра или ядра к каждому окну в последовательности, т.е. скалярное произведение между объединением векторов вложения в заданном окне и вектором весов u , за которым часто следует нелинейная функция активации. Сверточный фильтр применяется к каждому окну, что приводит к скалярным значениям для каждого i -го окна.

Векторы, полученные из сверточных слоев, объединяются с помощью операции объединения (pooling). Данная операция используется для объединения векторов, полученных из разных окон свертки, в один вектор размерности l . Это делается путем выбора максимального или среднего значения в полученных векторах из сверток. После используется функция активации, а полученный результат передается в полносвязный линейный слой для классификации.

Благодаря особенностям своей архитектуры, CNN [4] способны выявлять ключевые фразы и зависимости между словами. Применяя фильтры к word embedding, CNN может обнаруживать фразы и комбинации слов, несущие значительный смысл, тем самым улавливая локальные контекстуальные связи внутри текста.

CNN работают на нескольких уровнях, каждый из которых предназначен для распознавания всё более сложных паттернов. В классификации текста это означает, что нижние уровни могут идентифицировать базовые элементы, такие как части речи или простые фразы, в то время как более глубокие уровни могут интерпретировать более сложные конструкции, такие как идиоматические выражения или технический жаргон.

С помощью таких операций, как pooling, которое уменьшает размеры входных данных, CNN удаётся разделить текст на наиболее важные части.

Стоимость вычислений для выявления отношений между словами в предложении в CNN и в RNN увеличивается с увеличением длины предложения. Модели, основанные на Трансформерах [11] используют параллельные вычисления, не учитывая последовательную информацию, что делает их оптимальным выбором для задач обработки естественного языка на больших наборах данных.

Модель Трансформер использует отдельные модели энкодера (преобразует слова входного предложения в один или больше векторов в определенном пространстве) и декодера (генерирует из этих векторов последовательность слов). В качестве стандартных архитектур для энкодера и декодера Трансформер использует полносвязные слои. Трансформер уменьшает количество последовательных операций для привязки двух символов из последовательностей ввода-вывода. Это достигается при помощи механизма многозадачности, который позволяет моделировать зависимости независимо от их расстояния во входном или выходном предложении.

В Трансформер энкодер и декодер состоят из стека одинаковых слоёв. Каждый из этих слоёв состоит из двух общих типов подслоёв:

- механизма многослойного внимания (multi-head attention);
- позиционной полносвязной FFN.

Главное отличие декодера от энкодера в Трансформер – использование слоя с механизмом маскирующего многослойного внимания (masked multi-head attention), который позволяет «обращать внимание» на специфичные сегменты из энкодера. Это возможно благодаря тому, что masked multi-head attention маскирует будущие токены посредством блокирования информации токенов, которые находятся справа от вычисляемой позиции.

За счет уменьшения эффективного разрешения из-за усреднения взвешенных вниманием позиций с помощью механизма многослойного внимания модель Трансформер позволяет сократить количество операций, необходимых для связи (особенно дальних) позиций во входной и выходной последовательности.

Дополнительно, в энкодере и декодере модели Трансформер подслои внимания обрабатываются FNN. Она применяется к каждой позиции отдельно. Линейные преобразования одинаковы для каждой позиции, но используют разные параметры от слоя к слою, что аналогично двум операциям свертки с размером ядра 1.

Ко всему вышесказанному, перспективным направлением является использование предобученных языковых моделей, основанных на архитектуре Трансформер. Это направление особенно важно, если размер обучающих данных небольшой. Тогда необходимо дополнительно настроить нейронную сеть в режиме обучения без учителя под определенную задачу, описанную внутри структуры обучающих данных, а после использовать их непосредственно при обучении на размеченных данных.

Для задачи классификации текстов процесс извлечения признаков включает преобразование необработанных текстовых документов в набор значимых и информативных признаков, которые могут быть использованы в качестве входных данных для моделей машинного обучения.

Векторные представления текста. Процесс извлечения признаков для классификации текстов играет решающую роль в подготовке входных данных для моделей машинного обучения.

На сегодняшний день используется извлечение признаков текста на уровне слов и с помощью языковых моделей.

Известны следующие подходы для извлечения признаков текста на уровне слов:

1) Bag-of-words (BoW, мешок слов) – это подход, который представляет документ в виде вектора частот слов, не обращая внимания на порядок слов. Это простое представление может улавливать наиболее часто встречающиеся термины в документе, что может быть характерно для текстов определенных классов.

2) TF-IDF (TF – term frequency, IDF – inverse document frequency) – это расширение BoW, которое назначает более высокие веса терминам, более специфичным для конкретного документа, тем самым уменьшая влияние общих слов. Это помогает отличать разные классы текстов, так как специфичные для класса термины вероятнее всего будут иметь более высокий вес.

3) n-граммы слов представляют собой последовательности из n последовательных слов, захватывая локальные шаблоны и фразы, которые могут быть характерны для определенного класса.

Главный недостаток BoW заключается в том, что для больших словарей размер векторов будет велик, т.к. размерность вектора равна размеру словаря.

Методы извлечения, признаков, основанные на языковом моделировании используют скрытые состояния для создания векторного представления предложений.

Языковая модель (ЯМ) моделирует вероятностное распределение последовательности слов $P(w_1, w_2, \dots, w_N)$, т. е. позволяет сказать с какой вероятностью может встретиться последовательность слов w_1, w_2, \dots, w_N . Также языковым моделированием называют задачу предсказания следующего слова в последовательности $P(w_N | w_1, w_2, \dots, w_{N-1})$. Поэтому в литературе ЯМ называют модели описывающие как распределение (w_1, w_2, \dots, w_N) , так и распределение $P(w_N | w_1, w_2, \dots, w_{N-1})$.

В настоящее время для построения ЯМ используются глубокие нейросети вышеописанных архитектур. Вектора, описывающие текст, подаются вход нейросети, вычисляющей распределение вероятностей целевых классов, которое дополняет вектор признаков текста. Такое векторное представление (word embeddings) сохраняет контекстное сходство слов и не только поддается обработке с помощью алгоритмов обучения, но и имеет два важных и полезных свойства:

- снижение размерности – это более эффективное представление;
- контекстное подобие – это более выразительное представление.

Именно такой подход к представлению слов и документов можно считать одним из ключевых достижений глубокого обучения в решении сложных проблем обработки естественного языка.

В языковых моделях word embeddings применяется только в самом нижнем энкодере, который получает набор векторов определенной размерности. Размер этого набора векторов является гиперпараметром, который можно устанавливать, и, по сути, равен длине самого длинного предложения в обучающем корпусе. Типичная размерность такого вектора составляет 200-300.

Векторы слов используются для семантического анализа, чтобы извлечь значение из текста, чтобы обеспечить понимание естественного языка. Чтобы языковая модель могла предсказать значение текста, она должна учитывать контекстное сходство слов. Векторы, созданные с помощью word embeddings, сохраняют семантическое сходство, поэтому слова, которые регулярно встречаются рядом в тексте, также будут находиться в непосредственной близости в векторном пространстве.

Выбор техники извлечения признаков зависит от характера текстовых данных и конкретных требований задачи классификации текстов. Например, метод BoW или n-грамм может быть эффективен для захвата частоты и распределения слов в тексте, в то время как векторные представления слов с помощью языковых моделей могут захватывать более тонкие отношения между словами, такие как семантическая близость или аналогия.

Модель Трансформер позволяет добиться высокой точности в задачах классификации текстов. Данный тип поддерживает параллельные вычисления и эффективно извлекает признаки между различными словами через механизм self-attention, решая проблемы краткосрочной памяти, а также существует возможность использования предобученных языковых моделей на основе Трансформер. Однако, стоит учесть их огромное количество параметров и вычислительную мощность. Поэтому стоит подбирать архитектуру, в которой будет соблюден баланс между вычислительной мощностью и точностью. В задаче фильтрации спама применение архитектур на основе Трансформер для извлечения векторных представлений входного текста не является лучшим решением в силу незначительного улучшения метрик оценки качества по сравнению с более простыми моделями на фоне использования намного большего числа вычислительных ресурсов. В данной работе в качестве модели классификации использовался MLP, для получения векторного представления текста – TF-IDF.

Для качественного обучения глубоких моделей требуется большой датасет, содержащий сбалансированные классы.

Создание набора данных для обучения и тестирования модели фильтрации спама. В рамках данного исследования использовались данные из открытых источников, включающих как спам-сообщения, так и легитимные письма на английском и русском языках:

- Enron Email Dataset. Содержит более 500 000 электронных писем, сгенерированных 158 сотрудниками компании Enron, используемых для классификации писем. Из которых 60% являются спамом;

- SpamAssassin Public Corpus. Включает в себя около 6000 писем, 2000 из которых являются спамом;

- Ling-Spam Dataset. Содержит сообщения из почтовых рассылок лингвистов, корпус состоит из 2893 сообщений, из которых 2412 легальных и 481 спам-сообщения, полученное одним из авторов. Каждое сообщение этого датасета содержит более 500 слов;

- TREC 2007 Spam Track. Этот корпус содержит 50 199 спам-писем и 25 220 не спам-писем;

- SMS Spam Collection Dataset. Включает 5500 текстовых сообщений, из которых спам составляет 13%.

В результате интеграции этих текстовый корпусов в единый датасет для обучения и тестирования модели фильтрации спама получен размеченный набор данных, содержащий 50000 легальных сообщений и 40000 спам-сообщений.

Датасет представляет собой CSV-файл объемом 5.4 Мб и состоит из текстовых сообщений и их меток (0 – спам; 1 – легитимное сообщение). Полученный датасет разбит на обучающие/тестовые данные в отношении 80/20.

Разработка и тестирование модели фильтрации спама на основе многослойного персептрона. Для реализации модели фильтрации спама выбрано векторное представление текста TF-IDF, язык программирования Python 3.8, а также использованы следующие библиотеки:

- TensorFlow и Keras – высокоуровневые библиотеки для работы с нейронными сетями;
- Sklearn – для выполнения задач машинного обучения, в частности, для вычисления весов классов и кросс-валидации;
- NLTK – для вычисления TF-IDF и предварительной обработки текста, включающей в себя:
 - приведение текста к нижнему регистру;
 - удаление стоп-слов, не несущих смысловой информации (частицы, союзы, предлоги, местоимения, междометия, вводные слова и т.п.);
 - стемминг – процесс приведения слов к их основной форме путем удаления аффиксов;
- Pandas и Numpy – для загрузки данных и дальнейшей работы с ними, включая числовые массивы.

Для компенсации дисбаланса классов в процессе обучения модели в рамках данной работы используется метод взвешивания классов (Class Weighting), реализованный в Sklearn. Метод присваивает разные веса каждому классу в зависимости от его доли в данных, что позволяет модели лучше справляться с классом-меньшинством.

Для оптимизации процесса обучения и борьбы с overfitting используются Batch Normalization и dropout.

Batch Normalization нормализует выход каждого слоя. Нормализация заключается в преобразовании входных данных каждого слоя таким образом, чтобы среднее значение было приближено к нулю, а стандартное отклонение – к единице. Это делает сеть менее чувствительной к масштабу входных данных и улучшает общую стабильность процесса обучения.

Dropout во время каждого прохода обучения (эпохи) с вероятностью p исключает каждый нейрон. Это помогает предотвратить чрезмерную зависимость модели от конкретных путей и узлов в сети, что может привести к переобучению. Вероятность p является гиперпараметром модели. Для MLP обычно используются значения Dropout от 0.2 до 0.5. Это помогает предотвратить переобучение, поскольку полносвязные слои склонны к запоминанию шума в данных.

Для повышения производительности модели применяется алгоритм оптимизации Adam (adaptive moment estimation). Оптимизатор вносит изменения в веса и скорость обучения сети. Adam, в отличие от поддержания единой скорости обучения стохастического градиентного спуска для обновления весов, обновляет скорость обучения для каждого веса сети индивидуально. Этот оптимизатор выбран, поскольку он считается эталоном в глубоком обучении и рекомендуется в качестве алгоритма оптимизации по умолчанию. Кроме того, алгоритм прост в реализации, быстро отрабатывает, имеет низкие требования к памяти и требует меньше настроек, чем любой другой алгоритм оптимизации.

Предлагаемая архитектура MLP для фильтрации сама имеет следующие гиперпараметры:

- размер batch (обозначает количество сэмплов, которые необходимо взять для обновления параметров модели) – 64;
- число нейронов входного слоя – 100000;
- функция активации скрытых слоев – ReLU;
- функция активации выходного слоя – sigmoid [2];
- число нейронов первого скрытого слоя – 512;
- число нейронов второго скрытого слоя – 256;
- dropout – 0.5;
- оптимизатор – Adam;
- функция потерь – бинарная кросс-энтропия.

Для более надежной оценки модели используется кросс-валидация с разбиением данных методом k-Fold, реализованным в Sklearn. Метод k-Fold представляет из себя следующий алгоритм:

1. Фиксируется некоторое целое число k (обычно от 5 до 10), меньшее числа сэмплов в датасете. Это число является параметром алгоритма.
2. Датасет разбивается на k одинаковых частей (в последней части может быть меньше сэмплов, чем в остальных), называемых фолдами.
3. Осуществляется k итераций, во время каждой из которых один фолд выступает в роли тестового множества, а объединение остальных – в роли тренировочного. Модель учится на k–1 фолде и тестируется на оставшемся.
4. Финальная оценка модели получается либо усреднением k получившихся тестовых результатов, либо измеряется на отложенном тестовом множестве, не участвовавшем в кросс-валидации.

В рамках данной работы значение параметра k составило 5.

Точность является основной метрикой, используемой в данном исследовании для оценки модели. Точность определяется как доля правильно классифицированных примеров (как спам, так и не спам) от общего числа примеров.

В ходе обучения и оценки модели с использованием 5-фолдовой кросс-валидации получены следующие результаты:

Для каждого фолда вычислялась точность, которая затем сохранялась в список accuracies. Показатели точности приведены в таблице 1. Средняя точность модели по всем фолдам составляет 99.962%.

Таблица 1. Точность модели по каждому фолду

№ фолда	1 фолд	2 фолд	3 фолд	4 фолд	5 фолд
Точность, %	99.97	99.95	99.94	99.99	99.96

Заключение. Проведенный анализ базовых моделей глубокого обучения показал, что для больших наборов данных модели, основанные на Трансформерах, использующие параллельные вычисления и механизм многозадачности, являются лучшим выбором по сравнению с CNN и RNN, стоимость вычислений в которых существенно увеличивается с увеличением длины предложения. Для извлечения признаков из текста используются подходы на уровне слов и с помощью языковых моделей, которые моделируют вероятностное распределение последовательности слов в тексте. Векторные представления на уровне слов, хоть и легко вычисляемы, имеют

существенные ограничения. Они не отражают семантическую близость слов, а их главным недостатком становится огромная размерность вектора при увеличении словаря. Эффективным представлением текста, имеющим низкую размерность и сохраняющим контекстное сходство слов, является языковое моделирование.

Современные языковые модели основаны на архитектуре Трансформер, обладающей высокой точностью классификации текстов, но имеющей огромное количество параметров и вычислительную мощность. В задаче фильтрации спама применение таких архитектур требует больших вычислительных ресурсов, при этом незначительно улучшая точность классификации по сравнению с более простыми моделями. Для соблюдения баланса между вычислительной мощностью и точностью в данном исследовании в качестве модели классификации использовался двуслойный MLP, для получения векторного представления текста – TF-IDF.

Для обучения и тестирования модели сформирован размеченный датасет, содержащий 50000 легальных сообщений и 40000 спам-сообщений, взятых из открытых источников. Полученный набор данных разбит на обучающие/тестовые данные в отношении 80/20.

В процессе обучения модели для компенсации дисбаланса классов применен метод взвешивания классов, а также техники регуляризации, включая BatchNormalization и Dropout, что способствовало повышению устойчивости модели к переобучению и улучшению общей производительности.

Для более надежной оценки модели используется кросс-валидация с разбиением данных методом k-Fold с параметром $k=5$. Как показали численные исследования, средняя точность модели по всем фолдам составляет 99.962%, что говорит о высоком качестве предложенной модели.

СПИСОК ИСТОЧНИКОВ

1. Складенко, Н. С. Разработка спам-фильтра с использованием методов машинного обучения / Н. С. Складенко // Вестник компьютерных наук. – 2017. – № 4. – С. 45-58.
2. Иванов, П. Н. Фильтрация спама с использованием нейронных сетей / П. Н. Иванов // Программные системы: теория и приложения. – 2020. – Т. 11, № 3. – С. 32-41.
3. Петров, А. В. Методы глубокого обучения для распознавания спам-сообщений / А. В. Петров // Журнал искусственного интеллекта. – 2021. – Т. 12, № 2. – С. 12-24.
4. Григорьев, Д. С. Применение сверточных нейронных сетей для классификации спам-сообщений / Д. С. Григорьев // Труды научных конференций. – 2021. – № 5. – С. 77-85.
5. Тимофеев, В. С. Применение сверточных нейронных сетей для борьбы с рассылкой спама / В. С. Тимофеев // Информационные технологии в безопасности. – 2019. – Т. 14, № 2. – С. 23-36.
6. Гудфеллоу, Я. Глубокое обучение: пер. с англ. / Я. Гудфеллоу, И. Бенджио, А. Курвилль. – М.: ДМК Пресс, 2018. – 652 с.
7. Le, Q. Distributed representations of sentences and documents / Q. Le, T. Mikolov // 31st International Conference on Machine Learning, ICML 2014. – 2014. – V. 4. – P. 1188-1196.
8. Liu, P. Recurrent neural network for text classification with multi-task learning / P. Liu. – 2016. – URL: <https://arxiv.org/abs/1605.05101>.
9. Kim, Y. Convolutional neural networks for sentence classification / Y. Kim // EMNLP 2014 - 2014 Conference on Empirical Methods in Natural Language Processing, Proceedings of the Conference. – 2014. – P. 1746-1751.
10. Hochreiter, S. Long Short-Term Memory / S. Hochreiter, J. Schmidhuber // Neural Computation. – 2019. – V. 9. – № 8. – P. 1735-1780.
11. Huang, Y. A Comprehensive Review of Transformer Models in Natural Language Processing / Y. Huang, Y. Sui // ACM Computing Surveys. – 2022. – V. 54, № 12. – P. 1-38. – DOI 10.1145/3452805.

Поступила в редакцию 30.10.2024 г., рекомендована к печати 18.11.2024 г.

SPAM FILTERING USING DEEP LEARNING METHODS

Ermolenko T.V., SHalun N.A.

Spam filtering is a natural language text classification problem that can be effectively solved using deep learning models. The article analyzes the possibility of applying the main basic architectures of deep networks to spam filtering, such as convolutional and recurrent neural networks, as well as the Transformer architecture. Vector representations of text at the word level and using language models are also considered. Modern language models have the Transformer architecture, provide high classification accuracy, but have a huge number of parameters and computational power. In this regard, in this work, a multilayer perceptron was used as a classification model for spam filtering, and TF-IDF was used to obtain a vector representation of the text. This choice is dictated by the need to maintain a balance between computing power and accuracy. The accuracy of the model was assessed by cross-validation with data partitioning using the k-Fold method and amounted to an average of 99.962%.

Keywords: spam filtering, deep learning, word embedding, TF-IDF, language model, Transformer model, multilayer perceptron, k-Fold method.

Ермоленко Татьяна Владимировна

кандидат технических наук, доцент, доцент
кафедры компьютерных технологий
ФГБОУ ВО «Донецкий государственный
университет»,
Российская Федерация, ДНР, г. Донецк.
E-mail: t.ermolenko@donnu.ru

Ermolenko Tatiana Vladimirovna

Candidate of Technical Sciences, Associate Professor,
Associate Professor at Department of Computer
Technologies of Donetsk State University,
Russian Federation, DPR, Donetsk.

Шалун Никита Андреевич

магистрант кафедры компьютерных технологий
ФГБОУ ВО «Донецкий государственный
университет»,
Российская Федерация, ДНР, г. Донецк.
E-mail: lavandos1919@mail.ru

SHalun Nikita Andreevich

Master Student at Department of Computer
Technologies of Donetsk State University,
Russian Federation, DPR, Donetsk.

ОБОСНОВАНИЕ КОМПЛЕКСНОГО ПОКАЗАТЕЛЯ КАЧЕСТВА РАБОТЫ ТРАНСПОРТНОЙ СЕТИ МОБИЛЬНОГО ОПЕРАТОРА СВЯЗИ

© 2024 Чепцов М.Н., Лозинская В.Н., Климов В.В.

Настоящая работа направлена на обоснование использования функций «полезности» для получения зависимостей оценок качества работы транспортной сети мобильного оператора связи от количества подключаемых пользователей. Функции «полезности» позволяют перейти от разнородных показателей параметров качества обслуживания и оценок эффективности работы сети к безразмерным коэффициентам. Предложенный комплексный показатель позволяет получить зависимость эффективности работы от количества подключаемых пользователей. Комплексный показатель, в совокупности с соответствующими методами прогнозирования, позволяет эффективно как проектировать, так и оптимизировать работу транспортной сети мобильного оператора связи.

Ключевые слова: функции «полезности», транспортная сеть мобильного оператора связи, параметры качества обслуживания, оценки эффективности работы сети, комплексный показатель, методы прогнозирования.

Введение. Мобильный оператор связи (МОС) предоставляет инфокоммуникационные услуги, учитывая требования к параметрам качества обслуживания (QoS, Quality of Service) различных видов трафика (голосовой трафик, трафик мультимедийных потоков, web-трафик). Следовательно, существует проблема согласования этих требований с имеющимися ресурсами сети. Мерами использования ресурсов для МОС являются оценки качества ее работы. Данные оценки важны для перераспределения имеющейся нагрузки на магистральные каналы без потери качества предоставляемых услуг пользователям. Также для формирования оценок качества эффективности работы транспортной сети МОС необходимо учитывать определенное количество параметров QoS с учетом скорости их обновления в зависимости от скорости подключения к базовой станции. Это, в свою очередь, требует больших, как вычислительных, так и ресурсных затрат (постоянный опрос, передача и обработка служебной информации). Для уменьшения вышеприведенных негативных факторов можно воспользоваться комплексным показателем, учитывающим различные методы прогнозирования для получения зависимости оценок качества работы транспортной сети от количества подключаемых пользователей. Следовательно, целью настоящей работы является обоснование требований к комплексному показателю качества работы сети, учитывающего различные оценки на основе функций «полезности», что позволяет осуществить выбор эффективных параметров.

Постановка задачи. В общем виде задача прогнозирования может быть разделена на несколько подзадач: оценка объема выделяемой пропускной способности [1]; оценка задержек, вносимых планировщиком контроллера базовой станции [2]; оценка качества работы сети на основе многокритериальной теории полезности.

При подключении к базовой станции, терминальное оборудование производит анализ уровня получаемого сигнала. Так, транспортная сеть МОС получает дополнительную нагрузку без учета оценок качества эффективности работы. Чтобы учесть вышеперечисленное, представлено обоснование требований к комплексному показателю качества работы сети, учитывающего различные оценки на основе функций «полезности», что позволяет осуществить выбор эффективных параметров.

Обоснование данного критерия позволит облегчить задачу прогнозирования, которая включает в себя основные процедуры, отражающие следующую функциональную зависимость [3]:

- 1) сбор, обработка и прогнозирование трафика участка транспортной сети базовой станции;
- 2) прогнозирование ресурсов, которые необходимо выделить для качественного обслуживания;
- 3) разработка основных оценок качества работы транспортной сети оператора мобильной связи;
- 4) прогнозирование оценок качества работы участка транспортной сети оператора, в зависимости от количества подключаемых пользователей.

Рассмотрим процедуру прогнозирования ресурсов, необходимых для качественного обслуживания поступающего трафика. Переход к функциям полезности возможен благодаря следующим предположениям. Например, пользователь не почувствует разницу при использовании IP-телефонии, если задержка будет 1 мс или 50 мс, но если задержка будет превышать порог в 150 мс или 400 мс, то это будет уже чувствительно. Или при предоставлении видео услуг доступная скорость передачи должна быть выше скорости формирования потока видеокodeком. Выделение излишней пропускной способности не будет влиять на восприятие качества услуги. Следовательно, для каждого из параметров QoS можно выделить зону невосприимчивости либо слабой чувствительности пользователя к смене качества обслуживания. Данные функции – это зависимость качественной нормированной оценки от определенного параметра канала связи. Другими словами, значения функции характеризуют приемлемость канала связи для передачи этого трафика в данный момент. Обычно, в соответствии с типами потоков назначают классы обслуживания.

Для обеспечения лучшей модели свойств мультимедийных потоков введем функции полезности. Функции полезности можно определить, как для пользователя, так и для оператора, с точки зрения использования ресурса и его стоимости делятся на два типа: функции полезности, отображающие влияние конкретных параметров QoS на восприятие пользователя (представляют полезность между пороговыми значениями, а также изменение полезности в точках перехода) и функции полезности приложения. Последние вводятся для оценки взаимодействия с пользователем на основе анализа результатов функций полезности нарушения работы подсети.

Для того, чтобы разработать функции полезности предоставления ресурсов оператором связи, например, для магистрального канала, зададимся основными параметрами контроля при заданных характеристиках качества работы. Например, насколько для оператора связи будет критичным использование канала, загруженного на 90% или при каком уровне загрузки магистрального канала, необходимо осуществлять подключение нового? Ответ, на самом деле не так очевиден. Это связано, в первую очередь, с параметрами качества обслуживания, которые необходимо поддерживать при передаче по магистральному каналу, их разнородности и зависимости друг от друга. Так, обеспечение определенного уровня задержки, вызывает необходимость в выделении определенного ресурса для обслуживания очередей в буферах устройств. Обеспечение необходимого уровня потерь достигается за счет увеличения пропускной способности выделяемого ресурса по установлению виртуального канала. Чем больше выделяется ресурса под установление виртуального канала, тем меньше можно таких каналов организовать в физическом, а, с учетом задержки, скорость в таком канале уменьшается. Все перечисленные факторы приводят к неэффективному использованию ресурсов и потерей заработков для оператора связи.

Кроме того, оператору связи необходимо оценить каждую единицу предоставляемого ресурса. При этом, явным образом можно дать стоимостную оценку только определенным видам ресурсов. Для того, чтобы определить стоимостную оценку остальных ресурсов можно ввести функции полезности ресурсов оператора связи.

Традиционные параметры QoS могут быть легко определены с помощью числовых функций полезности, тогда как другие, такие как видеокодек и качество изображения, часто подходящим образом выражаются нечисловым и неоднородным образом. Рассмотрим механизм отображения между пространством качества и пространством числовой оценки полезности для измерений QoS, которые не подходят для моделирования исключительно математической функцией.

Рассмотрим такие виды трафика, как VoIP, видеотрафик и трафик, формируемый онлайн-играми.

VoIP – популярная услуга в IP-сети. К приложениям этой услуги относятся, помимо передачи речи, сетевое радио и голосовые сообщения. Передача голосовых услуг наиболее уязвима к задержкам, джиттеру и потерям, чем другие неинтерактивные голосовые услуги. Для пользователей с высокими требованиями к интерактивности, представим следующую общую функцию полезности для аудиопотоков. Функция имеет три области, отражающие средний уровень терпимости человека к искажениям:

$$U_{VoIP}(d) = \begin{cases} 100 - \gamma_1 d, & d < c_1 \\ b_1 \tanh(\beta(\chi - b_2)) + b_3, & c_1 \leq d \leq c_2, \\ \delta - \gamma_2 d, & \chi > c_2 \end{cases} \quad (1)$$

где d – оценка задержки;

γ_1, b_1, γ_2 – коэффициенты отражающие свойства ухудшения качества в каждой области

c_1, c_2 – нижний порог (граница первой области и второй области) и верхний порог (граница второй и третьей области) приложения соответственно;

b_1, b_2, b_3 – параметры настройки для обеспечения непрерывности между областями.

Подставляя эмпирические значения, перепишем (1) в следующий вид:

$$U_{VoIP}(d) = \begin{cases} 100 - 0,067 \cdot d, & d < 150 \\ -25 \cdot \tanh(0,04 \cdot (d - 205)) + 65, & 150 \leq d \leq 300. \\ 45,5 - 0,016 \cdot d, & d > 300 \end{cases}$$

Подставляя численные значения в приведенное выражение, получим графическую зависимость функции утилизации от величины задержки (рисунок 1).

Во время голосового разговора потеря одного пакета может сделать слово неузнаваемым, а отдельное слово может быть ключевым словом для всего предложения.

Согласно этому можно выразить общую функцию полезности для потери пакетов как:

$$U_{VoIP}(Pl) = b_1 + b_2 \cdot \ln(b_3 + \beta \cdot Pl),$$

где Pl – текущий уровень потерь;

β – крутизна ухудшения качества при увеличении потери пакетов.

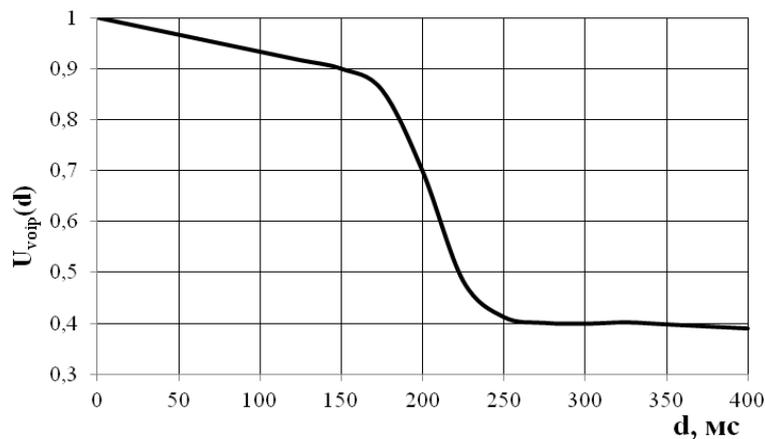


Рис. 1. График функции полезности $U_{VoIP}(d)$

Аналогично (1), подставим эмпирические значения и получим следующую численную зависимость от значения вероятности потерь для трафика приложения VoIP:

$$U_{VoIP}(PI) = 882 - 200 \cdot \ln(50 + 3 \cdot PI).$$

Подставляя численные значения в приведенное выражение, получаем графическую зависимость функции полезности от величины уровня потерь, рисунок 2.

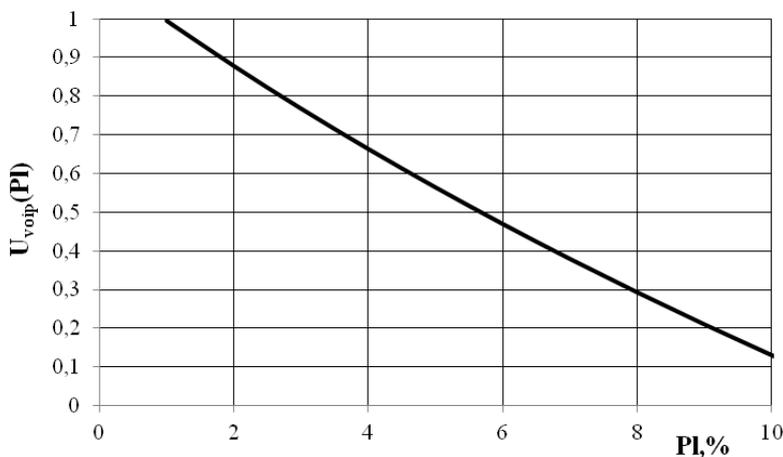


Рис. 2. График функции полезности $U_{VoIP}(PI)$

Для трафика VoIP качество обслуживания зависит от типа используемого кодека, поэтому при снижении минимально допустимой пропускной способности работа кодека останавливается и качество речи сильно ухудшается. Следовательно, функция полезности $U_{VoIP}(B)$ от величины пропускной способности будет иметь вид:

$$U_{VoIP}(B) = \frac{\text{sgn}(B - B_{\min}) + 1}{2},$$

где B_{\min} – минимально допустимая пропускная способность для заданного кодека.

Согласно общим характеристикам мультимедийных потоков, выражение общих функций полезности для VoIP могут использоваться для моделирования приложений потокового видео. Следующее выражение определяет функцию полезности задержки интерактивного видеопотока с порогами задержки 200 мс и 1 с:

$$U_{Video}(d) = \begin{cases} 100 - 0,004 \cdot d, & d < 200 \\ -25 \cdot \tanh((d - 625) / 200) + 75, & 200 \leq d \leq 1000 \\ 57 - 0,006 \cdot d, & d > 1000 \end{cases}$$

Для оценки задержки видеопотока, функция полезности выглядит следующим образом, рисунок 3.

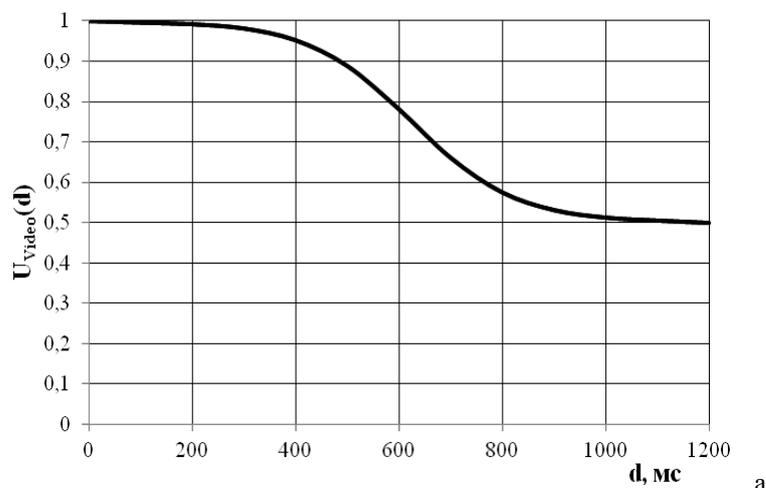


Рис. 3. График функции полезности $U_{Video}(d)$

Подставляя эмпирические данные, функция полезности уровня потерь для видеопотока примет вид:

$$U_{Video}(Pl) = 250 + 50 \cdot \ln(20 + 12,5 \cdot Pl) .$$

Функция полезности для видеопотока, в зависимости от величины пропускной способности, имеет следующий вид:

$$U_{Video}(B) = \frac{1}{1 + \left(\frac{1}{\varepsilon} - 1\right) \cdot e^{-r \cdot B}},$$

$$r = \frac{2 \ln\left(\frac{1}{\varepsilon} - 1\right)}{B_{\max}},$$

$$\varepsilon = \frac{B_{\min}}{B_{\max}} .$$

где B_{\min} – минимально допустимая пропускная способность;

B_{\max} – необходимая максимальная пропускная способность.

Благодаря изменениям стратегии игры для пользователя: «Стратегия в реальном времени» и «Шутер от первого лица» изменяются требования, предоставляемые конечными пользователями к задержке и времени отклика сети. Так, для второй стратегии предъявляются повышенные требования к времени отклика сети. Следовательно, именно на требования второй стратегии будут критериями принятия решений для формирования функций полезности. Задержка сильно влияет на производительность пользователя в онлайн-играх, особенно в играх типа «Шутер от первого лица». Из-за сильной чувствительности геймеров к сетевой задержке для моделирования подходит гиперболический тангенс (без расширенных переходов):

$$U_{Game}(D) = -45 \cdot \tanh(0,015 \cdot d - 1,35) + 60,$$

$$U_{Game}(Pl) = 100 - 50 \cdot \ln(1 + p_l),$$

$$U_{BFT}(B) = \frac{\log(B+1)}{\log(B_{\max}+1)}.$$

Выше представленные функции полезности позволяют задавать критерий использования соответствующего ресурса конечным пользователем. Кроме данных оценок, необходимо определить в какой мере, при исходных данных, сеть будет производительна для МОС.

Из-за разнородности оцениваемых критериев качества работы транспортной сети МОС предложено использовать комплексный показатель:

$$U = U_{QoS} \cdot U_{prf},$$

$$U_{QoS} = U_B(B) \cdot U_D(d) \cdot U_{Pl}(Pl), \quad (2)$$

$$U_{prf} = R \cdot Utl,$$

где U_{QoS} – композитная функция полезности;

U_{ef} – функция производительности;

$U_B(B), U_D(d), U_{Pl}(Pl)$ – функции «полезности» для пропускной способности, задержки, вероятности потерь соответственно;

Utl – оценка утилизации канала;

R – оценка надежности транспортной сети.

Для (2) действуют следующие выражения и ограничения:

$$\begin{aligned} U_B(B) &\rightarrow 1, \text{ при } B \leq B_{\max}; \\ U_D(d) &\rightarrow 1, \text{ при } d < d_{\text{дон}}; \\ U_{Pl}(Pl) &\rightarrow 1, \text{ при } Pl < Pl_{\text{дон}}. \end{aligned} \quad (3)$$

В (3) допустимые и максимальные значения выбираются согласно технологии используемой МОС.

Выражение (2) принимает максимальное значение «единица», если все параметры QoS удовлетворяют требованиям для этой услуги в полной мере. Если хотя бы один из параметров полностью не удовлетворен, то критерий принимает нулевое значение.

Для того, чтобы избежать ситуации, когда все наборы будут нулевыми, необходимо сортировать параметры качества обслуживания по важности для каждого класса услуг (чем важнее услуга, тем меньше ее номер). Это позволит пренебречь малозначающими показателями QoS. Кроме того, комплексный критерий учитывает нелинейную чувствительность трафика определенной услуги к параметрам QoS.

Проведена оценка эффективности предложенного алгоритма путем исследования соответствующих имитационных моделей. Исследования проводились для случая подключения любого абонента к одной базовой станции (рисунок 4).

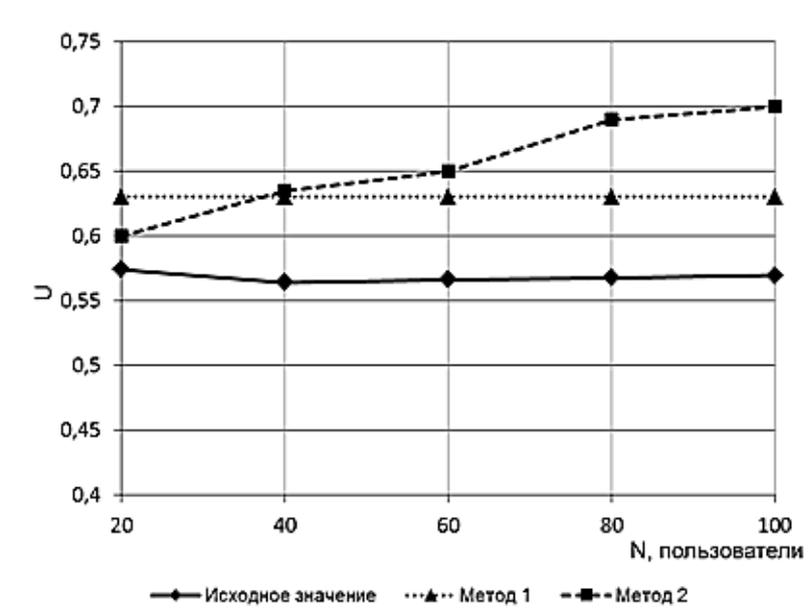


Рис. 4. Зависимость оценки качества работы транспортной сети от количества подключаемых пользователей

На рисунке: «Исходное значение» – практически измеренные значения оценок качества работы сети. В «Метод 1» представлен один из существующих методов прогнозирования. «Метод 2» – метод прогнозирования, использующий обоснованный комплексный критерий. Для «Метод 1» средняя абсолютная процентная ошибка составляет 36,3%, а для метода прогнозирования, использующего обоснованный критерий, «Метод 2», – 21%. Следовательно, используемый комплексный критерий, применяемый совокупно с соответствующими методами прогнозирования, позволяет повысить точность прогнозирования оценок качества работы транспортной сети мобильного оператора связи на 15%, по сравнению с существующими, при допустимых параметрах качества обслуживания.

Выводы. В статье предложено решение проблемы согласования параметров качества обслуживания и выделения для этого ресурсов сети мобильного оператора связи. Отмечено, что для уменьшения затрат на сбор, обработку и передачу служебных данных, можно воспользоваться комплексным показателем, учитывающим различные методы прогнозирования для получения зависимости оценок качества работы транспортной сети от количества подключаемых пользователей. Приведен и обоснован комплексный критерий, который, применяемый совокупно с соответствующими методами прогнозирования, позволяет повысить точность прогнозирования оценок качества работы транспортной сети мобильного оператора связи на 15%, по сравнению с существующими, при допустимых параметрах качества обслуживания.

СПИСОК ИСТОЧНИКОВ

1. Чепцов, М. Н. Подход к прогнозированию пропускной способности канала связи / М. Н. Чепцов, В. В. Климов // Вестник Донецкого национального университета. Серия Г: Технические науки. – 2024. – № 3. – С. 119-125. – DOI 10.5281/zenodo.14018650. – EDN DDOXJJ.
2. Климов, В. В. Метод прогнозирования оценки размера буфера контроллера базовой станции / В. В. Климов, М. Н. Чепцов // Проблемы искусственного интеллекта. – 2020. – № 3(18). – С. 49-59. – EDN QJUXUU.
3. Mu, M. A. Utility-Based QoS Model for Emerging Multimedia Applications / M. Mu, A. Mauthe, F. Garcia // The Second International Conference on Next Generation Mobile Applications, Services, and Technologies, 2008. – P. 521-528.

Поступила в редакцию 05.11.2024 г., рекомендована к печати 20.11.2024 г.

MOBILE COMMUNICATION OPERATOR'S TRANSPORT NETWORK QUALITY COMPREHENSIVE INDICATOR JUSTIFICATION

СHeptsov M.N., Lozinskaia V.N., Klimov V.V.

This work is aimed at substantiating the use of "utility" functions to obtain dependencies of the quality assessments of the mobile operator's transport network on the number of connected users. "Utility" functions allow moving from heterogeneous indicators of service quality parameters and network efficiency assessments to dimensionless coefficients. The proposed complex indicator allows obtaining the dependence of the efficiency on the number of connected users. The complex indicator, in combination with the corresponding predicting methods, allows both to effectively design and optimize the operation of the mobile operator's transport network.

Keywords: utility functions, mobile operator transport network, quality of service parameters, network performance assessments, complex indicator, predicting methods.

Чепцов Михаил Николаевич

доктор технических наук, профессор,
ректор ФГБОУ ВО «Донецкий институт
железнодорожного транспорта»,
Российская Федерация, ДНР, г. Донецк
E-mail: institut-transporta@mail.ru

СHeptsov Mikhail Nikolaevich

Doctor of Technical Science, Full Professor,
Rector of Donetsk Institute of Railway Transport,
Russian Federation, DPR, Donetsk.

Лозинская Виктория Николаевна

кандидат технических наук, доцент, доцент
кафедры автоматизации и телекоммуникации ФГБОУ
ВО «Донецкий национальный технический
университет»,
Российская Федерация, ДНР, г. Донецк.
E-mail: lvn_04@mail.ru

Lozinskaia Viktoriia Nikolaevna

Candidate of Technical Sciences, Associate
Professor, Associate Professor at department of
Automatics and Telecommunication of Donetsk
National Technical University,
Russian Federation, DPR, Donetsk.

Климов Владимир Владимирович

старший преподаватель кафедры автоматизации и
телекоммуникаций ФГБОУ ВО «Донецкий
национальный технический университет»,
Российская Федерация, ДНР, г. Донецк.
E-mail: boban_cc@mail.ru

Klimov Vladimir Vladimirovich

Senior Lecturer at Department of Automatics and
Telecommunication of Donetsk National Technical
University,
Russian Federation, DPR, Donetsk.

ИССЛЕДОВАНИЕ РАБОТЫ СИСТЕМЫ ПОМОЩИ ВОДИТЕЛЮ ПРИ ПЕРЕСТРОЕНИИ НА ОСНОВЕ РАДАРНОГО ДАТЧИКА

© 2024 *Киреев В.А., Бобынцев Д.О.*

Работа посвящена оценке эффективности ранее разработанного метода распознавания аварийных ситуаций в системе помощи водителю при перестроении путём использования известного радарного датчика. Проведены натуральные испытания метода на 4-полосной автомобильной дороге с участием двух автомобилей и использованием известного радарного датчика. Установлено, что время перестроения в реальных условиях вдвое больше измеренного путём имитационного моделирования, а предложенный ранее критерий определения манёвра перестроения не в полной мере отвечает реальным условиям дорожного движения. Предложены корректировки критерия для повышения эффективности распознавания аварийных ситуаций, сформулирован модифицированный метод обнаружения аварийной ситуации при перестроении на основе предложенных корректировок.

Ключевые слова: имитационная модель, системы помощи при перестроении, системы помощи водителю, системы активной безопасности, радарные датчики, безопасность дорожного движения, системы помощи водителю, условный боковой интервал, условная дистанция.

Введение. На сегодняшний день в мире автомобилестроения активно внедряются различные системы, направленные на повышение комфорта и безопасности водителей и пассажиров. Автопроизводители оснащают свои транспортные средства разнообразными датчиками мониторинга окружающей обстановки, включая радарные, которые играют важную роль в обеспечении безопасности на автомобильных дорогах. Радарные датчики используют радиоэлементы излучения и приёма сигнала для измерения расстояния, скорости, угла азимута и направления движения объектов [1]. С их помощью реализованы различные системы помощи водителю. Эти технологии не только расширяют арсенал средств безопасности для водителей, но также повышают уровень комфорта в различных ситуациях дорожного движения.

В частности, радарные датчики применяются в системах, позволяющих выявлять помехи в соседних полосах, в особенности при их попадании в так называемые «слепые зоны» боковых зеркал, когда помеха не видна в боковое зеркало из-за особенностей отражения света [2-5]. Применение радарных датчиков является более эффективным для распознавания манёвра перестроения по сравнению с системами, основанными на контроле полосы движения посредством видеодатчика и требующими наличия и достаточной видимости дорожной разметки [6, 7]. При этом стоит отметить, что даже на размеченных многополосных дорогах ширина одной полосы может допускать движение автомобилей по ней в два ряда, то есть быть достаточной, чтобы водители могли обеспечить безопасный боковой интервал при таком движении. В этом случае перемещение из ряда в ряд внутри такой полосы будет считаться перестроением, соответственно, требует соблюдения правил выполнения данного манёвра. Таким образом, полагаться при определении манёвра перестроения только на дорожную разметку нерационально.

Однако существующие системы помощи при перестроении, основанные на радарных датчиках, имеют предупреждающе-рекомендательный характер действия, не вмешиваются в процесс управления для предотвращения дорожно-транспортного происшествия, и определяют манёвр перестроения по включению указателей поворота

или повороту руля, тогда как водители не всегда пользуются указателями поворота при перестроении, а поворот руля осуществляется также при движении по закруглениям дороги большого радиуса. Для выработки реакции системы на ошибки водителя необходимо повышение точности распознавания манёвра перестроения. Для решения описанных проблем были разработаны метод и алгоритм распознавания аварийной ситуации при перестроении, основанные на контроле динамики условного бокового интервала между транспортными средствами в соседних полосах при помощи измерений, осуществляемых радарным датчиком, а также оценке безопасности условной дистанции между транспортными средствами [8]. Проведено также имитационное моделирование разработанного решения [9] и разработан программно-аппаратный комплекс для экспериментальной проверки решения в реальных условиях дорожного движения [10].

В соответствии с вышесказанным целью данного исследования является оценка эффективности и точности разработанного метода, сравнение с результатами имитационного моделирования, полностью описанного в работе [11], а также модификация разработанного метода при необходимости. Экспериментальная проверка проводится в реальных условиях дорожного движения с использованием известного радарного датчика. Полученные результаты позволят определить степень соответствия между данными, полученными входе имитационного моделирования, и реальными данными, скорректировать метод и алгоритм при необходимости, что является важным шагом к применению данного метода на практике для повышения безопасности на дорогах.

Для достижения цели исследования решены следующие задачи:

1. Тестирование расчёта введенных в работе [7] показателей, на которых базируется распознавание манёвра перестроения, в реальном времени в реальных условиях дорожного движения: условного бокового интервала, условной дистанции и скорости изменения условного бокового интервала. В системе помощи водителю данная задача будет возложена на процессорный блок управления.
2. Графическое представление результатов тестирования аналогично представлению результатов имитационного моделирования для сравнительного анализа.
3. Модификация разработанного метода по результатам проведенного исследования.

Экспериментальное исследование датчика, метода и алгоритма распознавания аварийной ситуации при перестроении. На основе разработанного ранее программно-аппаратного комплекса, описанного в статье [10], а также теоретических положений разработанного метода и алгоритма распознавания аварийной ситуации при перестроении проведены натуральные испытания с использованием реальных автомобилей и радарного датчика.

Натуральные испытания представляют собой экспериментальные перестроения, аналогично исследованиям имитационной модели в работе [11]. Перестроения совершены на 4-полосной автомобильной дороге крупного населённого пункта, имеющей по 2 полосы движения в каждом направлении, а также закругление большого радиуса на одном участке. Для испытаний было использовано два автомобиля. На один из автомобилей был установлен разработанный программно-аппаратный комплекс, а второй автомобиль использовался в качестве помехи, двигаясь в зоне видимости радарного датчика, в соседней полосе.

Предварительно на этом же автомобиле в качестве цели, ведомой датчиком, была проверена корректность измерений, выполняемых датчиком, после чего проведено тестирование работы метода и алгоритма распознавания аварийной ситуации. В результате была получена выборка данных объемом $n = 30$, равным количеству экспериментальных перестроений. Математическое описание соответствующего датасета S аналогично работе [11]:

$$S = \begin{pmatrix} b'_{11}(t) & b'_{12}(t) & \dots & b'_{1m}(t) \\ b'_{21}(t) & b'_{22}(t) & \dots & b'_{2m}(t) \\ \dots & \dots & \dots & \dots \\ b'_{n1}(t) & b'_{n2}(t) & \dots & b'_{nm}(t) \end{pmatrix},$$

где m – количество измерений скорости изменения условного бокового интервала $b'(t)$, а n – объём датасета (количество экспериментальных перестроений).

При этом окончанием манёвра перестроения считается полное занятие автомобилем соседнего ряда и возврат к только прямолинейному движению, без целенаправленного смещения вправо или влево. В результате анализа полученной выборки подтверждена выявленная имитационным моделированием вариативность общего времени манёвра, однако самое быстрое перестроение в проведенных тестах прошло за $t_n = 1,6$ с, что в 2 раза больше полученного при имитационном моделировании. Рассмотрим график изменения средних значений $b'(t)$, полученный по натуральным экспериментам, и сравним с аналогичным графиком, полученным путём имитационного моделирования (рис. 1).

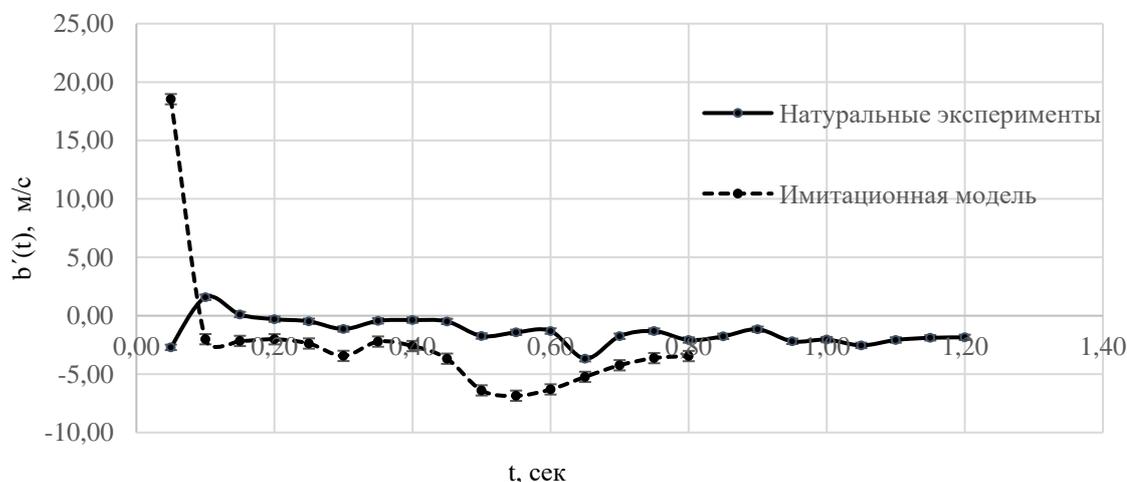


Рис. 1. График изменения $b'(t)$

Для вычисления доверительного интервала среднего генеральной совокупности применяется распределение Стьюдента. Уровень значимости (надёжности) принят равным 95%. При данном уровне надёжности доверительный интервал составляет 0,23. Как можно заметить по графику на основе натуральных экспериментов, динамика изменения $b'(t)$ в реальных условиях менее гладкая, чем на имитационной модели что может быть обусловлено различными факторами, не учтёнными в имитационной модели:

1. Поведение водителя при манёвре.
2. Курсовая устойчивость автомобиля.
3. Изъяны дорожного полотна, вынуждающие водителя совершать дополнительные микроманёвры уклонения.
4. Погрешности в работе датчика.

Указанные причины в ряде случаев приводят к тому, что датчик во время манёвра может показывать кратковременное увеличение условного бокового интервала после начала его снижения, что можно потенциально воспринимать как прекращение манёвра (таблица 1). Поэтому представляется целесообразным дополнительно фиксировать постоянное сокращение интервала на протяжении некоторого периода времени.

Таблица 1. Фрагмент датасета маневра перестроения

Время маневра, с	Значения производной (скорости изменения условного бокового интервала), м	Значения условного бокового интервала, м
0,05	-5,2	2,23
0,1	4,4	2,45
0,15	-1,6	2,37
0,2	0	2,37
0,25	-2,4	2,25
0,3	-0,8	2,21
0,35	1,4	2,28
0,4	-2,6	2,15
0,45	-0,8	2,11
0,5	-1	2,06
0,55	0	2,06
0,6	-1,2	2,00
0,65	-1,4	1,93
0,7	-1,4	1,86
0,75	-1,6	1,78
0,8	-1,6	1,70

Также можно сделать вывод, что выбранное по результатам имитационного моделирования пороговое значение $b'_n(t) = 3,39$ м/с не подходит для фиксации манёвра, так как перестроение может происходить с меньшей скоростью изменения бокового интервала на протяжении всего манёвра, и говорить о гарантированном достижении данного значения невозможно.

Кроме того, собственные аналитические исследования поведения водителей при перестроении привели к выводу, что для удобства вождения целесообразно оставлять водителю время на самостоятельную реакцию на предупреждение об опасности, если его быстрота реакции позволяет это сделать. Тогда, чем быстрее система обнаружит манёвр, тем лучше.

Для выбора предлагаемого времени на обнаружение манёвра разделим время перестроения на три основных временных отрезка:

1. Обнаружение потенциально опасного манёвра и выдача предупреждения, например, вибромотором руля.
2. Ожидание реакции водителя. После выдачи предупреждения система ожидает реакцию водителя. Этот временной отрезок отражает интервал, в течение которого водитель должен обнаружить предупреждение и предпринять соответствующие меры. Индивидуальное время реакции водителя может варьироваться от 0,4 до 1,5 секунд [12].

Однако при принятии решения об ожидании реакции водителя после распознавания маневра и выдачи предупреждения необходимо учитывать текущую условную дистанцию и боковой интервал. Например, если помеха уже находится в «слепой зоне» боковых зеркал, реагировать нужно немедленно, и данный этап должен быть исключён.

3. Реакция системы при отсутствии реакции водителя. Если водитель не реагирует на предупреждение или не предпринимает достаточных действий, система должна автоматически вмешаться в процесс управления, чтобы предотвратить столкновение.

Продолжительность данных отрезков необходимо выбирать исходя из примерного минимального времени манёвра, а также среднего времени реакции водителя и времени срабатывания тормозной системы и рулевого управления. Предполагается, что время срабатывания гидравлической тормозной системы составляет 0,1 – 0,4 с, а на поворот руля машина реагирует практически мгновенно, тогда на этап 3 отведём 0,5 с.

Чтобы определить минимальное время на два других временных отрезка, выбрано минимальное время перестроения из выборки S , которое составило 1,6 с. Для дополнительной проверки проведено ручное измерение времени перестроения двумя водителями по следующей схеме:

1. Водитель 1 находится за рулём и совершает 20 перестроений.
2. Водитель 2 находится в кресле пассажира и выступает в роли хронометриста.
3. Водитель 1, совершая перестроение, сообщает хронометристу о начале и конце манёвра, хронометрист фиксирует измеренное секундомером время.

После завершения цикла водители меняются местами. Были получены следующие результаты: $1,73 \leq T_1 \leq 2,90$, $1,65 \leq T_2 \leq 3,28$, $T_{cp1} = 2,25$ с, $T_{cp2} = 2,33$ с. Таким образом, в реальных условиях перестроение быстрее, чем за 1,5 с, маловероятно, поэтому предлагается считать это значение нижним порогом продолжительности манёвра. Тогда на распознавание манёвра и ожидание реакции водителя остаётся 1 с.

Детальный анализ данных каждого экспериментального перестроения из датасета S показал наличие в каждом случае временных отрезков в 0,3 с, на протяжении которых фиксируется постоянное сокращение условного бокового интервала. Самый поздний случай первого появления такого временного отрезка – 1,05 с от начала манёвра при общей продолжительности этого манёвра 2,1 с. Исходя из такого факта предлагается ограничение времени распознавания маневра в 0,3 с от первого отрицательного значения $b'(t)$, указывающего на сокращение условного бокового интервала. Комбинирование проверки постоянного сокращения условного бокового интервала с проверкой по пороговому значению в 1 м/с позволит распознавать манёвр быстрее и эффективнее по сравнению с критерием, разработанным ранее.

Следует отметить, что различные типы транспортных средств имеют различные динамические характеристики, что влияет на выбор безопасной дистанции между ними. Например, у грузовых автомобилей с пневматическими тормозами время срабатывания тормозной системы больше, чем у легковых автомобилей с гидравлическими тормозами. Исходя из этого, важно иметь представление о типе помехи при оценке безопасной дистанции во время движения в транспортных потоках смешанного типа, где могут присутствовать разные типы транспортных средств. Для идентификации типов транспортных средств предлагается использовать лидарный датчик, с помощью которого можно получить представление о габаритах цели, позволяющих классифицировать её.

Исследуемый метод обнаружения аварийной ситуации при перестроении также предполагает оценку безопасности условной дистанции по правилу 3 секунд, которое позволяет в большинстве случаев избежать фронтального столкновения независимо от множества внешних факторов. Однако наблюдения показывают, что, если учитывать по этому правилу только абсолютную скорость автомобиля, который движется позади, такая дистанция для многих водителей будет существенно больше реально безопасной, и может доставлять серьёзные неудобства при управлении. Например, при скорости 60 км/час безопасная дистанция по правилу 3 секунд равна 50 м. Поэтому предлагается при оценке безопасности условной дистанции использовать не абсолютную скорость помехи, а относительную, измеряемую радарным датчиком, которая показывает разницу в абсолютных скоростях помехи и автомобиля с радаром, и выглядит более адекватным показателем, чем абсолютная.

На завершающем этапе исследования проведена аналитическая оценка влияния погрешности измерения угла азимута при расположении помехи позади автомобиля с радаром и при движении по криволинейным участкам дороги с большим радиусом поворота. Для этого выбрано среднее значение бокового интервала между автомобилями в соседних полосах $b_{\text{ср}} = 2$ м. По данным экспериментальных перестроений погрешность измерения угла азимута после перестроения помехи в полосу автомобиля с радаром находится в пределах 5 градусов. Определим аналитически, на каких расстояниях R и условных дистанциях D при боковом интервале $b = 2$ м будут такие же значения угла азимута θ , что может ввести бортовой компьютер в заблуждение. Результаты представлены в таблице 2.

Таблица 2. Значения расстояния и условной дистанции при $\theta \leq 5^\circ$ и $b = 2$ м

Угол азимута	SIN	COS	Расстояние R, м	Условная дистанция D, м
1	0,017	1,000	114,597	114,580
2	0,035	0,999	57,307	57,273
3	0,052	0,999	38,215	38,162
4	0,070	0,998	28,671	28,601
5	0,087	0,996	22,947	22,860

Таким образом, на расстояниях, превышающих зону видимости датчика (30 м), $D \approx R$, то есть треугольник определения необходимых геометрических величин сжимается в прямую линию, что говорит не только о возможной двусмысленности понимания угла азимута, выдаваемого радаром, но и о сложности отслеживания динамики изменения бокового интервала, поскольку требуется измерение угла с более высокой точностью, до сотых долей градуса. В связи с этим возникает вопрос о способности какого-либо радарного датчика обеспечить требуемую точность измерений.

Также на расчёт необходимых величин может оказывать влияние искривление траекторий движения автомобилей в результате движения по криволинейным участкам дороги с большим радиусом поворота. Тем не менее, проведённые натуральные эксперименты показали, что на расстоянии, соответствующем зоне видимости данного радарного датчика, движение автомобилей по таким участкам не оказывает существенное воздействие на результаты измерений. Это связано с тем, что, как правило, на всём протяжении таких участков дороги отрезок длиной 30 м выглядит почти как прямая линия (рис. 2). При большем удалении автомобилей друг от друга на таком участке дороги говорить о параллельности их траекторий движения, необходимой для корректного расчёта условной дистанции и условного бокового интервала, будет невозможно. Поэтому оба описанных фактора приводят к выводу, что использование датчика большей дальности нецелесообразно.



Рис. 2. Зона обнаружения радарного датчика на криволинейном участке дороге

Исследования также показали, что в случае перестроения помехи в полосу автомобиля, оснащённого радарным датчиком, регистрируемые значения радарного датчика совпадают с показателями, фиксируемыми при перестроении самого автомобиля с установленным радаром, что означает аналогичную реакцию бортового компьютера на перестроения, совершаемые соседями по полосам. Для исключения подобных ложных срабатываний при определении аварийной ситуации предлагается использовать комбинацию критерия сокращения условного бокового интервала и критерия изменения угла поворота руля в сторону помехи, для чего в систему помощи необходимо интегрировать датчик угла поворота руля. Сформулируем окончательные варианты данных критериев.

Предлагаемый вариант критерия сокращения условного бокового интервала выглядит следующим образом: условный боковой интервал непрерывно уменьшается на протяжении 0,3 с или скорость сокращения условного бокового интервала в произвольный момент времени более 1 м/с по модулю. Непрерывное уменьшение условного бокового интервала означает, что $b'(t) < 0$. Тогда критерий математически описывается формулой:

$$(b'(t) < 0) \vee (b'(t) < -1), \\ \Delta t = 0,3$$

где Δt – заданный интервал времени наблюдения $b'(t)$.

Определим критерий изменения угла поворота руля:

1. Установлено опытным путем, что при перестроении в соседнюю полосу угол поворота руля α изменяется на величину от 5 градусов. То же самое наблюдается при движении автомобиля по закруглению дороги большого радиуса.
2. Таким образом, $\alpha \geq 5^\circ$ означает одно из двух: начало маневра перестроения или вхождение в поворот большого радиуса. Данный фактор предлагается комбинировать с фактором сокращения условного бокового интервала.

На основе полученных определений выведем логическую формулу комбинированного критерия распознавания манёвра перестроения. Пусть X_1 – значение критерия сокращения условного бокового интервала, а X_2 значение критерия изменения угла поворота руля. X_1 и X_2 имеют логический формат. Рассмотрим варианты комбинаций X_1 и X_2 :

1. $X_1 = 0, X_2 = 0$ (сокращения условного бокового интервала нет, и изменения угла поворота руля нет): ничего потенциально опасного не происходит.
2. $X_1 = 0, X_2 = 1$ (сокращения условного бокового интервала нет, а изменение угла поворота руля есть): происходит только движение по закруглению дороги большого радиуса, опасности нет.
3. $X_1 = 1, X_2 = 0$ (сокращение условного бокового интервала есть, а изменения угла поворота руля нет): манёвр перестроения происходит со стороны помехи, опасности нет.
4. $X_1 = 1, X_2 = 1$ (сокращение условного бокового интервала есть, и изменение угла поворота руля есть): происходит манёвр перестроения со стороны автомобиля с системой помощи, необходимо оценить его опасность.

Таким образом, манёвр перестроения фиксируется только при выполнении обоих критериев, что соответствует функции «логическое И» и описывается математически формулой:

$$Y = X_1 \wedge X_2,$$

где Y - логическое значение, указывающее на фиксацию манёвра перестроения, X_1 - критерий сокращения условного бокового интервала, X_2 - критерий изменения угла поворота руля.

Или в развернутом виде формулой:

$$Y = ((b'(t) < 0) \vee (b'(t) < -1)) \wedge (\alpha \geq 5), \quad (1)$$

$\Delta t = 0,3$

где Y - логическое значение, указывающее на фиксацию манёвра перестроения,

Δt – заданный интервал времени наблюдения $b'(t)$, α – угол поворота руля.

В соответствии с изложенным выше, по результатам проведённого исследования метод обнаружения аварийной ситуации скорректирован следующими положениями:

1. Изменить критерий распознавания манёвра: необходимо использовать уменьшенное пороговое значение $b'_n(t) = 1$ м/с и одновременно фиксировать постоянное снижение бокового интервала на протяжении 0,3 секунд от первого снижения, выбирая то из двух, что наступит раньше, чтобы получить достаточно времени на реакцию со стороны водителя, использовать комбинацию критерия сокращения условного бокового интервала и критерия изменения угла поворота руля в сторону помехи.
2. Оценивать безопасность условной дистанции по относительной скорости помехи, получаемой от радарного датчика.

Таким образом, модифицированный метод обнаружения аварийной ситуации выглядит следующим образом:

1. Обнаружение помехи в соседней полосе с использованием двухантенного радарного датчика.
2. Измерение угла поворота руля соответствующим датчиком в малые промежутки времени, фиксация изменений угла поворота процессорным блоком.

3. Определение относительной скорости помехи, расстояния до неё и угла азимута датчиком, расчёт условной дистанции по расстоянию и углу азимута процессорным блоком, оценка безопасности условной дистанции по правилу трёх секунд.
4. Расчёт условного бокового интервала процессорным блоком по расстоянию и углу азимута в малые промежутки времени, вычисление скорости сокращения условного бокового интервала.
5. При первом $b'(t) < 0$ проведение мониторинга $b'(t)$ на протяжении 0,3 с на предмет постоянства отрицательного значения, либо достижения $b'(t) < -1$. Если выполняется критерий (1), и условная дистанция при этом не безопасна, принимается решение об аварийной ситуации, на которую необходимо отреагировать. Иначе манёвр считается безопасным и не требует вмешательства в процесс управления.

Выводы. На основании проделанной работы можно вынести следующие выводы:

1. Установлено, что время перестроения в реальных условиях в 2 раза больше, чем полученное в имитационной модели.
2. На расстояниях в десятки метров до помехи угол азимута $\theta \rightarrow 0$, что можно принять за погрешность измерения, затрудняющую определение взаимного расположения автомобиля. Это значит, что сканирование пространства на расстояние более чем 30 м нецелесообразно.
3. Значительных отличий в измерениях на участках прямолинейного движения и на криволинейных участках дороги не наблюдается.
4. Выявлена необходимость коррекции критериев, используемых в методе обнаружения аварийной ситуации, с целью повышения эффективности распознавания и устранения потенциальных проблем в управлении транспортным средством, предложены варианты коррекции.

СПИСОК ИСТОЧНИКОВ

1. Сысоева, С. Актуальные технологии и применения датчиков автомобильных систем активной безопасности. Часть 6 Радары / С. Сысоева // Компоненты и технологии. – 2007. – № 3(68). – С. 67-76.
2. Троицкий, А. М. Системы мониторинга дорожного движения и помощи водителю. Радары диапазонов 24 ГГц и 77 ГГц / А. М. Троицкий // перспективные разработки по приоритетным направлениям развития: сборник статей III Международного научно-исследовательского конкурса, Петрозаводск, 15 марта 2022 года. – Петрозаводск: Международный центр научного партнерства «Новая Наука» (ИП Ивановская И.И.), 2022. – С. 78-87. – EDN BLJWFQ.
3. Система контроля слепых зон автомобиля / А. А. Сидоров, В. С. Шувалов, М. Д. Свиная, Р. В. Безносюк // Научно-техническое обеспечение технологических и транспортных процессов в АПК : Материалы Международной научно-практической конференции, посвященной памяти доктора технических наук, профессора, заслуженного деятеля науки и техники РСФСР, академика РАТ Николая Николаевича Колчина, Рязань, 24 мая 2024 года. – Рязань: Рязанский государственный агротехнологический университет имени П. А. Костычева, 2024. – С. 163-170. – EDN QWYEVK.
4. Пономарев, М. В. Продвинутое системы помощи водителю / М. В. Пономарев // Направления развития технического сервиса : Материалы региональной студенческой научно-практической конференции, посвященной 80-летию Уральского государственного аграрного университета и 70-летию факультета инженерных технологий, Екатеринбург, 27 ноября – 27 декабря 2019 года. – Екатеринбург: Уральский государственный аграрный университет, 2020. – С. 81-87. – EDN IOPUUE.
5. Ковалев, О. Ф. Система контроля за движением автомобиля по полосам дороги / О. Ф. Ковалев, К. И. Зеленков // Компьютерные технологии и анализ данных (СТДА'2022): материалы III Международной научно-практической конференции, Минск, 21–22 апреля 2022 года. – Минск: ГУО «Республиканский институт высшей школы», 2022. – С. 207-210. – EDN HJXIFI.
6. Киреев, В. А. Исследование аварийных ситуаций при попутном движении на многополосных автомобильных дорогах / В. А. Киреев, Д. О. Бобынцев // Интеллектуальные информационные системы: тенденции, проблемы, перспективы (ИИС - 2022) : Сборник материалов IX Всероссийской научно-практической конференции, Курск, 13–15 сентября 2022 года. – Курск: Юго-Западный государственный университет, 2022. – С. 76-80.

7. Киреев, В. А. Метод и алгоритм распознавания аварийной ситуации при перестроении автомобильными системами активной безопасности / В. А. Киреев, Д. О. Бобынцев // Экономика. Информатика. – 2022. – Т. 49, № 4. – С. 798-809.
8. Бобынцев, Д. О. Обнаружение аварийной ситуации при перестроении автомобильными системами активной безопасности / Д. О. Бобынцев, В. А. Киреев // Автомобили, транспортные системы и процессы: настоящее, прошлое и будущее : Сборник статей 4-й Международной научно-технической конференции, Курск, 20 мая 2022 года / Отв. редактор Е.В. Агеев. – Курск: Юго-Западный государственный университет, 2022. – С. 27-30.
9. Киреев, В. А. Имитационное моделирование метода обнаружения аварийной ситуации при перестроении на автомобильных дорогах / В. А. Киреев, Д. О. Бобынцев // Оптико-электронные приборы и устройства в системах распознавания образов и обработки изображений : сборник материалов XVII Международной научно-технической конференции, Курск, 12–15 сентября 2023 года. – Курск: Юго-Западный государственный университет, 2023. – С. 121-123.
10. Бобынцев, Д. О. Программно-аппаратный комплекс тестирования алгоритма работы вычислительной системы помощи водителю при перестроении / Д. О. Бобынцев, В. А. Киреев // Информационные системы и технологии. – 2024. – № 5(145). – С. 83-92.
11. Бобынцев, Д. О. Имитационная модель распознавания аварийной ситуации при перестроении на автомобильных дорогах / Д. О. Бобынцев, В. А. Киреев // Информационные системы и технологии. – 2023. – № 5(139). – С. 5-14. – EDN WUZPMG.
12. Кольчурин, А. Г. Определение времени реакции водителя при расследовании дорожно-транспортных преступлений / А. Г. Кольчурин // Общество и право. – 2016. – № 4(58). – С. 128-130.

Поступила в редакцию 07.11.2024 г., рекомендована к печати 22.11.2024 г.

EXPERIMENTAL STUDY OF THE METHOD OF EMERGENCY DETECTION BY RADAR SENSOR DURING REALIGNMENT

Kireev V.A., Bobyntsev D.O.

The study focuses on evaluating the effectiveness of a previously developed method for recognizing hazardous situations in driver assistance systems during lane changes using a well-known radar sensor. Field tests of the method were conducted on a four-lane highway involving two vehicles and employing the specified radar sensor. It was found that the lane change duration in real-world conditions is twice as long as the time measured through simulation modeling. Additionally, the previously proposed criterion for determining lane-change maneuvers does not fully correspond to real-world traffic conditions. Adjustments to the criterion are proposed to enhance the recognition efficiency of hazardous situations, and a modified method for detecting hazardous situations during lane changes based on these adjustments is formulated.

Keywords: simulation model, lane-change assistance systems, driver assistance systems, active safety systems, radar sensors, road traffic safety, driver support systems, conditional lateral gap, conditional distance.

Киреев Вадим Андреевич

аспирант кафедры вычислительной техники
ФГБОУ ВО «Юго-Западный государственный
университет»,
Российская Федерация, г. Курск.
E-mail: avadimsa@yandex.ru

Kireev Vadim Andreevich

Postgraduate Student at Department of Computer
Engineering of South-West State University,
Russian Federation, Kursk.

Бобынцев Денис Олегович

кандидат технических наук, доцент кафедры
вычислительной техники ФГБОУ ВО «Юго-
Западный государственный университет»,
Российская Федерация, г. Курск.
E-mail: danis-artist@vt-swsu.ru.

Bobyntsev Denis Olegovich

Candidate of Technical Sciences, Associate
Professor at Department of Computer Engineering of
South-West State University
Russian Federation, Kursk.

ОДНОФАЗНЫЙ КОНВЕКТИВНЫЙ ТЕПЛОПЕРЕНОС В КОМПАКТНОМ ТЕПЛООБМЕННИКЕ, ЗАПОЛНЕННОМ АНИЗОТРОПНОЙ ПОРИСТОЙ СРЕДОЙ© 2024 *Малеваный М.В., Коновалов Д.А., Кожухов Н.Н., Черниченко В.В.*

В работе предложена оригинальная конструкция пористого компактного теплообменного элемента, состоящего из слоев различной пористости, изменяющейся как по высоте теплообменника, так и по его длине. Проведено численное моделирование различных конструкций теплообменных элементов, а также показано влияние расположения пористых слоев на эффективность интенсификации. Сравнение с известными результатами в предельных случаях позволили сделать вывод о корректности применения численной модели.

Ключевые слова: пористая среда, развитая поверхность теплообмена, высокопористые материалы, интенсификация, анизотропность.

Введение. Компактные теплообменники на основе пористых сред нашли широкое применение в различных предметных областях. Это системы термостабилизации для микроэлектроники и радиоэлектроники, компактные энергетические установки, в т.ч. для «зеленой» энергетики, системы тепловой защиты авиационных и космических систем. Особое внимание уделяется средам с высокой пористостью, которые характеризуются компактностью, высоким значением коэффициента теплоотдачи и гидравлическим сопротивлением, сопоставимым со значением для гладкого канала. Отдельное место занимает исследование так называемых градиентных пористых сред, когда теплообменный элемент представляет собой послойную конструкцию состоящей из нескольких горизонтальных слоев или вставок с различной пористостью, расположенных по ходу течения теплоносителя (рис. 1) [1].

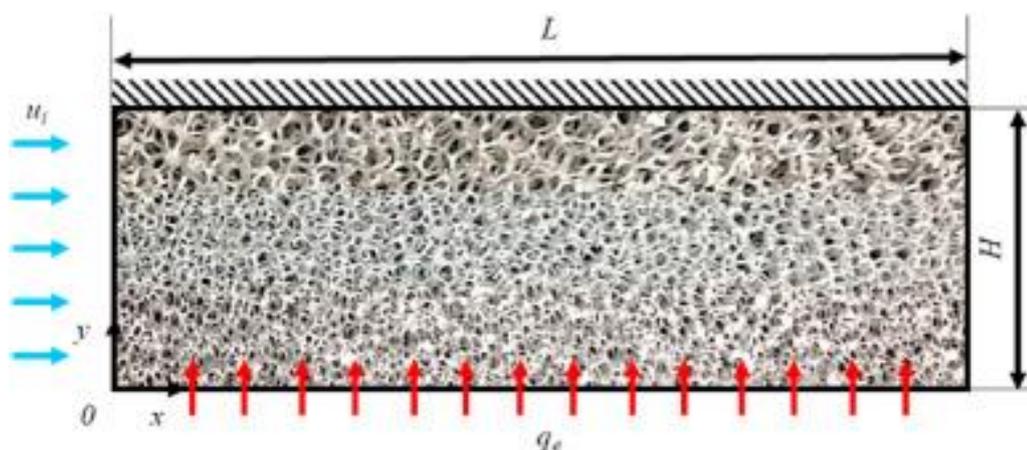


Рис. 1. Пример среды с переменной пористостью [1]

Исследования отечественных и зарубежных авторов показывают, что использование такого способа интенсификации является оправданным [2-4].

Однако по мере течения теплоносителя через канал тепловые поля меняются в продольном и поперечном направлениях. Актуальной является оценка эффективной интенсификации теплообмена в теплообменнике заполненном пористой средой при наличии геометрической анизотропии как по высоте канала, так и по его длине. Экспериментальная модель гибридного теплообменника может быть сравнительно просто реализована за счет использования готовых образцов металлических пен или с применением современных технологий 3D печати на промышленных металлических принтерах.

Постановка задачи. Физическая область решения представлена как плоский прямоугольный канал, заполненный анизотропной пористой средой (рис. 2).

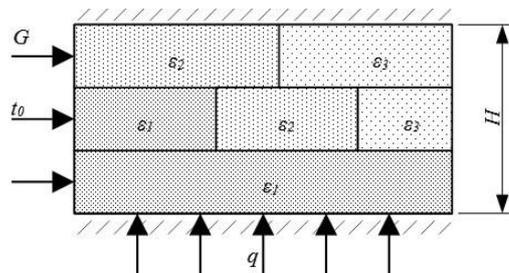


Рис. 2. Физическая область решения для теплообменника с переменной пористостью. Соотношения длин, высот и пористостей канала являются условными

Теплоноситель с массовым расходом G движется с начальной температурой t_0 по каналу высотой H , заполненным средой с изменяющейся пористостью. Нижняя стенка канала является теплонапряженной и на ней заданы граничные условия второго рода (удельный тепловой поток q). Высота канала разделена на три части, каждая из которых имеет свою структуру пористости. В целом для всего теплообменного элемента пористость изменяется как по длине канала, так и по его высоте. Фазовые переходы отсутствуют, коэффициент смачиваемости поверхностей одинаков и остается постоянным. Контактным термическим сопротивлением на границах пористых сред пренебрегаем, полагая, что технология 3д печати позволяет получить единую монолитную структуру, указанную на рис. 2. Задача решается в двухмерной стационарной постановке. В качестве модели теплопередачи заложена двухтемпературная модель, предполагающая локальное тепловое неравновесие между пористым каркасом и теплоносителем.

Результаты численного моделирования. Расчеты производились в среде Comsol для двух моделей, представленных на рис. 3 (а, б) и отличающихся между собой комбинациями слоев по высоте и значениями пористости.

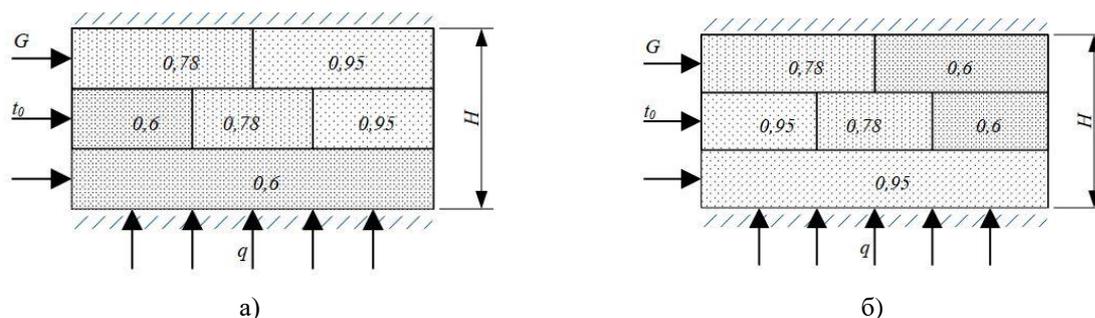


Рис. 3. Расчетные модели (цифровые обозначения – величина пористости):
а) слой с низкой пористостью прилегает к теплонапряженной поверхности (тип матрицы А)
б) слой с высокой пористостью прилегает к теплонапряженной поверхности (тип матрицы Б)

При расчетах использовались следующие исходные данные. Теплоноситель – вода. Исследуемый диапазон скорости теплоносителя составлял $u_0 = 0,02 \dots 0,2$ м/с, начальная температура теплоносителя $t_0 = 20^\circ\text{C}$, высота пористого канала $H = 6$ мм (высота каждого подслоя 2 мм), длина пористого канала 40 мм, удельный тепловой поток на тепловыделяющей поверхности $q = 100$ Вт/см², материал пористого канала – алюминий $\lambda = 238$ Вт/(м·К). Диапазон пористости отдельных элементов изменялся от 0,6 до 0,95.

Для моделирования рассматриваемой задачи рассматривалось неизотермическое течение. Для решения гидродинамической подзадачи использовалась модель Бринкмана для несжимаемой жидкости и уравнение неразрывности, влиянием сил свободной конвекции пренебрегали. Наиболее значимой характеристикой пористой среды является проницаемость, модель которой была взята из модели Козени-Кармана [5]:

$$\kappa = \frac{d_p^2}{180} \cdot \frac{\varepsilon_p^3}{(1 - \varepsilon_p)^2}.$$

Для построения расчетной сетки использовался автоматический способ с учетом физики задачи (течение сплошной среды). Показателем ее качества выбран способ skewness. Сама расчетная сетка и некоторые ее параметры представлены на рис. 4 и в таблице 1.

Таблица 1. Некоторые характеристики качества расчетной сетки

Название	Значение
количество элементов	4258
минимальное значение качества сетки	0,36
среднее значение качества сетки	0,81

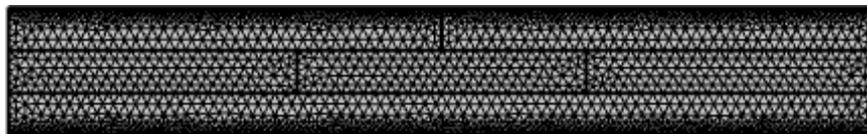


Рис. 4. Расчетная сетка

Для обобщения гидравлических характеристик пористого элемента находилась зависимость $\xi(\text{Re})$. Критерий Рейнольдса и коэффициент гидравлического сопротивления при этом определялись по следующим формулам:

$$\text{Re} = \frac{\rho_f u_0 h}{\mu_f \varepsilon^2},$$

$$\xi = \frac{2d_r \Delta p}{l \rho_f u_0^2}.$$

Для пористой среды принималась двухтемпературная модель, поэтому в модели для расчета необходимо было определить дополнительные параметры. Площадь поверхности контакта между пористой матрицей и теплоносителем A_{fs} , 1/м [6]:

$$A_{fs} = \frac{6(1-\varepsilon)}{d_p}.$$

Коэффициент теплоотдачи между пористой матрицей и теплоносителем h_{fs} , Вт/(м²·К) [6]:

$$h_{fs} = \frac{\lambda_f}{d_p} \left[2 + 1,1Pr^{1/3} \left(\rho_f \frac{u_0 d_p}{\mu_f} \right)^{0,6} \right],$$

где $Pr = \frac{\mu_f c_{pf}}{\lambda_f}$ [6].

При этом скорость теплоносителя определялась из уравнения неразрывности потока. Результаты решения гидродинамической задачи в виде полей скоростей теплоносителя и давления для скорости $u_0 = 0,2$ м/с представлены на рис. 5-7.

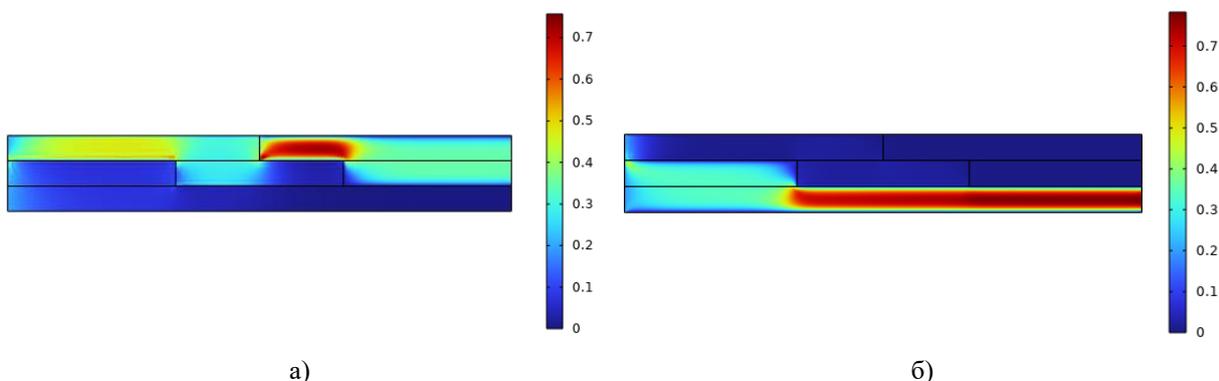


Рис. 5. Поля скоростей в компактном теплообменнике:
а) тип матрицы А б) тип матрицы Б

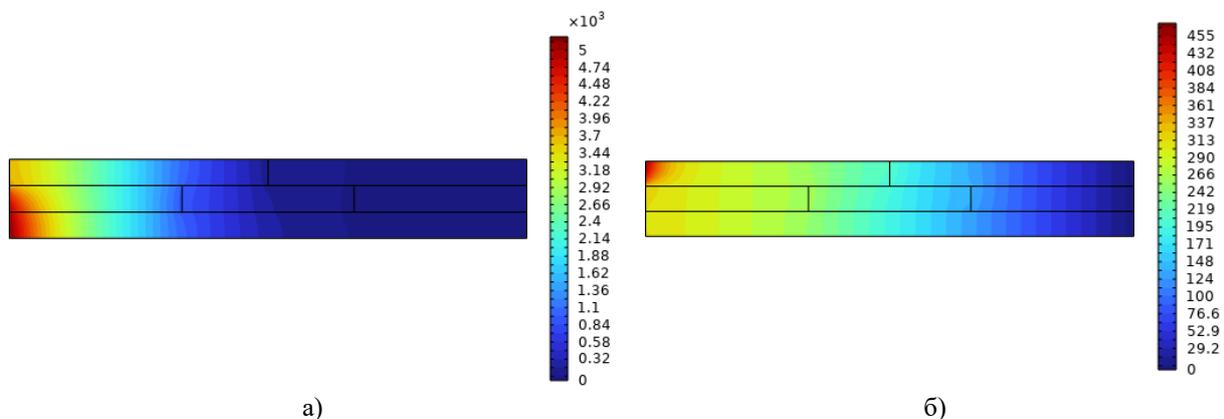


Рис. 6. Поля давлений в компактном теплообменнике:
а) тип матрицы А б) тип матрицы Б

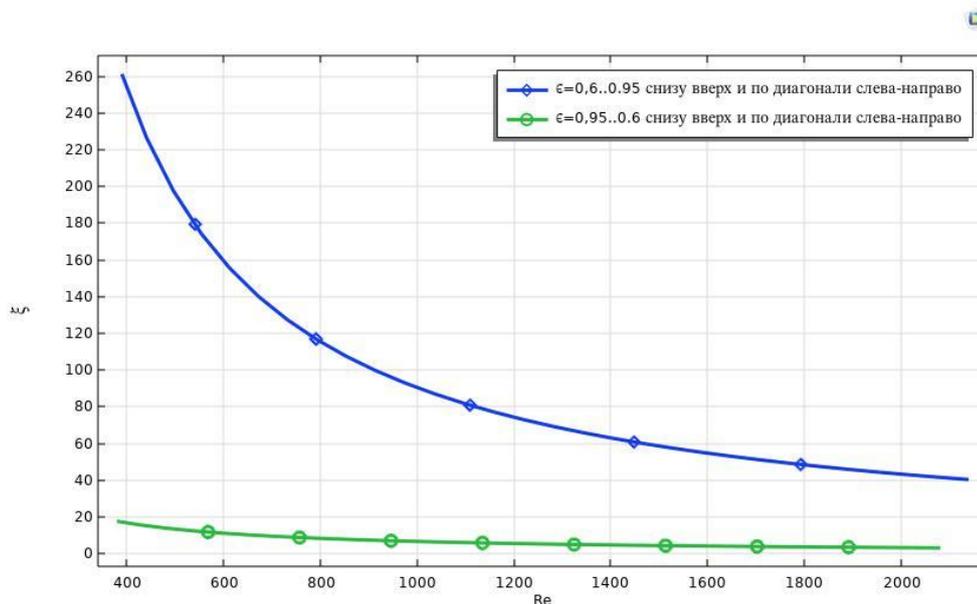


Рис. 7. Зависимость коэффициента гидравлического сопротивления от числа Рейнольдса для различных типов матриц

Результаты показали, что минимальным гидравлическим сопротивлением обладает пористая матрица типа Б, т.е. слои с большим значением пористости располагаются непосредственно с теплонапряженной поверхностью. Результаты численного моделирования тепловой подзадачи представлены на рис. 8-10.

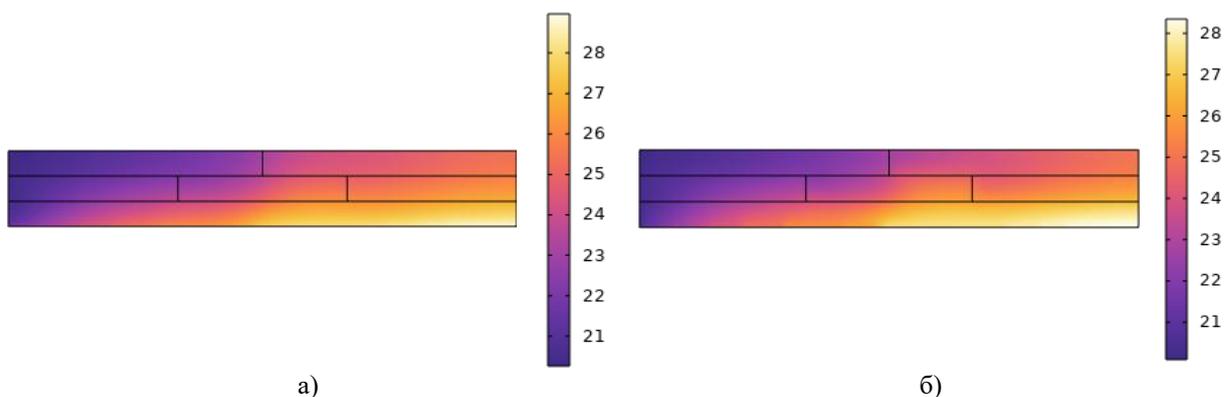


Рис. 8. Поля температур в пористой матрице (а) и теплоносителе (б). Тип матрицы А

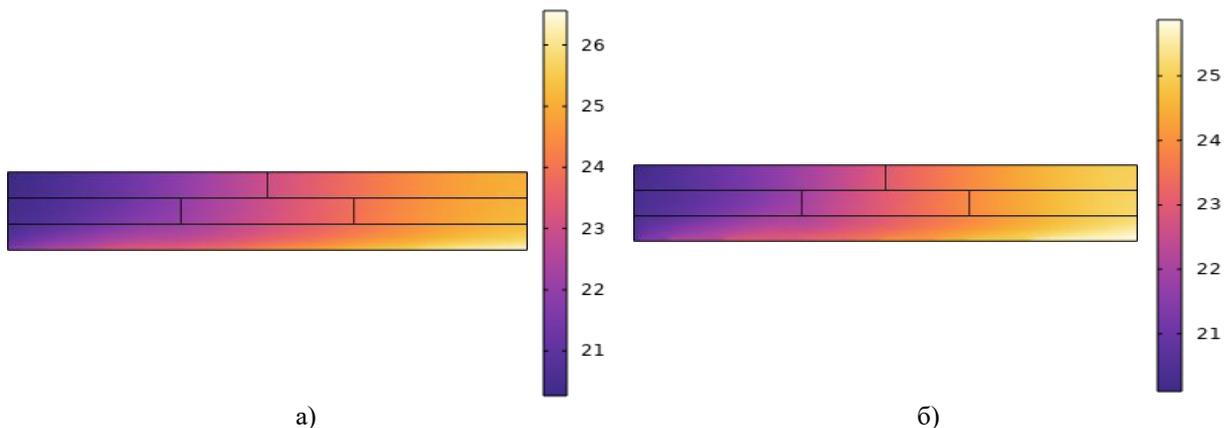


Рис. 9. Поля температур в пористой матрице (а) и теплоносителе (б). Тип матрицы Б

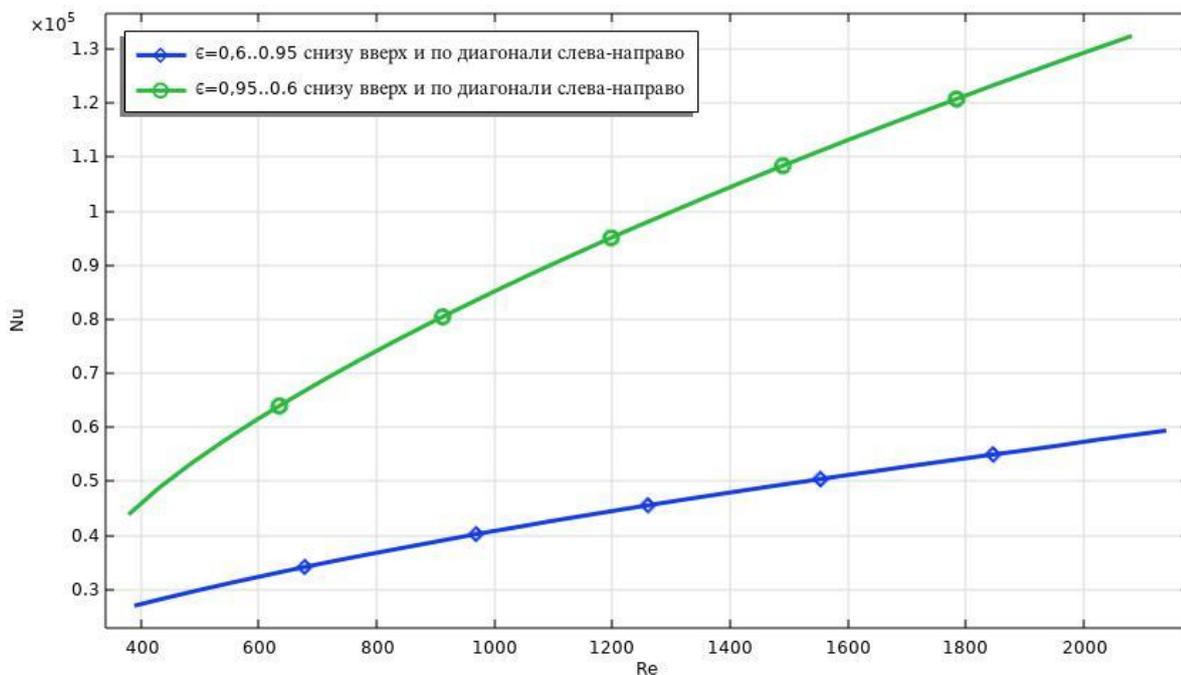


Рис. 10. Зависимость числа Нуссельта от числа Рейнольдса для различных типов матриц

Результаты показали, что наилучший результат по теплообмену также дает матрица типа Б. Для верификации полученных результатов проводилось сравнение с результатами работ для послойного (по высоте канала) изменения пористости [1]. В качестве исходной рассматривалась модель, приведенная на рис. 1. Изменение пористости происходит только лишь по высоте канала.

Установлено, что послойное изменение пористости канала по высоте дает эффект интенсификации теплообмена по сравнению со случаем, если бы пористость оставалась постоянной (увеличение коэффициента теплоотдачи порядка 15-20%). В то же время, увеличение пористости в целом несколько снижает коэффициент теплоотдачи, что возможно связано со снижением коэффициента эффективной теплопроводности каркаса в целом. Результаты верификации представлены на рис. 11.

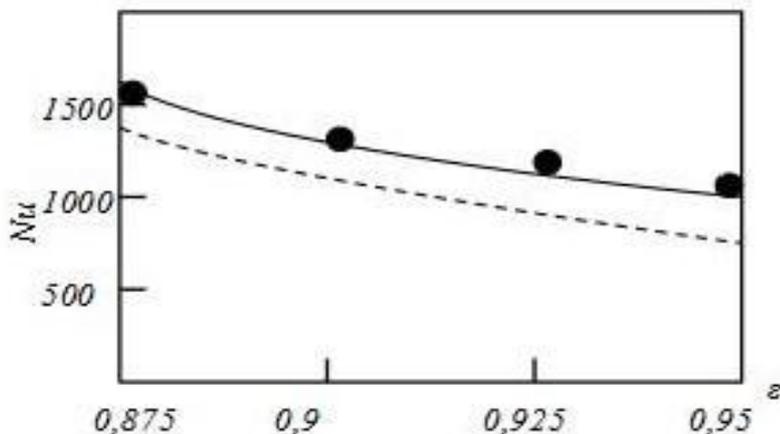


Рис. 11. Зависимость числа Нуссельта от пористости:
 — - пористость возрастает по высоте канала; --- - пористость постоянна по высоте канала;
 ● - результаты работы [7] для канала с переменной пористостью

Заключение. В результате выполнения работы получены значения полей температур и давлений в широком диапазоне изменения пористости канала не только по высоте, но и по длине. При этом чередование зон с различной пористостью по длине компактного теплообменника и по его высоте позволяет реализовать технические решения, обеспечивающие минимальное гидравлическое сопротивление при максимально возможных коэффициентах теплоотдачи.

Список обозначений:

- ε – пористость;
- t_0 – температура теплоносителя на входе, °С;
- d_p – диаметр пор, м.
- λ_f – коэффициент теплопроводности теплоносителя, Вт/(м·К);
- Pr – критерий Прандтля;
- ρ_f – плотность теплоносителя, кг/м³;
- u_0 – скорость теплоносителя, м/с;
- μ_f – коэффициент динамической вязкости, Па·с.
- c_{pf} – теплоемкость теплоносителя, Дж/(кг·К);
- d_r – гидравлический диаметр, м;
- Δp – перепад давления, Па;
- l – длина пути потока, м;
- H – высота пористого теплообменника, м.

СПИСОК ИСТОЧНИКОВ

1. The effects of variable porosity and cell size on the thermal performance of functionally-graded foams / M. Iasiello, N. Bianco, W. K. S. Chiu, V. Naso // International Journal of Thermal Sciences. – 2021. – V. 160. – P. 106696. – DOI 10.1016/j.ijthermalsci.2020.106696.
2. Modeling of heat transfer in an element with anisotropic porosity / N. N. Kozhukhov, E. A. Kozhukhova, D. A. Kononov [et al.] // Journal of Physics: Conference Series: XIII School-Seminar of Young Scientists and Specialists (XXIII SSSYSS 2021). – V. 2039. – Yekaterinburg: IOP Publishing Ltd, 2021. – P. 012011. – DOI 10.1088/1742-6596/2039/1/012011. – EDN DSHVHT.
3. Ryazhskikh, V. I. On the structure of the orthotropic 3D permeability tensor of an anisotropic porous body in heat and mass transfer problems / V. I. Ryazhskikh, A. V. Nikolenko, D. A. Kononov // Journal of Physics: Conference Series: 6th International Conference On Information Technology and Nanotechnology (ITNT 2020). – Samara, 2021. – P. 012082. – DOI 10.1088/1742-6596/1745/1/012082. – EDN AVBGGZ.
4. Numerical assessment of thermal characteristics of metal foams of orderly varied pore density and porosity under different convection regimes / G. Trilok, K. K. Kumar, N. Gnanasekaran, M. Mobedi // International Journal of Thermal Sciences. – 2022. – V. 172. – P. 107288. – DOI 10.1016/j.ijthermalsci.2021.107288.
5. Beji, H. Influence of Thermal Dispersion on Natural-Convection Heat-Transfer in Porous Media / H. Beji, D. Gobin // Numerical Heat Transfer, Part A. – 1992. – V. 22. – P. 487-500.
6. Amiri, A. Analysis of Dispersion Effects and Non-Thermal Equilibrium, Non-Darcian, Variable Porosity Incompressible Flow Through Porous Media / A. Amiri, K. Vafai // International Journal of Heat and Mass Transfer. – 1994. – V. 37. – № 6. – P. 939-954
7. Разработка методов интенсификации теплообмена в микроканальных теплообменниках гибридных систем термостабилизации / Д. А. Коновалов, И. Н. Лазаренко, Н. Н. Кожухов, И. Г. Дроздов // Вестник ВГТУ. – 2016. – Т. 12. – № 3. – С. 21-30.

Поступила в редакцию 30.08.2024 г., рекомендована к печати 09.10.2024 г.

SINGLE-PHASE CONVECTIVE HEAT TRANSFER IN A COMPACT HEAT EXCHANGER FILLED WITH ANISOTROPIC POROUS MEDIUM

Malevanyi M.V., Konovalov D.A., Kozhukhov N.N., CHernichenko V.V.

The work proposes an original design of a porous compact heat exchange element, consisting of layers of different porosity varying both along the height of the heat exchanger and along its length. Numerical modeling of various designs of heat exchange elements was carried out, and the influence of the arrangement of porous layers on the efficiency of intensification was shown. Comparison with known results in limiting cases allowed us to draw a conclusion about the correctness of the application of the numerical model.

Keywords: porous medium, developed heat exchange surface, highly porous materials, intensification, anisotropy.

Малеваный Михаил Владимирович

аспирант кафедры теоретической и промышленной теплоэнергетики ФГБОУ ВО «Воронежский государственный технический университет»,
Российская Федерация, г. Воронеж.
E-mail: 89304110859@mail.ru

Malevanyi Mikhail Vladimirovich

Postgraduate Student of the Department of Theoretical and Industrial Thermal Power Engineering of the Voronezh State Technical University,
Russian Federation, Voronezh.

Коновалов Дмитрий Альбертович

доктор технических наук, доцент, профессор кафедры теоретических основ теплотехники им. М.П. Вукаловича ФГБОУ ВО «Национальный исследовательский университет "Московский Энергетический Институт"»,
Российская Федерация, г. Москва,
профессор кафедры теоретической и промышленной теплоэнергетики ФГБОУ ВО «Воронежский государственный технический университет»,
Российская Федерация, г. Воронеж.
E-mail: dmikonovalov@yandex.ru

Konovalov Dmitrii Albertovich

Doctor of Technical Sciences, Associate Professor, Professor of the Department of Theoretical Foundations of Thermal Engineering named after M.P. Vukalovich of the National Research University «Moscow Power Engineering Institute»,
Russian Federation, Moscow,
Professor of the Department of Theoretical and Industrial Thermal Power Engineering of the Voronezh State Technical University,
Russian Federation, Voronezh.

Кожухов Николай Николаевич

кандидат технических наук, доцент, доцент кафедры теоретической и промышленной теплоэнергетики ФГБОУ ВО «Воронежский государственный технический университет»,
Российская Федерация, г. Воронеж.
E-mail: k0zhukhov@yandex.ru

Kozhukhov Nikolai Nikolaevich

Candidate of Technical Sciences, Associate Professor, Associate Professor of the Department of Theoretical and Industrial Thermal Power Engineering of the Voronezh State Technical University,
Russian Federation, Voronezh.

Черниченко Владимир Викторович

кандидат технических наук, доцент кафедры теоретической и промышленной теплоэнергетики ФГБОУ ВО «Воронежский государственный технический университет»,
Российская Федерация, г. Воронеж.
E-mail: vlad1427@yandex.ru

CHernichenko Vladimir Viktorovich

Candidate of Technical Sciences, Associate Professor of the Department of Theoretical and Industrial Thermal Power Engineering, Voronezh State Technical University,
Russian Federation, Voronezh.

МИКРОСКОПИЧЕСКОЕ ИЗУЧЕНИЕ ФОРМ МЕЛКОДИСПЕРСНЫХ ЧАСТИЦ ВЕЩЕСТВ, ПРИМЕНЯЕМЫХ В ТЕХНИЧЕСКИХ АГРЕГАТАХ

© 2024 Батухтин А.Г., Рудой В.И., Савиных Н.К., Кузютин К.А.

В данной статье проведено визуальное исследование форм мелкодисперсных частиц угля и магнетитового концентрата под микроскопическим увеличением. Данное исследование позволяет уточнить различные вопросы аэродинамического моделирования в исследованиях технических агрегатов, где могут в едином потоке находиться как частицы угля, так и частицы магнетитового концентрата, например сушильных печей. Представлены результаты микроскопических наблюдений измельченных частиц прошедших рассейвания через лабораторные сита размерами 71 мкм, 90 мкм, 125 мкм, 140 мкм, 200 мкм и 1000 мкм.

Ключевые слова: уголь, магнетитовый концентрат, аэродинамика частиц, моделирование процессов, микроскопическое исследование, тепломассообмен.

Введение. Современные исследования в области энергетики довольно часто вызывают необходимость моделирования различных процессов, происходящих внутри тех или иных агрегатов [1-3]. К таковым исследованиям можно отнести исследования процессов горения, движения потоков, тепломассообмена, гидравлического и аэродинамического сопротивления и т.п. Точность и верифицируемость модели по сравнению с реальным процессом во многом зависит от точности исходных граничных условий модели. Наибольшие трудности в моделировании реальных процессов вызывают процессы фазового перехода, химического превращения, аэродинамика и тепломассообмен в многокомпонентных разнородных смесях [4]. Встречаются же такие процессы во многих технических агрегатах, в частности в котельном оборудовании как в отопительном и энергетическом, так и в промышленном и ином, где происходят процессы горения топлива и последующее движение дымовых газов.

Примером такого устройства может являться барабанная сушильная печь [5]. Данные устройства применяются в горнодобывающей промышленности и включают в себя теплогенератор, сушильный барабан и вспомогательное оборудование. В частности данные устройства могут применяться при сушке магнетитового концентрата. Процесс сушки заключается в омывании горячими газами частиц магнетита во вращающемся барабане и последующем отделении высушенного вещества от дымовых газов. Процесс разделения компонентов происходит не полностью, и часть осушаемого вещества уносится вместе с потоком дымовых газов, где кроме них могут находиться недогоревшие частицы топлива, в случае твердотопливного питания теплогенератора. Частицы магнетита в последствии удаляются в очистных устройствах, таких как циклоны и электрофилтры. Однако моделирование процесса от этапа разделения основной части высушенного материала и до этапа удаления оставшегося в дымовых газах представляет довольно сложную задачу.

Работа выполнена в рамках государственного задания Министерства науки и высшего образования Российской Федерации (тема № 123102000012-2 Комплексное исследование аэродинамических характеристик плазменных систем термохимической подготовки топлива, соглашение № 075-03-2023-028/1 от 05.10.2023 г.)

Основная часть. Основная сложность моделирования заключается в правильной и точной оценке процессов тепломассопереноса многокомпонентной смеси, имеющей в себе как газовую фазу в виде дымовых газов, так и частицы осушенного материала и недогоревшие частицы угля. Процессы обмена теплотой между различными частицами в газовой смеси во многом зависят от площади поверхностей данных частиц.

Однако, моделирование является наиболее удобным способом исследования сложных аэродинамических и тепломассообменных процессов, поэтому решение такой задачи представляется необходимым шагом.

Для моделирования и исследования аэродинамических и тепломассообменных свойств, прежде всего, были исследованы формы частиц, находящихся в потоке газов, а именно частицы угля и частицы магнетита. Наиболее часто применяется для описания дисперсных материалов, имеющих неправильную форму, классификация частиц на три группы по трём измерениям [6]:

- 1) изометрические частицы, у которых все три размера примерно одинаковы.
- 2) пластины – частицы, у которых два размера преобладают над третьим.
- 3) волокна – частицы, у которых преобладает один размер.

Размеры частицы чаще всего определяются с помощью её диаметра, но в идеале это применимо только для частиц правильной формы, если геометрия частицы неправильная, то её диаметр находится как среднее арифметическое всех трёх измерений [7]. Однако высоту частицы определить на фото микроскопа довольно трудно, поэтому диаметр оценивают как среднее арифметическое длины и ширины частицы.

Для обеспечения репрезентативности данных исследования применялся метод квартования проб [8]. При данном способе берут сыпучий материал из разных мест и смешивают, далее рассыпают ровным слоем и делят на 4 квадрата. После чего противоположные квадраты по диагонали удаляют, оставляя два квадрата, снова смешивают, делят на 4 части и т. д. Таким образом, количество материала уменьшают, пока его не станет столько, сколько необходимо для проведения микроскопического анализа. Квартование в нашем случае применялось как для отбора проб, так и для отбора рассматриваемых частиц для каждого сито отдельно.

Для исследования были взяты проба энергетического угля и магнетитового концентрата. Перед тем, как приступить к микроскопическому анализу, уголь прошёл размол в барабанно-шаровой мельнице и был просеян на ситах с размером ячеек 1000 мкм, 200 мкм, 140 мкм, 125 мкм, 90 мкм, 71 мкм как одни из наиболее характерных размеров из стандартных сит [9] применяемых в энергетической отрасли. Магнетит также прошёл аналогичный просев. С помощью цифрового микроскопа Inskan 315W была проведена оценка формы частиц в соответствующих диапазонах отсева частиц угля и магнетитового концентрата (соответственно 71 мкм и меньше, от 71 мкм до 90 мкм, от 90 мкм до 125 мкм, от 125 мкм до 140 мкм, от 140 мкм до 200 мкм, от 200 мкм до 1000 мкм).

Для каждого интервала размера частиц угля наиболее встречающаяся форма – изометрическая. Фото всех интервалов размеров частиц представлены на рисунках 1-6 и выполнены на фоне сито, которое является нижней границей размера для каждого отдельного случая.

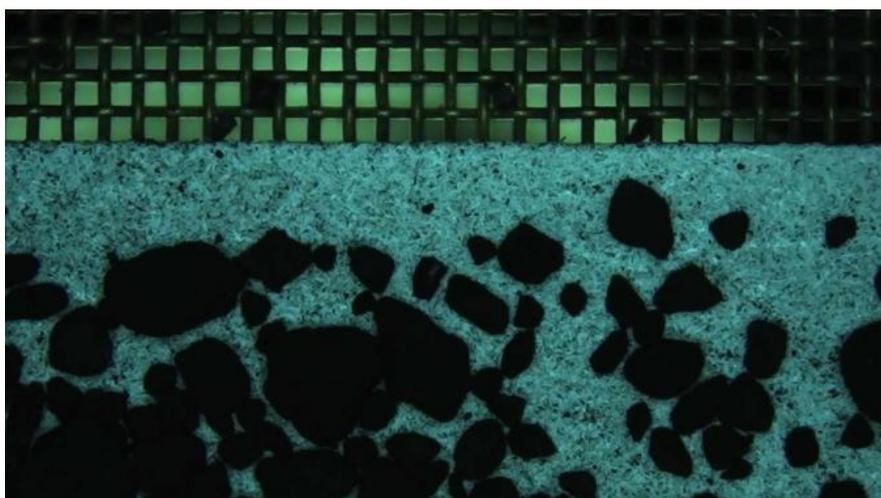


Рис. 1. Частицы в интервале 200 мкм - 1000 мкм

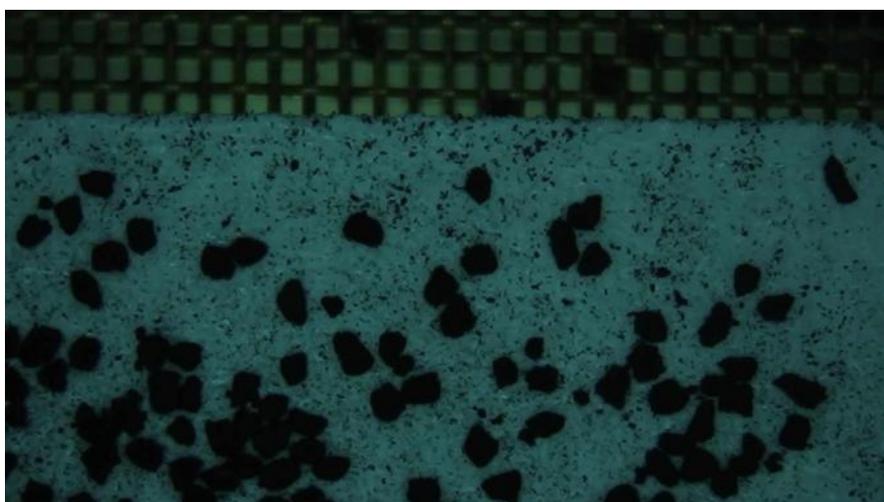


Рис. 2. Частицы в интервале 140 мкм – 200 мкм

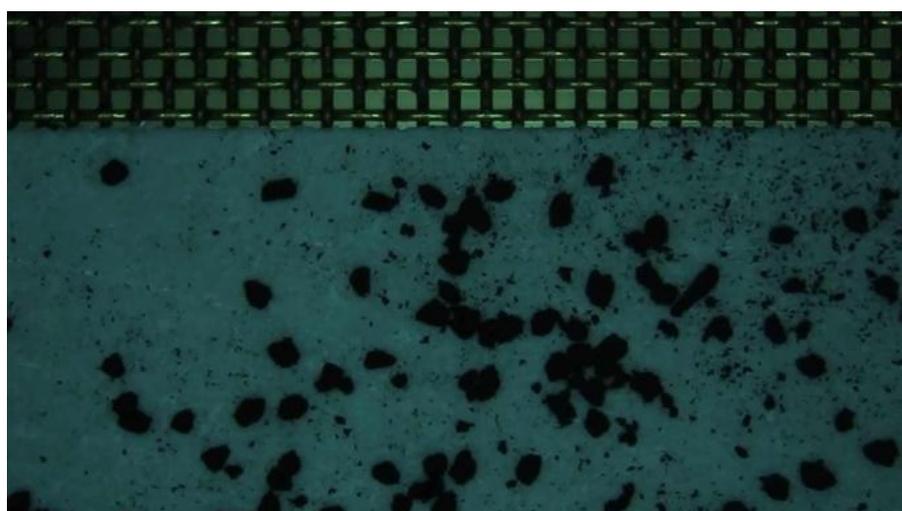


Рис. 3. Частицы в интервале 125 мкм – 140 мкм

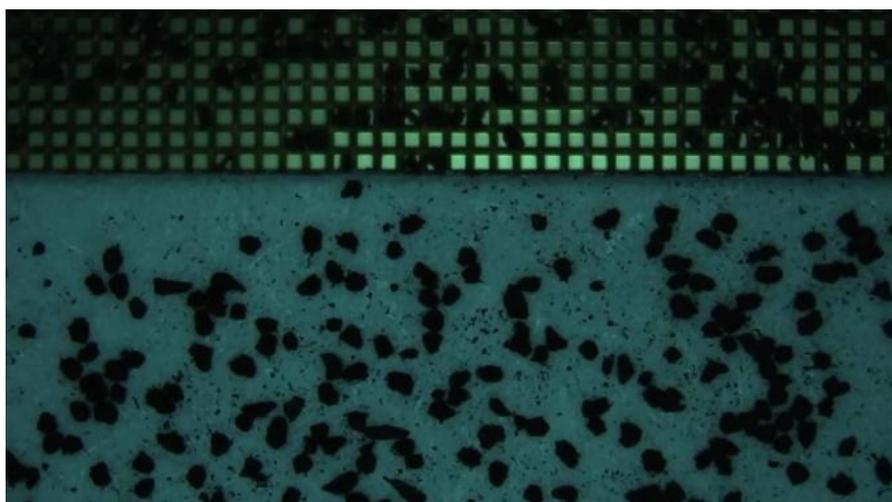


Рис. 4. Частицы в интервале 90 мкм – 125 мкм

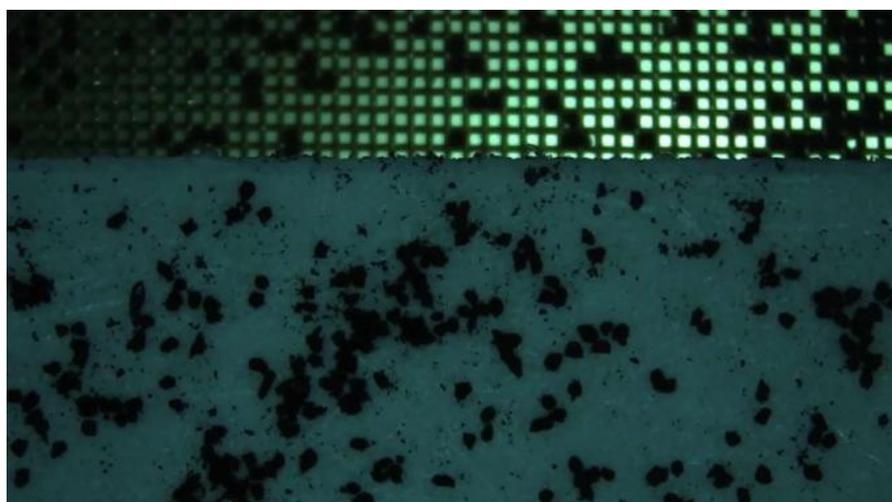


Рис. 5. Частицы в интервале 71 мкм – 90 мкм

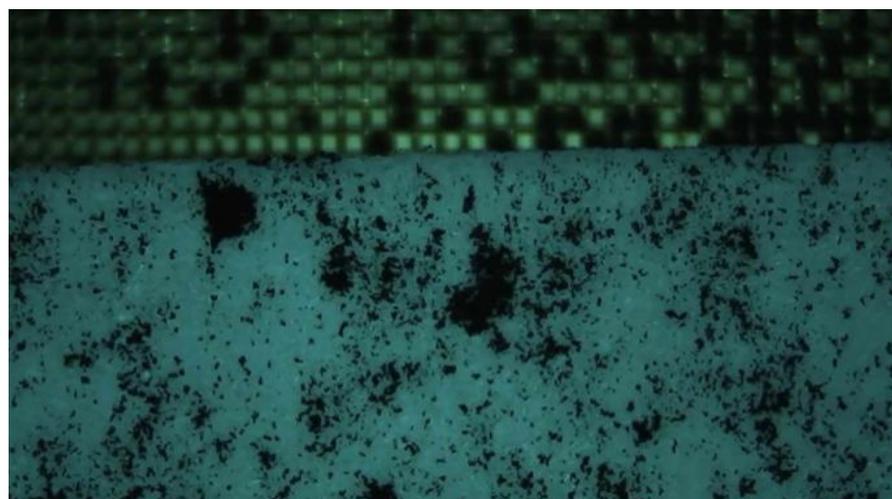


Рис. 6. Частицы меньше 71 мкм

В случае магнетита, частицы прошли все уровни сита, поскольку их размер после помола не превышает 60 мкм [10]. А наибольшая тенденция формы частиц магнетита – пластины. Фото частиц магнетита представлено на рисунке 7 на фоне сита размером 71 мкм.

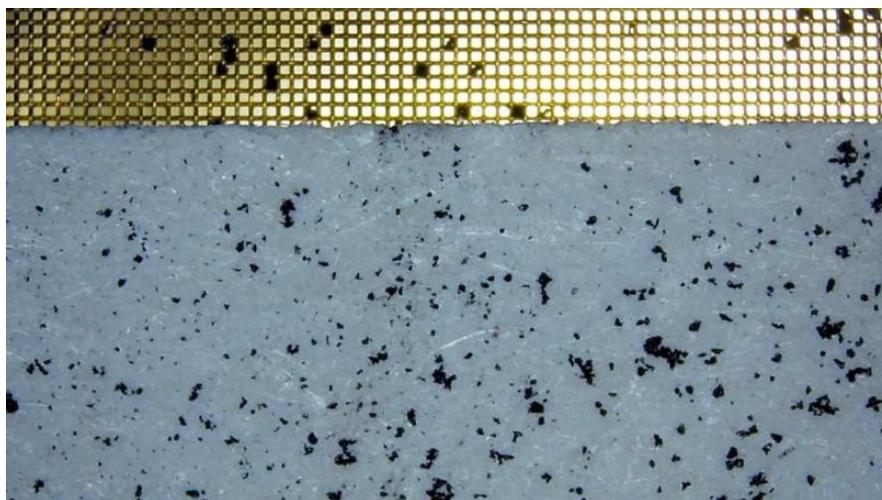


Рис. 7. Частицы меньше 71 мкм

Выводы. Полученные данные могут быть использованы в последующем моделировании запыленных газовых потоков, включающем в себя компоненты магнетита и угля. Задание в модели формы и размера одиночной частицы потока в полном соответствии с реальным процессом позволяет значительно повысить точность данных, получаемых с модели.

СПИСОК ИСТОЧНИКОВ

1. Корченко, М. Д. Способы применения нейросетей в энергетике / М. Д. Корченко // Вестник науки. – 2024. – Т. 1, № 6(75). – С. 1444-1448.
2. К вопросу о прогнозировании амплитуд пульсации давления в системах трубопроводов / А. В. Николаева, А. П. Скибин, А. В. Чернышев [и др.] // Вестник Московского государственного технического университета им. Н.Э. Баумана. Серия Машиностроение. – 2012. – № 3(88). – С. 3-16.
3. Изучение способов построения профиля скорости потока в трубе малого диаметра / А. Г. Батухтин, М. С. Басс, М. В. Кобылкин [и др.] // Международный журнал гуманитарных и естественных наук. – 2023. – № 11-4(86). – С. 7-13. – DOI 10.24412/2500-1000-2023-11-4-7-13.
4. Корценштейн, Н. М. Объемная конденсация в запыленном парогазовом потоке с учетом распределения пылинок по размерам / Н. М. Корценштейн, А. К. Ястребов // Коллоидный журнал. – 2016. – Т. 78, № 4. – С. 444. – DOI 10.7868/S0023291216040108.
5. Чернобыльский, И. И. Сушильные установки химической промышленности / И. И. Чернобыльский, Ю.М. Тананайко. – Киев: Техника, 1969. – 280 с.
6. Коузов, П. А. Основы анализа дисперсного состава промышленных пылей и измельченных материалов / П. А. Коузов. – Л.: Химия, 1987. – 264 с.
7. Кларк, Э. Р. Микроскопические методы исследования материалов / Э. Р. Кларк, К. Н. Эберхардт. – М.: Техносфера, 2007. – 376 с.
8. Васильев, В. П. Аналитическая химия. В 2 ч. Ч. 1. Гравиметрический и титриметрический методы анализа: учеб. для химико-технол. спец. вузов / В. П. Васильев. – М.: Высш. шк., 1989. – 320 с.
9. Артеменко, А. И. Справочное руководство по химии: справ. пособие / А. И. Артеменко, В. А. Малеванный, И. В. Тикунова. – М.: Высш. шк., 1990. – 303 с.
10. Матюхин, П. В. Основные физико-механические характеристики магнетита, подвергнутого воздействию высоких давлений прессования / П. В. Матюхин, Р. Н. Ястребинский, А. В. Широков // Вестник Белгородского государственного технологического университета им. В.Г. Шухова. – 2016. – № 9. – С. 189-195.

Поступила в редакцию 03.09.2024 г., рекомендована к печати 25.10.2024 г.

MICROSCOPIC STUDY OF THE SHAPES OF FINE PARTICLES OF SUBSTANCES USED IN TECHNICAL AGGREGATES

This article provides a visual study of the shapes of fine particles of coal and magnetite concentrate under microscopic magnification. This study makes it possible to clarify various issues of aerodynamic modeling in the research of technical aggregates, where both coal particles and particles of magnetite concentrate can be in a single stream, for example, drying furnaces. The results of microscopic observations of crushed particles dispersed through laboratory sieves measuring 71 microns, 90 microns, 125 microns, 140 microns, 200 microns and 1000 microns are presented.

Keywords: coal, magnetite concentrate, particle aerodynamics, process modeling, microscopic examination, heat and mass transfer.

Batukhtin A.G., Rudoï V.I., Savinykh N.K., Kuziutin K.A.

Батухтин Андрей Геннадьевич

доктор технических наук, доцент кафедры энергетики ФГБОУ ВО «Забайкальский государственный университет», Российская Федерация, Забайкальский край, г. Чита.
Email: batuhkina_ir@mail.ru

Batukhtin Andrei Gennadevich

Doctor of Technical Sciences, Associate Professor at Department of Power Engineering of Transbaikal State University,
Russian Federation, Transbaikal Territory, Chita.

Рудой Валерий Игоревич

аспирант, ассистент кафедры энергетики ФГБОУ ВО «Забайкальский государственный университет», Российская Федерация, Забайкальский край, г. Чита.
Email: valera_rud_99@mail.ru

Rudoï Valerii Igorevich

Postgraduate Student, Assistant at Department of Power Engineering of Transbaikal State University,
Russian Federation, Transbaikal Territory, Chita.

Савиных Надежда Константиновна

магистрант кафедры энергетики ФГБОУ ВО «Забайкальский государственный университет», Российская Федерация, Забайкальский край, г. Чита.
Email: sivaeva21@yandex.ru

Savinykh Nadezhda Konstantinovna

Master Student at Department of Power Engineering of Transbaikal State University,
Russian Federation, Transbaikal Territory, Chita.

Кузютин Константин Андреевич

магистрант кафедры энергетики ФГБОУ ВО «Забайкальский государственный университет», Российская Федерация, Забайкальский край, г. Чита.
Email: kuzutinkostia@mail.ru

Kuziutin Konstantin Andreevich

Master Student at Department of Power Engineering of Transbaikal State University,
Russian Federation, Transbaikal Territory, Chita.

МЕТОДИКА АНАЛИЗА АВАРИЙНЫХ РЕЖИМОВ РАБОТЫ КОНТАКТНЫХ СОЕДИНЕНИЙ ЭЛЕКТРОСЕТИ 0,4/0,22 КВ

© 2024 Бершадский И.А., Мых А.Д., Згарбул А.В., Якимишина В.В.

В данной статье анализируется статистический уровень пожарной опасности в жилых помещениях. Оценивается его влияние на надежность электросистемы квартир и обозначается роль устройств искрозащиты в этом процессе. Кроме того, показана и описана методика определения вероятности появления источника возгорания в электропроводке с учетом различных повреждающих факторов, интенсивности их появления и устранения.

Ключевые слова: риск, пожарная безопасность, система электроснабжения, узел нагрузки, электрифицированное помещение, средства защиты.

Введение. Согласно статистических данных, имеющихся в отделе пожарной статистики ВНИИПО (Всероссийского научно-исследовательского института противопожарной обороны) МЧС России, в период с 2018 по 2022 годы в РФ, несмотря на общее снижение количества пожаров (рис. 1), наблюдается увеличение количества пожаров связанных с нарушением правил устройства и эксплуатации электрооборудования (рис. 2) [1].

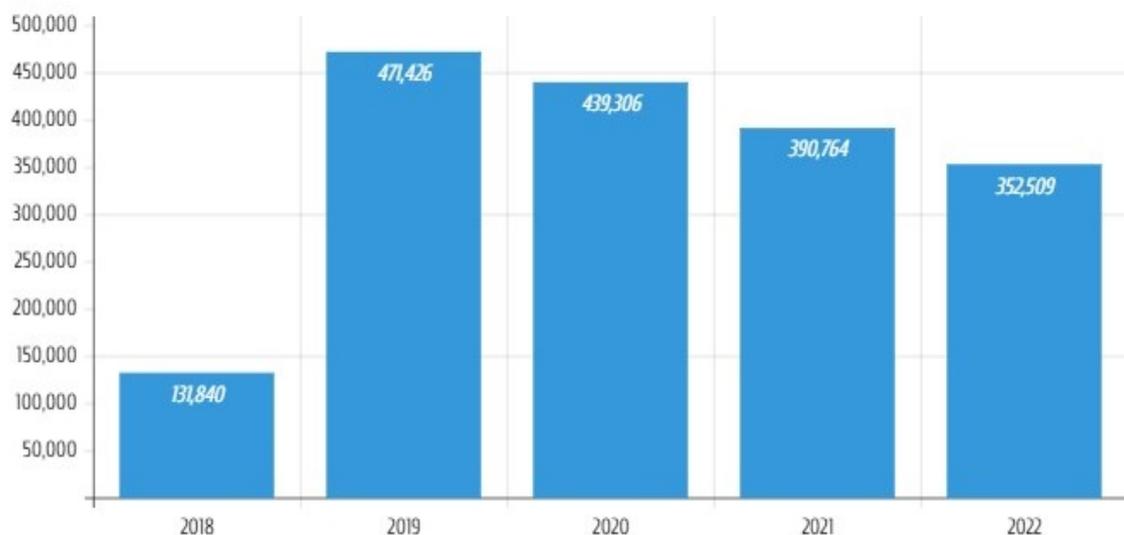
Количество пожаров за 2018-2022 гг.

Рис. 1. Сравнительный анализ количества пожаров за 2018-2022 гг.

Существенной проблемой обеспечения пожарной безопасности в низковольтных электроустановках жилых и общественных зданий являются отказы, связанные с раскаленными соединениями и искрениями, которые обычно возникают из-за перегрева некачественных электрических соединений, характеризующихся высоким электрическим сопротивлением [2].

Количество пожаров за 2018–2022 гг.

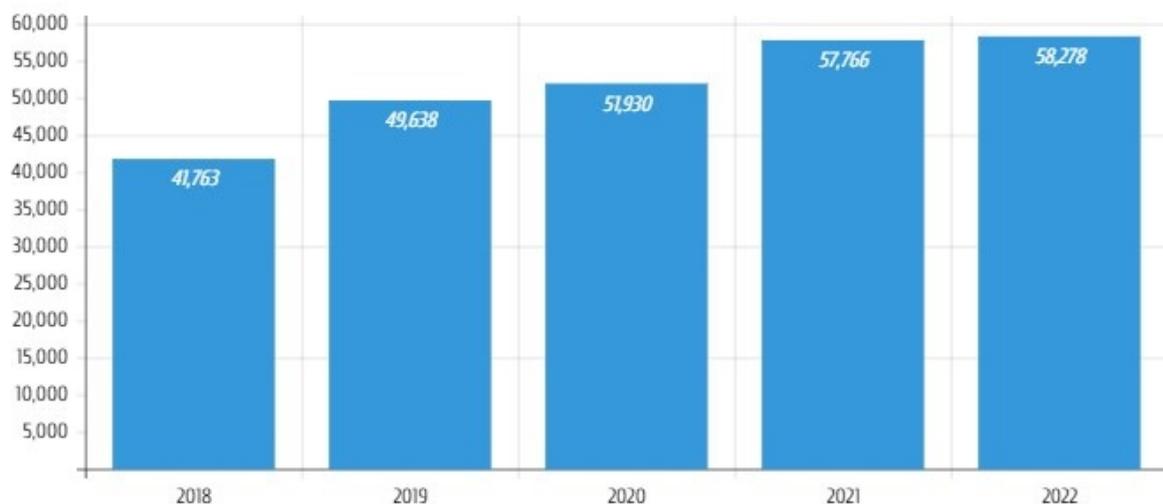


Рис. 2. Сравнительный анализ количества пожаров по причинам, связанным с электрооборудованием за 2018–2022 гг.

Вот почему разработка защитных и профилактических методов, организация технических осмотров горячих, искрящихся контактов является актуальной проблемой в настоящее время.

Целью исследований является получение аналитической зависимости частоты появления пожаров в электрифицированных помещениях от частоты появления опасно нагретого силового контактного соединения (штепсельная электроколодка с включенной штепсельной электровилкой), частоты и длительности нахождения горючего материала на поверхности штепсельной электроколодки, надежности средств защиты сети 0,38/0,22 кВ.

Данную задачу разбиваем на две части. В первой части определим вероятность появления опасного источника воспламенения, вызванного нарушениями в контактных соединениях электропроводок электрифицированного помещения. Во второй части, используя параметры интенсивности появления искрения или нагрева контактного соединения по данным, полученным в первой части, оцениваем риск возникновения пожаров в этом помещении при эксплуатации его электропроводки [3].

Модель разрушения электрических контактов под воздействием окружающей среды. В настоящее время математическое моделирование этапов возникновения пожаров в электросети соответствует математическим детерминированным моделям [4, 5]. В соответствии с ГОСТом (Р 51901.15-2005: Менеджмент риска. Применение марковских методов [6]), можно оценить показатели надежности контактных соединений электросети квартиры.

Цепь Маркова представляет собой случайный процесс, который переходит из состояния в состояние в дискретные моменты времени, называемыми шагами и обозначаемые через $n = 0, 1, 2, \dots$, с определенной вероятностью, так называемой вероятностью перехода. Число состояний конечно, а значение переходной вероятности полностью определяется тем, в каком состоянии находится процесс, то есть она является условной.

Для оценки показателей надежности контактных соединений рассмотрим электросеть как совокупность контактов, каждый из которых может существовать только в одном из крайних состояний: неработоспособном или работоспособном. Электрическая сеть в целом может существовать в различных состояниях, каждое из которых определяется специфической комбинацией работоспособного и неработоспособного состояний ее элементов. В момент отказа или восстановления элемента электросеть переходит из одного состояния в следующее. Такая модель является моделью дискретных состояний с непрерывным временем.

На рис. 3 показан граф состояний показателей надежности контактных соединений электросети жилого многоквартирного дома.

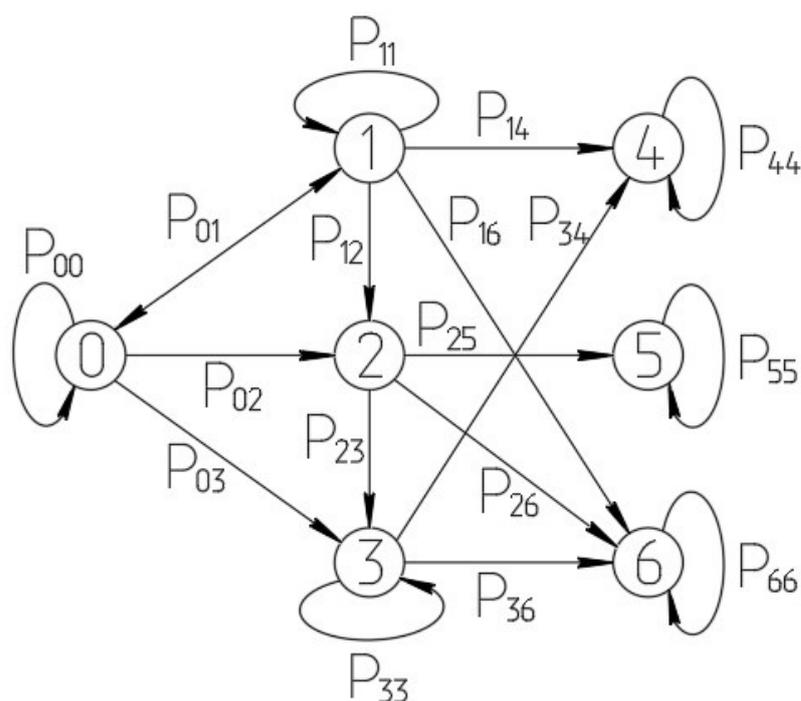


Рис. 3. Граф состояний надежности контактных соединений

Состояние 0 соответствует рабочему состоянию контактных соединений электросети жилого многоквартирного дома.

Состояние 1 – периодически возникаемые БПС (попадание воды, вибрация...).

Состояние 2 – случайное невосстанавливаемое воздействие, приводящее к разрушению контакта (окисление, физическое воздействие...).

Состояние 3 – постоянное невосстанавливаемое воздействие (скрутка проводов, заводской брак, некачественный монтаж...).

Состояние 4 – теплового воздействия на изоляцию от падения напряжения на контакте.

Состояние 5 – искрение, возникновение дуги и ее воздействие на легкогоряемые материалы.

Состояние 6 – выделение тепловой энергии в результате возникновения перегрузки или короткого замыкания.

Данной математической модели будет соответствовать следующая матрица интенсивности переходов [5]:

$$Px = \begin{pmatrix} P_{00} & P_{01} & P_{02} & P_{03} & 0 & 0 & 0 \\ P_{10} & P_{11} & P_{12} & 0 & P_{14} & 0 & P_{16} \\ 0 & 0 & 0 & 0 & 0 & P_{25} & P_{26} \\ 0 & 0 & P_{32} & P_{33} & P_{34} & 0 & P_{36} \\ 0 & 0 & 0 & 0 & P_{44} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & P_{55} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & P_{66} \end{pmatrix}$$

Состояния $S_0 - S_3$ принадлежат эргодическому множеству, в котором можно из любого состояния попасть в любое [5, 7-9]. В данной модели используются поглощающие цепи Маркова - цепи, в которых часть состояний являются поглощающими, то есть, попав в эти состояния, нельзя из них выйти. Поглощающими состояниями являются $S_4 - S_6$.

Известно, что в любой поглощающей конечной цепи Маркова, независимо от того, где начался процесс, вероятность после n шагов оказаться в невозвратном эргодическом состоянии стремится к 1 при $n \rightarrow \infty$ [4, 10, 11].

Удобно придать матрице Px канонический вид, объединив все эргодические состояния в одну группу и все невозвратные состояния – в другую группу. Тогда каноническая форма будет:

$$Pz = \begin{pmatrix} S & \vdots & 0 \\ \dots & \dots & \dots \\ R & \vdots & Q \end{pmatrix},$$

или:

$$Pz = \begin{pmatrix} P_{66} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & P_{55} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & P_{44} & 0 & 0 & 0 & 0 \\ P_{36} & 0 & P_{34} & P_{33} & 0 & 0 & 0 \\ P_{26} & P_{25} & 0 & P_{23} & 0 & 0 & 0 \\ P_{16} & 0 & P_{14} & 0 & P_{12} & P_{11} & P_{10} \\ 0 & 0 & 0 & P_{03} & P_{02} & P_{01} & P_{00} \end{pmatrix},$$

Подматрица S размерности 3×3 :

$$S = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

относится к процессу после достижения пожаром поглощающих состояний, подматрица Q размерностью 4×3 представляет собой совокупность нулей, матрица R размерности 3×4 имеет следующий вид:

$$R = \begin{bmatrix} P_{36} & 0 & P_{34} \\ P_{26} & P_{25} & 0 \\ P_{16} & 0 & P_{14} \\ 0 & 0 & 0 \end{bmatrix},$$

и показывает вероятность переходов из невозвратных в эргодические состояния.

Подматрица Q размерности 4×4 :

$$Q = \begin{bmatrix} P_{33} & 0 & 0 & 0 \\ P_{23} & 0 & 0 & 0 \\ 0 & P_{12} & P_{11} & P_{10} \\ P_{03} & P_{02} & P_{01} & P_{00} \end{bmatrix},$$

описывает поведение процесса до выхода из множества эргодических состояний [12, 13].

Матрица A получается из матрицы интенсивности переходов Px путем исключения единицы из главной диагонали:

$$A = Q - I.$$

Подматрица I представляет собой единичную матрицу. Из теории Марковских цепей следует, что среднее время, соответствующее аварийному режиму работы электропроводки в каждом конкретном состоянии, всегда конечно, и при этом средние значения времени определяются фундаментальной матрицей N . Каждый элемент матрицы соответствует среднему числу попадания системы в то или иное состояние до остановки процесса (поглощения). Матрица имеет вид:

$$N = (I - Q)^{-1}.$$

Среднее время до перехода электропроводки в одно из поглощающих состояний, если в начальный момент она находилась в одном из непоглощающих состояний будет определяться выражением [14]:

$$\tau = N\xi,$$

где ξ - вектор-столбец с единичными элементами:

$$\xi_{\langle 4 \rangle} = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}.$$

Марковская модель, включающая несколько поглощающих состояний, позволят рассчитать, в какое из поглощающих состояний цепь попадет раньше (или позже); в каких из них процесс будет останавливаться чаще, а в каких – реже. Введем величину B , которая будет характеризовать величину вероятности того, что процесс завершится в некотором поглощающем состоянии при условии, что начальным было состояние Pz . При этом множество состояний образует матрицу, строки которой соответствуют эргодическим состояниям, а столбцы - всем поглощающим состояниям.

$$B = NR.$$

Вероятность появления одного из опасных поглощающих состояний 6, 5, 4 определяется из выражения:

$$Q(t) = 1 - P_0(t) - P_1(t) - P_2(t) - P_3(t),$$

где $P_0(t)$, $P_1(t)$, $P_2(t)$, $P_3(t)$ находят из решения системы линейных дифференциальных уравнений вида:

$$P'(t) = A^T \cdot P(t). \quad (1)$$

Система дифференциальных уравнений (1) решается при начальных условиях: $P_3(0)=0; P_2(0) = 0; P_1(0) = 0; P_0(0) = 1$, где $P'(t)$ – вектор-столбец [$P'_3(t), P'_2(t), P'_1(t), P'_0(t)$]; $P(t)$ – вектор-столбец [$P_3(t), P_2(t), P_1(t), P_0(t)$].

Пример. Для электрифицированного помещения, в котором эксплуатируется сеть 0,38/0,22 кВ заданы следующие величины интенсивности переходов в смежные состояния (см. рис.3):

$$\begin{array}{cccc} \lambda_{0,1} = 0,1 \text{ год}^{-1} & \lambda_{1,2} = 0,2 \text{ год}^{-1} & \lambda_{2,3} = 0,2 \text{ год}^{-1} & \lambda_{3,4} = 0,4 \text{ год}^{-1} \\ \lambda_{0,2} = 0,1 \text{ год}^{-1} & \lambda_{1,4} = 0,2 \text{ год}^{-1} & \lambda_{2,5} = 0,4 \text{ год}^{-1} & \lambda_{3,6} = 0,4 \text{ год}^{-1} \\ \lambda_{0,3} = 0,1 \text{ год}^{-1} & \lambda_{1,6} = 0,2 \text{ год}^{-1} & \lambda_{2,6} = 0,4 \text{ год}^{-1} & \end{array}$$

и:
$$\mu_{1,0} = 0,2 \text{ год}^{-1}$$

При этом параметр λ характеризует интенсивность или скорость, с которой пожаробезопасное состояние силового контактного соединения перейдет в пожароопасное; μ – интенсивность или скорость, с которой пожароопасное состояние силового контактного соединения переходит в пожаробезопасное.

Необходимо оценить вероятность перехода системы в одно из отказавших состояний при эксплуатации помещения в течение года, а также оценить время до наступления отказавшего состояния.

Решение: Матрица интенсивности переходов Px согласно рис. 3 будет иметь вид:

$$Px = \begin{pmatrix} 0,7 & 0,1 & 0,1 & 0,1 & 0 & 0 & 0 \\ 0,2 & 0,2 & 0,2 & 0 & 0,2 & 0 & 0,2 \\ 0 & 0 & 0 & 0,2 & 0 & 0,4 & 0,4 \\ 0 & 0 & 0 & 0,2 & 0,4 & 0 & 0,4 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

Матрица интенсивности переходов канонического вида Pz :

$$Pz = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0,4 & 0 & 0,4 & 0,2 & 0 & 0 & 0 \\ 0,4 & 0,4 & 0 & 0,2 & 0 & 0 & 0 \\ 0,2 & 0 & 0,2 & 0 & 0,2 & 0,2 & 0,2 \\ 0 & 0 & 0 & 0,1 & 0,1 & 0,1 & 0,7 \end{pmatrix}.$$

Матрица R размерности 4x3:

$$R = \begin{pmatrix} 0,4 & 0 & 0,4 \\ 0,4 & 0,4 & 0 \\ 0,2 & 0 & 0,2 \\ 0 & 0 & 0 \end{pmatrix}.$$

Матрица Q размерности 4x4:

$$Q = \begin{pmatrix} 0,2 & 0 & 0 & 0 \\ 0,2 & 0 & 0 & 0 \\ 0 & 0,2 & 0,2 & 0,2 \\ 0,1 & 0,1 & 0,1 & 0,7 \end{pmatrix}.$$

Матрица A представляет собой разность между матрицей Q и единичной:

$$A = \begin{pmatrix} -0,8 & 0 & 0 & 0 \\ 0,2 & -1 & 0 & 0 \\ 0 & 0,2 & -0,8 & 0,2 \\ 0,1 & 0,1 & 0,1 & -0,3 \end{pmatrix}.$$

Определим среднее число попадания системы в одно из эргодических состояний до поглощения:

$$N = \begin{pmatrix} 1,25 & 0 & 0 & 0 \\ 0,25 & 1 & 0 & 0 \\ 0,205 & 0,364 & 1,364 & 0,909 \\ 0,568 & 0,455 & 0,455 & 3,636 \end{pmatrix}.$$

Находим среднее время нахождения электропроводки в одном из эргодических (непоглощающих) состояний 3, 2, 1, 0 до перехода в то или иное состояние до остановки процесса (поглощения):

$$\tau = \begin{pmatrix} 1,25 \\ 1,25 \\ 2,841 \\ 5,114 \end{pmatrix}, \text{ год.}$$

Определим величину B , которая будет характеризовать вероятность завершения процесса в некотором поглощающем состоянии 6, 5, 4 при условии, что начальным было одно из состояний 3, 2, 1, 0.

$$B = \begin{matrix} & \underbrace{\quad 6 \quad 5 \quad 4} & \\ \begin{pmatrix} 0,5 & 0 & 0,5 \\ 0,5 & 0,4 & 0,1 \\ 0,5 & 0,145 & 0,355 \\ 0,5 & 0,182 & 0,318 \end{pmatrix} & \begin{matrix} \left. \begin{matrix} 3 \\ 2 \\ 1 \\ 0 \end{matrix} \right\} \end{matrix} \end{matrix}$$

При этом множество состояний образует матрицу, строки которой соответствуют эргодическим состояниям, а столбцы - всем поглощающим состояниям.

В начальный момент времени $t = 0$ вероятности нахождения системы в исходных состояниях 3, 2, 1, 0 будет иметь значения:

$$P_3(0) = 0$$

$$P_2(0) = 0$$

$$P_1(0) = 0$$

$$P_0(0) = 1$$

Решив систему дифференциальных уравнений (1), находим как меняется вероятность пребывания системы в работоспособном состоянии в течение заданного времени (рис. 4 и 5).

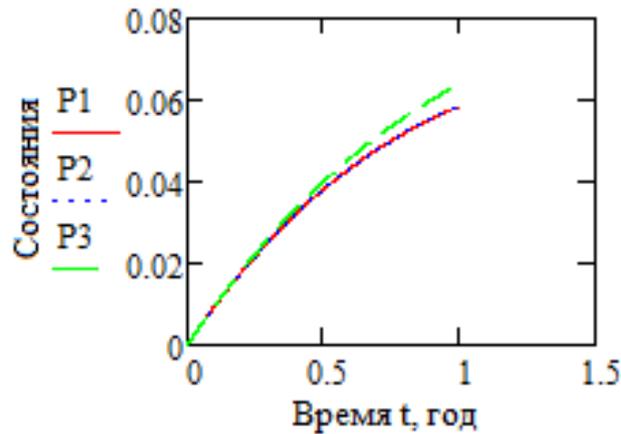


Рис. 4. Вероятности нахождения системы в состояниях 1, 2, 3 за время $t = 1$ год

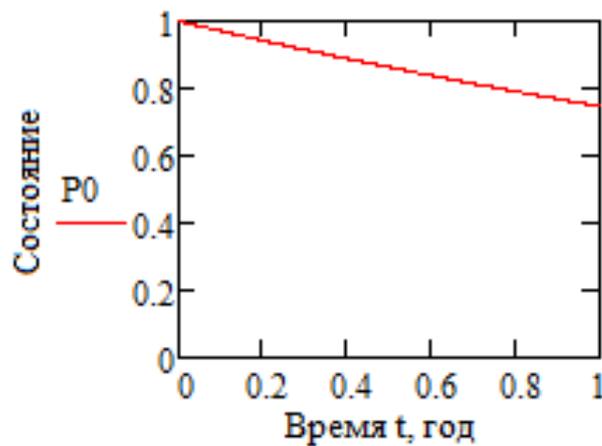


Рис. 5. Вероятности нахождения системы в состоянии 0 за время $t = 1$ год

Определим вероятности нахождения электропроводки в различных состояниях за время $t = 1$ год.

$$P_0(1) = 0,747$$

$$P_1(1) = 0,058$$

$$P_2(1) = 0,058$$

$$P_3(1) = 0,064$$

Вероятность возникновения источника пожара в электропроводке определим как вероятность группы несовместных событий $\overline{4, 5, 6}$ за $t = 1$ год:

$$\tilde{P}_{456}(1) = 1 - P_0(1) - P_1(1) - P_2(1) - P_3(1) = 0,072$$

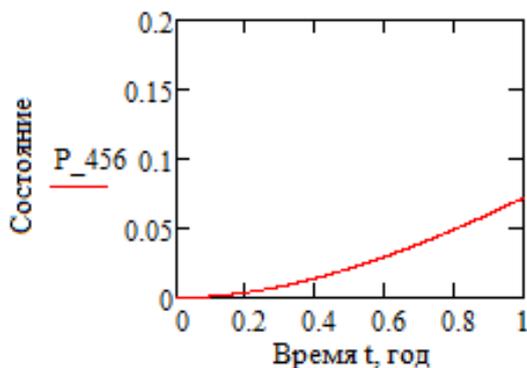


Рис. 6. Вероятность нахождения системы в состоянии отказов (\tilde{P}_{456}) за время $t = 1$ год

Найдём вероятности переходов из состояния 0 в состояния отказов 4, 5 и 6:

$$P_4(1) = B_{3,2} \cdot P_{456}(1) = 0,023$$

$$P_5(1) = B_{3,1} \cdot P_{456}(1) = 0,013$$

$$P_6(1) = B_{3,0} \cdot P_{456}(1) = 0,036$$

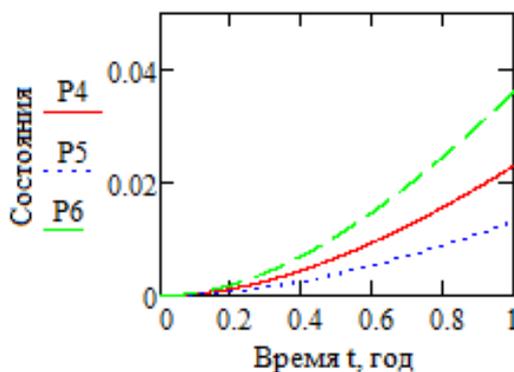


Рис. 7. Вероятность нахождения системы в состояниях 4, 5, 6 за время $t = 1$ год

Для однородного марковского процесса и времени наблюдения 1 год можно принять допущение, что интенсивности отказов 4, 5 и 6:

$$\lambda_4 = P_4(1) = 0,023 \text{ год}^{-1}$$

$$\lambda_5 = P_5(1) = 0,013 \text{ год}^{-1}$$

$$\lambda_6 = P_6(1) = 0,036 \text{ год}^{-1}$$

Выводы. На основе однородных марковских процессов с дискретным числом состояний и непрерывным временем была предложена математическая вероятностная модель оценки трёх аварийных режимов электропроводки квартир:

- тепловое воздействие на изоляцию от падения напряжения на контакте;
- искрение, возникновение дуги и ее воздействие на легкоосгораемые материалы;
- выделение тепловой энергии в результате возникновения перегрузки или короткого замыкания.

На одном примере показано, каким образом можно определить интенсивности переходов в поглощающие пожароопасные состояния электропроводки квартиры. Эти данные могут использоваться для расчета риска возникновения пожаров в электрифицированном помещении при эксплуатации его электропроводки

СПИСОК ИСТОЧНИКОВ

1. Пожары и пожарная безопасность в 2022 г.: статистика пожаров и их последствий. – ФГБУ ВНИИПО МЧС России, г. Балашиха, Московская обл., 2023. – 81с.
2. Бершадский, И. А. Обоснование способа защиты электрической сети напряжением 220 В от перегрузок с учетом нестационарного нагрева электропроводки / И. А. Бершадский, А. В. Згарбул, А. А. Кулиш // Вестник Донецкого национального технического университета. – 2019. – 2(16). – С. 71-79.
3. Оценка рисков возникновения пожаров при эксплуатации электрифицированных помещений / А. П. Ковалев, И. А. Бершадский, А. Д. Мых, А. В. Згарбул // Научный вестник НИИГД Респиратор. – 2024. – 1(61). – С. 29-37.
4. Моторыгин, Ю. Д. Математическое моделирование процессов возникновения и развития пожаров: монография / Ю. Д. Моторыгин. – СПб.: Санкт-Петербургский университет Государственной Противопожарной Службы МЧС России, 2011. – 345 с.
5. Ramachandran, G. Probabilistic Approach to Fire Risk Evaluation / G. Ramachandran // Fire Technology. – 1988. – V. 24(3). – P. 204-226.
6. ГОСТ Р 51901.15-2005 (МЭК 61165:1995). Менеджмент риска. Применение марковских методов: национальный стандарт Российской Федерации: введен 2006-02-01. – М.: Стандартинформ, 2005 – 15 с.
7. Мирясов, Е. Ю. Моделирование развития пожара на специальных хранилищах вооружения / Е. Ю. Мирясов, Ю. Д. Моторыгин, Д. В. Косенко // Материалы Первого международного научно-практического семинара «Системы комплексной безопасности и физической защиты». – СПб: СПб государственной политехнической университет, 2013.
8. Разливанов, И. Н. Математическое моделирование процессов развития пожара и пожаротушения в условиях ограниченности сил и средств: дис. ... канд. техн. наук : 05.13.18 / Разливанов Игорь Николаевич ; СПб УГПС МЧС Росси. – Санкт-Петербург, 2009. – 185 с.
9. Моделирование пожаров и взрывов: монография / И. Ф. Астахова, В. П. Беляцкий, Н. Н. Брушлинский [и др.]. – М.: Ассоц. "Пожнаука", 2000. – 498 с.
10. Методы прикладной математики в пожарно-технических задачах: лекции по курсу "Прикл. математика" / Н. Н. Брушлинский, Б. М. Пранов, Н. Н. Соболев [и др.]. – М.: ВИПТШ, 1983. – 134 с.
11. Пузач, С. В. Методы расчета тепломассообмена при пожаре в помещении и их применение при решении практических задач пожаровзрывобезопасности зданий: монография / С. В. Пузач. – М.: Академия ГПС МЧС России, 2005. – 336 с.
12. Чешко, И. Д. Применение расчетных методов в пожарно-технической экспертизе: учебное пособие / И. Д. Чешко, Ю. Д. Моторыгин, М. А. Галишев. – СПб.: СПб Университет МВД РФ, 2000. – 125 с.
13. Broadbent, S. R. Percolation Processes / S. R. Broadbent, J. M. Hammersley // Mathematical Proceedings of the Cambridge Philosophical Society. – 1957. – V. 53. – P. 629–641.
14. Ковалев, А. П. О проблемах оценки безопасности электротехнических объектов / А. П. Ковалев // Электричество. – 1991. – №8. – С. 50-54.

Поступила в редакцию 09.09.2024 г., рекомендована к печати 30.09.2024 г.

THE METHOD OF ANALYZING EMERGENCY MODES OF OPERATION OF CONTACT CONNECTIONS OF THE 0.4/0.22 KV POWER GRID

Bershadskii I.A., Mykh A.D., Zgarbul A.V., IAKimishina V.V.

This article analyzes the statistical level of fire danger in residential premises. Its influence on the reliability of the electrical system of apartments is evaluated and the role of spark protection devices in this process is indicated. In addition, a method for determining the probability of a source of ignition in electrical wiring, taking into account various damaging factors, the intensity of their occurrence and elimination, is shown and described.

Keywords: risk, fire safety, power supply system, load node, electrified room, protective equipment.

Бершадский Илья Адольфович

доктор технических наук, доцент, заведующий кафедрой электроснабжения промышленных предприятий и городов ФГБОУ ВО «Донецкий национальный технический университет», Российская Федерация, ДНР, г. Донецк.
E-mail: iberish164@yandex.ru

Bershadskii Iliia Adolfovich

Doctor of Technical Sciences, Associate Professor, Head of the Department of Power Supply of Industrial Enterprises and Cities of Donetsk National Technical University, Russian Federation, DPR, Donetsk.

Мых Анатолий Дмитриевич

аспирант кафедры электроснабжения промышленных предприятий и городов ФГБОУ ВО «Донецкий национальный технический университет», Российская Федерация, ДНР, г. Донецк.
E-mail: vip.myh@mail.ru

Mykh Anatolii Dmitrievich

Postgraduate Student at Department of Power Supply of Industrial Enterprises and Cities of the Donetsk National Technical University, Russian Federation, DPR, Donetsk.

Згарбул Андрей Викторович

кандидат технических наук, доцент кафедры электроснабжения промышленных предприятий и городов ФГБОУ ВО «Донецкий национальный технический университет», Российская Федерация, ДНР, г. Донецк.
E-mail: zgarbul.andrey@gmail.com

Zgarbul Andrei Viktorovich

Candidate of Technical Sciences, Associate Professor at Department of Power Supply of Industrial Enterprises and Cities of the Donetsk National Technical University, Russian Federation, DPR, Donetsk.

Якимишина Виктория Викторовна

кандидат технических наук, доцент кафедры электроснабжения промышленных предприятий и городов ФГБОУ ВО «Донецкий национальный технический университет», Российская Федерация, ДНР, г. Донецк.
E-mail: yvsm@list.ru

IAkimishina Viktoriia Viktorovna

Candidate of Technical Sciences, Associate Professor at Department of Power Supply of Industrial Enterprises and Cities of the Donetsk National Technical University, Russian Federation, DPR, Donetsk.

АНАЛИЗ РАЗЛИЧНЫХ СХЕМ ИСПОЛЬЗОВАНИЯ ТЕПЛОВОЙ ЭНЕРГИИ: МИКРОУРОВЕНЬ ТЕПЛОПОТРЕБЛЕНИЯ

© 2024 *Бирюков А.Б., Гаращенко В.Н.*

Проведен анализ различных схем использования тепловой энергии конечными ее потребителями и на микроуровне теплотребления разработано определение модели единой микро-энергетической системы (UMES - unified micro-energy system) энергоснабжения конечного потребителя, которая синтезирована для представления общего энергопотребления как квартиры или дома, так и любого другого объекта, который используется Человеком, в течении определенного периода. Предлагаемая модель единой микро-энергетической системы включает в себя, в качестве подсистем генерации энергии централизованные и децентрализованные системы энергоснабжения. Также, для комплексного рассмотрения необходимой и достаточной потребности в тепловой, электрической и других видах энергии разработано определение единицы условной энергии (unit of conditional energy U.C.E.). Определены направления использования предлагаемых решений: сравнение фактических количественных и качественных значений состояния энергетической системы с наиболее рациональными; определение этапов развития системы; формирование состава работ и необходимого оборудования для реконструкции и нового строительства; снижение негативного воздействия на окружающую среду, в том числе от “парниковых газов”, до минимально возможного уровня; повышение эффективности использования невозобновляемых источников энергии, до тех пор, пока они используются; минимизирование совокупных энергетических затрат на производство оборудования для использования возобновляемых источников энергии; улучшение комплексной энергетической эффективности.

Ключевые слова: возобновляемые источники энергии, невозобновляемые источники энергии, когенерация, централизованные системы энергоснабжения, децентрализованные системы энергоснабжения, микроуровень теплотребления, единая микро-энергетическая система, единица условной энергии, комплексная энергетическая эффективность, типовые схемы теплоснабжения.

Введение. Рост использования возобновляемых источников энергии предопределяет необходимость рационального решения задачи определения необходимой мощности энергогенерирующего оборудования и повышения энергоэффективности работы существующих систем централизованного энергоснабжения, которые вырабатывают тепловую и электрическую мощность в когенерационном режиме. Возобновляемые источники энергии используются как в централизованных системах энергоснабжения, так и в децентрализованных системах энергоснабжения. При этом потребность в тепловой и других видах энергии для различных схем теплоснабжения и теплотребления изменяется многократно в течение суток и генерация энергии с использованием возобновляемых источников энергии зависит от времени суток также. В связи с этим, для улучшения работы энергосистем, важно согласовывать использование возобновляемых и невозобновляемых источников энергии в централизованных и децентрализованных энергетических системах.

Использование возобновляемых источников приобретает все большую значимость в связи с экологическими проблемами и исчерпаемостью невозобновляемых источников энергии. Минимальные выбросы веществ загрязняющих окружающую среду при максимальной энергетической эффективности обеспечиваются только в определенном диапазоне работы теплогенерирующего оборудования и поэтому значение способа производства тепловой энергии, как и схемы теплоснабжения становится все более определяющим.

Технологии совершенствуются с учетом места установки тепло- и энергогенерирующего оборудования в централизованных и децентрализованных системах энергоснабжения, а конечный Потребитель, в большинстве случаев, один и тот же. При этом имеются неиспользованные резервы традиционных технологий – это развитие когенерации, снижение потерь и использование вторичной энергии.

Основные направления эволюционного и энергоэффективного развития систем энергоснабжения и энергопотребления, в структуре которых тепловая энергия составляет большую часть, можно сгруппировать следующим образом:

- комбинированная генерация тепловой и электрической энергии в децентрализованных системах на объектах жилой [1], коммерческой и промышленной недвижимости;

- организация согласованной работы децентрализованных и централизованных энергетических систем путем встраивания распределенной энергетики в традиционную систему там, где это экономически целесообразно с учетом конкретных условий [2];

- комбинированное использование различных видов энергоносителей для использования на объектах потребления тепловой [3] и другой энергии;

- комбинированное (гибридное) использование возобновляемых и невозобновляемых источников энергии с учетом инвентаризация потоков энергии и сравнения фактического с необходимым и достаточным количеством энергии, что позволяет определять возможные и необходимые диапазоны регулирования элементов подсистем системы генерации тепловой энергии [4].

В ряде случаев приходится решать проблему избыточной доли генерации при использовании ветровой и фотоэлектрической энергии [5], что свидетельствует о сложности определения оптимальных параметров системы.

Цель функционирования системы энергоснабжения - обеспечение конечного потребителя необходимым и достаточным количеством тепловой и другими видами энергии.

Существующие противоречия между возобновляемыми и невозобновляемыми источниками энергии, а также между децентрализованными и централизованными системами тепло- и энергоснабжения являются диалектическими и разрешение этих противоречий, в проводимых исследованиях, направлено на эволюционное развитие энергетических систем в целом.

Систематизация основных схем энергоснабжения для использования тепловой энергии. Потребность в тепловой энергии и соответствие генерации этой потребности определяет, в значительной степени, эффективность работы всей системы и ее подсистем. Гармонизация потребления и генерации повышает энергетическую эффективность и приводит к снижению вредных выбросов в окружающую среду.

Определение необходимого и достаточного количества тепловой энергии для потребителей, которые сгруппированы на одном объекте или в одной системе может определяться как суммарная потребность соответствующих микро-энергетических систем.

Для рассмотрения основных типов микро-энергетических систем UMES (unified micro-energy system) в части использования тепловой энергии, систематизируются следующим образом:

- UMES-1 - отопление и ГВС централизованные, кондиционирование децентрализованное;

- UMES-2 - отопление централизованное, ГВС и кондиционирование децентрализованные;

- UMES-3 – отопление, ГВС и кондиционирование децентрализованное, при наличии централизованного газоснабжения;
- UMES-4 – отопление, ГВС и кондиционирование децентрализованное, при отсутствии централизованного газоснабжения;
- UMES-5 – отопление, ГВС и кондиционирование – все централизованные.

В схеме 1 (рисунок 1) отопление централизованное (работает только в отопительный сезон, горячее водоснабжение (ГВС) централизованное и, в большинстве случаев, осуществляется круглогодично, кондиционирование децентрализованное и осуществляется только в жаркие месяцы года. Централизованное электроснабжение используется для децентрализованного кондиционирования. Централизованное водоснабжение (ХВС) и централизованное топливоснабжение (газоснабжение) для производства тепловой энергии на объектах теплоснабжения не используются.

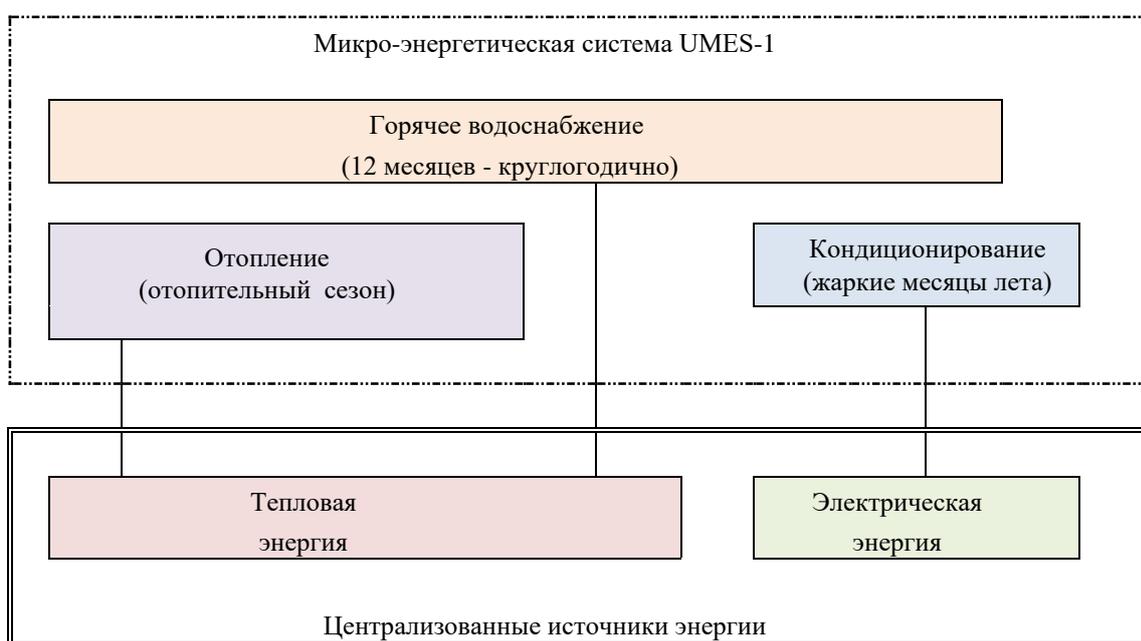


Рис. 1. Схема использования тепловой энергии при централизованных отоплении и горячем водоснабжении, тепловая энергия при кондиционировании производится децентрализованно (с использованием электрической энергии)

В схеме UMES-1 для нужд теплоснабжения микросистем (лицевых счетов) в многоквартирных домах тепловая энергия генерируется теплоэлектроцентралями (ТЭЦ) или котельными, как источниками в централизованных системах теплоснабжения. Схема UMES-1 распространена в крупных городах, энергоснабжение в которых осуществляют ТЭЦ и производится комбинированная выработка тепловой и электрической энергии.

В схеме 2 (рисунок 2) отопление централизованное (работает только в отопительный сезон). Централизованное водоснабжение (ХВС) и централизованное электроснабжение используются для децентрализованного горячего водоснабжения (ГВС). Централизованное электроснабжение используется для децентрализованного кондиционирования (в жаркие месяцы года). Для нужд теплоснабжения микросистем, так называемых, лицевых счетов в многоквартирных домах выше 5-ти этажей и при отсутствии сетей газоснабжения для горячего водоснабжения используются электрические бойлеры.

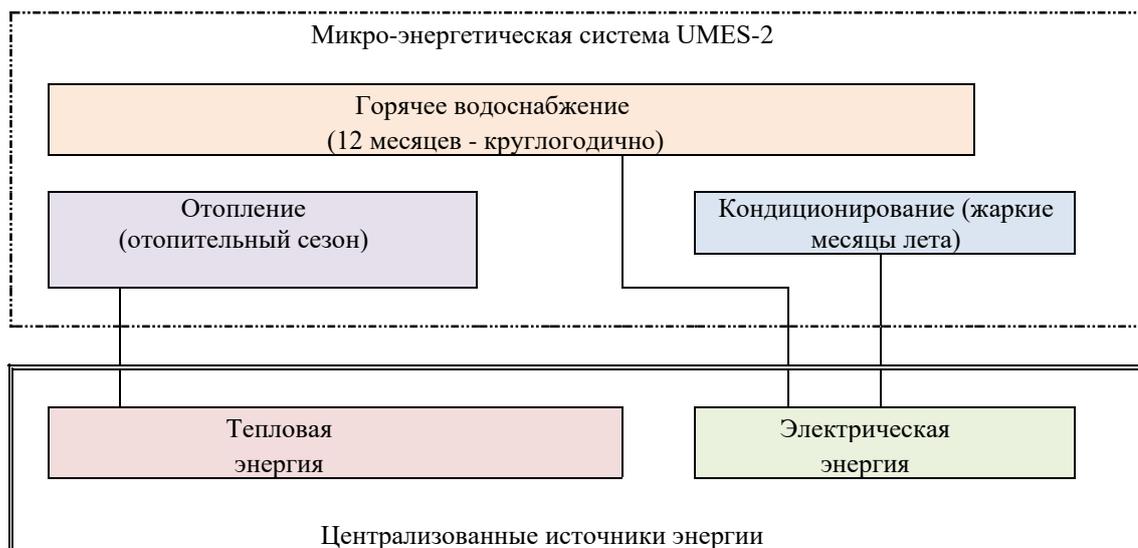


Рис. 2. Схема использования тепловой энергии при централизованном отоплении, тепловая энергия для кондиционирования и горячего водоснабжения производится децентрализованно (в рассматриваемом случае, с использованием электрической энергии)

Для нужд теплоснабжения в многоквартирных домах до 5-ти этажей и при наличии сетей газоснабжения для горячего водоснабжения используются газовые колонки. В обоих вариантах UMES-2 для децентрализованной генерации тепловой энергии используются централизованные источники энергии. Схема UMES-2 распространена в городах, теплоснабжение в которых осуществляют котельные, доля комбинированной выработки тепловой и электрической энергии на которых незначительна, либо отсутствует.

В схеме 3 (рисунок 3) децентрализованное отопление (когда это необходимо, что повышает качество теплоснабжения) и децентрализованное горячее водоснабжение ГВС (круглогодично и в необходимом количестве) производятся с использованием централизованного топливоснабжения (газоснабжения).

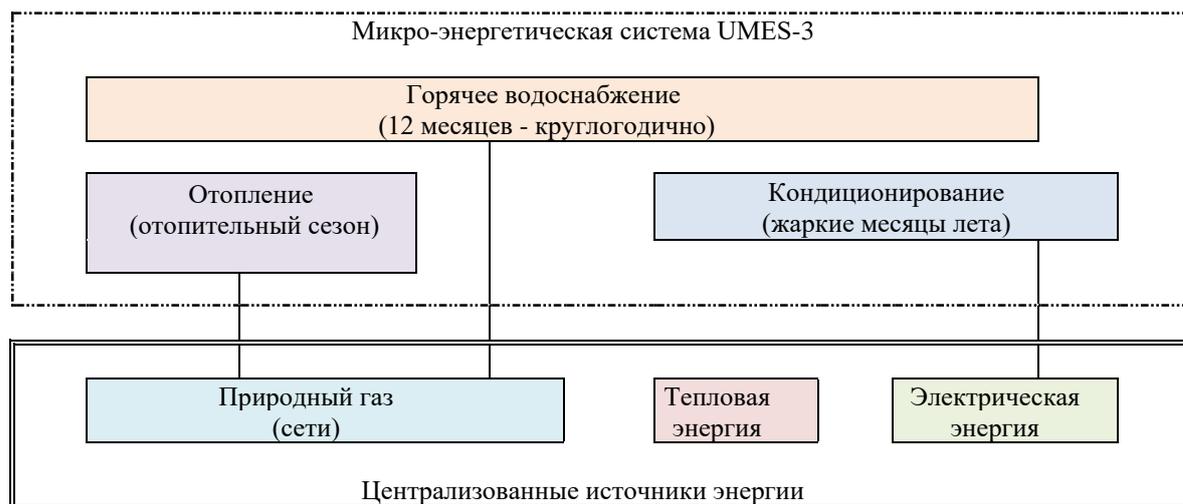


Рис. 3. Схема использования тепловой энергии при децентрализованном отоплении и децентрализованном горячем водоснабжении (с использованием централизованного газоснабжения), тепловая энергия для кондиционирования производится децентрализованно (с использованием централизованного электроснабжения)

Централизованное электроснабжение используется для децентрализованного кондиционирования (в жаркие месяцы года). Таким образом, тепловая энергия, и для отопления, и для горячего водоснабжения, и для кондиционирования в UMES-3 вырабатывается децентрализованно, но источники исходной энергии (природный газ и электроэнергия) транспортируются централизованно.

Такая система используется в тех случаях, когда, например, двухконтурный котел работает для отопления и горячего водоснабжения квартиры, индивидуального дома или аналогичного объекта. Схема UMES-3 позволяет наиболее эффективно использовать природный газ, так как и тепловая энергия для отопления, и нагрев воды для горячего водоснабжения производятся (генерируются) в необходимом и достаточном количестве непосредственно в месте потребления.

Схема UMES-3 снижает потери энергии при передаче от теплогенерирующего к теплоиспользующему оборудованию до минимальных значений.

В схеме 4 (рисунок 4) при отсутствии централизованного газоснабжения отопление и горячее водоснабжение (ГВС) производятся децентрализованно. Централизованное электроснабжение используется для децентрализованного кондиционирования (в жаркие месяцы года), а так же для отопления и ГВС при использовании электрокотлов и электроводонагревателей.

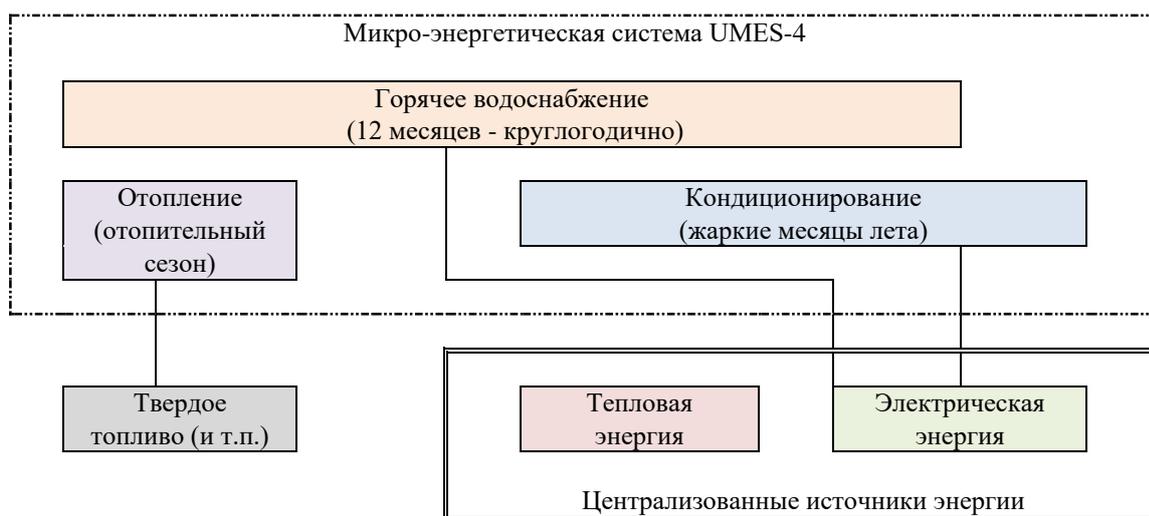


Рис. 4. Схема использования тепловой энергии при децентрализованном отоплении (с использованием твердого топлива) и децентрализованном горячем водоснабжении (с использованием централизованного электроснабжения), тепловая энергия для кондиционирования производится децентрализованно (с использованием централизованной электрической энергии)

UMES-4 используется, преимущественно, для теплоснабжения индивидуальных жилых домов, не подключенных к газовым сетям. В качестве основного топлива используется, преимущественно твердое топливо. В системе UMES-4 использование кондиционера в режиме «теплового насоса» в начале и по завершении отопительного сезона, а также в период оттепели позволяет, при необходимости, осуществлять теплоснабжение на нужды отопления без запуска твердотопливного котла, как основного отопительного децентрализованного оборудования. Также, наряду с кондиционерами (в режиме отопления) могут использоваться электрокотлы, работа которых может быть согласована с твердотопливным котлом для удобства эксплуатации, повышения качества отопления и снижения общих энергорасходов.

В схеме 5 (рисунок 5) отопление, горячее водоснабжение (ГВС) и кондиционирование производятся централизованно. Централизованное водоотведение стоков ГВС производится централизованно.

Схема UMES-5, при которой централизованно осуществляется и отопление, и кондиционирование с использованием тепловой энергии, используют компоновку оборудования «чиллер-фанкойл». При централизованном производстве горячей воды для UMES-5 все источники тепловой энергии будут централизованными. В такой схеме электрическая энергия внутри микро-энергетической системы не используется для трансформации в тепловую энергию.

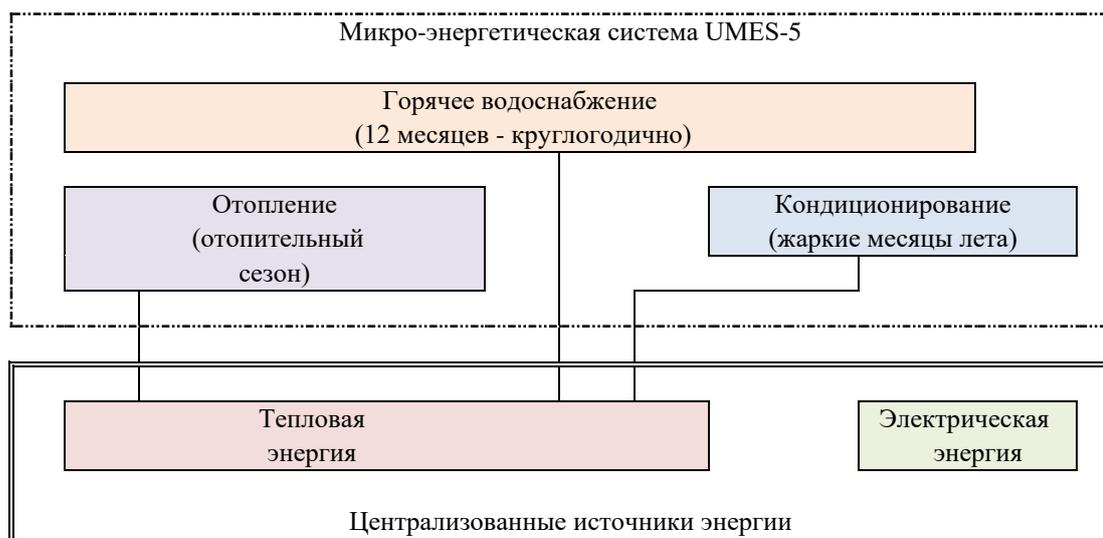


Рис. 5. Схема использования тепловой энергии при централизованном отоплении, централизованном горячем водоснабжении, тепловая энергия для кондиционирования производится также централизованно

Единая микро-энергетическая система. Для определения рациональных схем и режимов необходимого диапазона работы энергогенерирующего оборудования предлагается использовать синтезированную виртуальную модель единой микро-энергетической системы (ЕМЭС), в которой объединяются равнозначные подсистемы децентрализованного потребления энергии возобновляемых источников на самом объекте потребления и централизованного энергоснабжения (в соответствующей доле) с комбинированным использованием невозобновляемых и возобновляемых источников. В этой модели потребление всех видов энергии, в том числе тепловой, объединяется и рассматривается в единицах условной энергии (unit of conditional energy - U.C.E.). Единица условной энергии, в зависимости от конкретного случая, может соответствовать 1 Вт*час, 1 кВт*час, 1 МВт*час, 1 Гкал и другим физическим единицам измерения энергии. Структурная схема единой микро-энергетической системы UMES ЕМЭС представлена на рисунке 6.

Синтезирование идеальной модели системы теплоснабжения с учетом временной потребности в энергии (которая получена от возобновляемых источников энергии и от самых эффективных схем использования невозобновляемых источников энергии) позволяет, сравнивая с текущим состоянием, определять потенциал повышения эффективности и необходимые сроки для планируемых этапов развития этой системы.

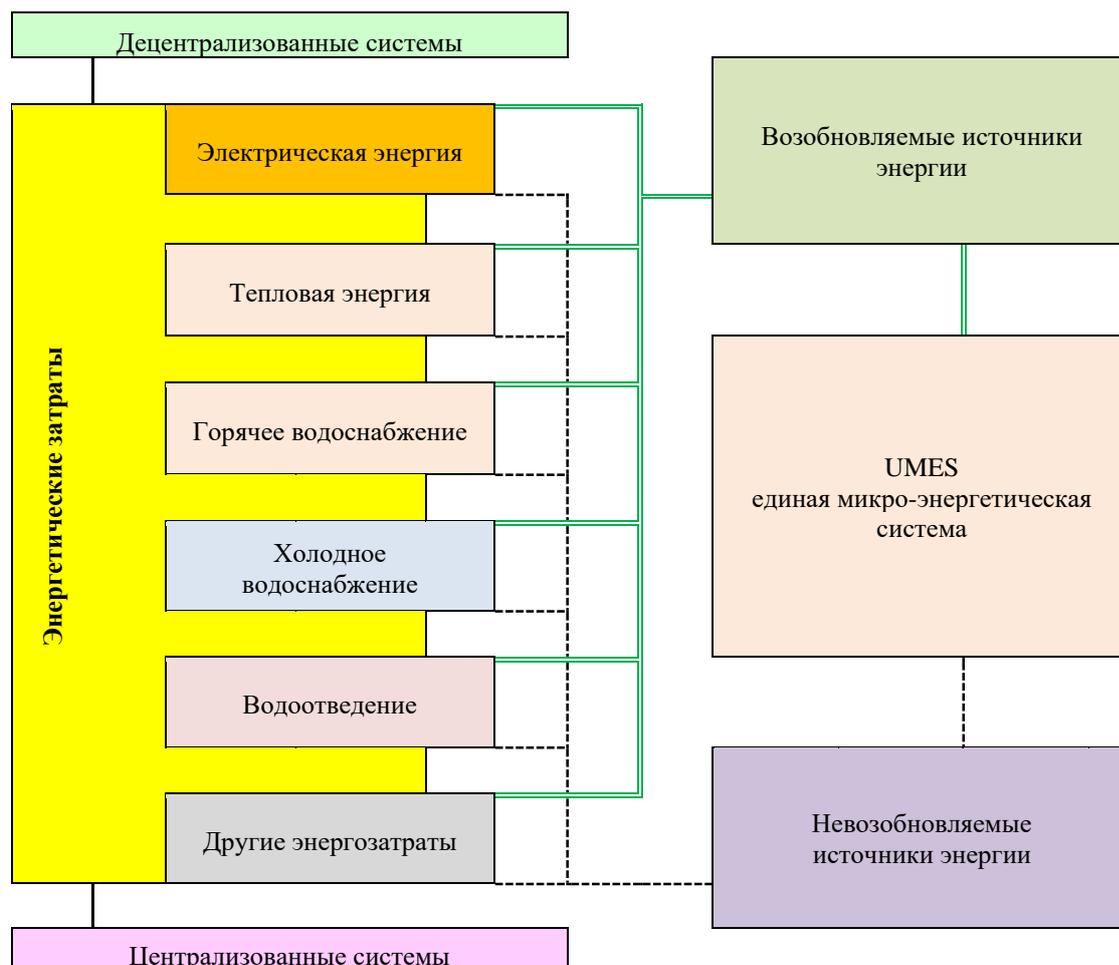


Рис. 6. Принципиальная схема UMES (unified micro-energy system)

Для анализа существующих и перспективных систем энергоснабжения могут учитываться возможные изменения в генерации, соответствие или несоответствие потребности в тепловой и других видах энергии суммарно для каждого из периодов.

Особенностями единой микро-энергетической системы UMES являются:

- определение общей потребности в энергии, в том числе тепловой с использованием единиц условной энергии;
- границы определяются общим количеством использованной энергии конечным потребителем и соответствующей долей генерации;
- наибольшее соответствие необходимому и достаточному количеству энергии с минимальными возможными отклонениями.

Рассмотрение схем теплоснабжения и теплопотребления с использованием единых микро-энергетических систем позволяет совершенствовать методики оценки энергоэффективности, определять перспективные схемы реконструкции и нового строительства и упрощать совместимость теплоэнергетических и других цифровых моделей, в том числе при строительстве зданий [6].

Необходимо учитывать, что энергетические расходы тепловой энергии на нужды кондиционирования, как разновидности тепловой энергии, растут. Эффективное использование возобновляемых источников энергии с учетом возможностей «виртуальных хранилищ» [7] для кондиционирования и другого использования тепловой энергии связано с оптимальным планированием [8] и большей эффективностью использования энергии солнца в жаркий период года.

Система централизованного теплоснабжения состоит из множества микро-систем UMES и для каждой из них генерируется соответствующая доля тепловой энергии на исходном источнике преобразования энергии, что может быть использовано в динамической модели оптимизации мощности системы, в том числе с использованием возобновляемых источников энергии [9].

Особенности рассматриваемых типовых схем теплоснабжения. Рассмотрение основных принципиальных схем UMES позволяет определить особенности рассматриваемых схем теплоснабжения, которые обусловлены составом оборудования, определенным исходной проектной документацией, а также внесенными изменениями и дополнениями проведенных реконструкций, замен и нового строительства:

- Схема UMES-5 по числу структурных элементов предварительно рассматривается как оптимальная, так как число ее элементов минимально. При когенерации электрическая энергия может использоваться для зарядки электромобилей, с учетом использования солнечной энергии [10];

- Схема UMES-4 – это схема, в которой объект из централизованных систем энергоснабжения подключен только к системе электроснабжения (в основном это малые города и другие населенные пункты);

- Схема UMES-3 – это схема, при которой теплоснабжение осуществляется индивидуальным теплогенерирующим оборудованием для ГВС, отопления и кондиционирования (перспективное направление для улучшения состояния тепловых сетей за счет перевода проблемных участков на индивидуальное теплоснабжение) с возможностью использования возобновляемых источников энергии [11];

- Схемы UMES-1 и UMES-2 – это самые распространенные схемы в больших городах и вблизи крупных станций (ТЭЦ и ТЭС), нормативные значения для которых зависят от климатических условий [12].

Выводы. Таким образом, исследование различных схем использования тепловой энергии с целью классификации таких систем на микро-уровне привело к следующим результатам:

1. Проведен анализ существующих схем тепло- и энергоснабжения для использования при определении текущей и возможной энергетической эффективности, планирования реконструкций и нового строительства. Особенности распределения энергии между типовыми микро-энергетическими системами позволяют классифицировать их, как базовые схемы UMES энергопотребления для моделирования различных режимов работы и выбора наиболее рациональных схем из конечного числа возможных вариантов.

2. Разработано определение единой микро-энергетической системы UMES, в формировании измеримых параметров которой используются единицы условной энергии U.C.E. для определения необходимых технических решений для каждого из этапов развития энергосистемы.

3. Концепция UMES позволяет рассматривать противоречия между децентрализованными и централизованными системами энергоснабжения, а также между использованием возобновляемых и невозобновляемых источников энергии как диалектические, разрешение которых позволяет управлять эволюционным развитием энергосистемы.

4. Согласование работы децентрализованных и централизованных систем энергоснабжения в наиболее рациональной схеме (для конечного потребителя тепловой и других видов энергии) может позволить при увеличении доли используемых возобновляемых источников энергии до оптимального значения, повысить эффективность использования невозобновляемых источников энергии до максимально необходимого и возможного уровня. Дальнейшие прикладные исследования позволят разработать решения, которые позволят получать наряду с энергетическим, значимый экологический эффект.

5. Определены направления дальнейших исследований и применения предлагаемых решений на основе модели UMES для объединения генерации, транспортировки и использования тепловой энергии со смежными системами использования других видов энергии в едином общем энергопотреблении:

- разработка основных положений методики расчёта эффективности циклов работы установок и систем теплоснабжения и энергоснабжения;
- развитие когенерации и мультигенерации;
- минимизация несоответствия генерации и потребления необходимой и достаточной энергии для улучшения энергетической эффективности системы в целом и снижения выбросов вредных веществ, в том числе парниковых газов.

Дальнейшие исследования направлены на определение измеримых значений и условий согласованной работы централизованных и децентрализованных систем энергоснабжения (с учетом состава оборудования, климатических и других условий), что позволит совершенствовать рациональные схемы комбинированного использования возобновляемых и невозобновляемых источников энергии.

СПИСОК ИСТОЧНИКОВ

1. Кодолов, Г. О. Перспективы применения инновационных автономных систем энергоснабжения в малоэтажной жилой застройке / Г. О. Кодолов, З. К. Петрова // Биосферная совместимость: человек, регион, технологии. – 2014. – № 4(8). – С. 100-112. – EDN TCBOJT.
2. Маркова, В. М. Децентрализация энергетики: интеграция и инновации / В. М. Маркова, В. Н. Чурашев // ЭКО. – 2020. – № 4(550). – С. 8-27. – DOI 10.30680/ECO0131-7652-2020-4-8-27. – EDN QCCJKR.
3. Торопов, А. Л. О перспективах применения электродкотлов для автономных систем теплоснабжения малой мощности / А. Л. Торопов // Сантехника, Отопление, Кондиционирование. – 2022. – № 2(242). – С. 32-35. – EDN ECLKHS.
4. Бирюков, А. Б. Направления повышения энергоэффективности систем генерации тепловой энергии в рамках теплоэнергетики России / А. Б. Бирюков, В. Н. Гаращенко // Энергетические системы. – 2023. – № 2. – С. 17-27. – DOI 10.34031/es.2023.2.002. – EDN KIVSMX.
5. Optimal scheduling of the CSP–PV–wind hybrid power generation system considering demand response / Q. Zhao [et al.] // Frontiers Energy Research. – 2023. – V. 11. – DOI 10.3389/fenrg.2023.1252311.
6. Digital twin modeling method based on IFC standards for building construction processes / C. Dai [et al.] // Frontiers Energy Research. – 2024. – V. 12. – DOI 10.3389/fenrg.2024.1334192.
7. Dynamic pricing optimization for commercial subcontracting power suppliers engaging demand response considering building virtual energy storage / H. Huang [et al.] // Frontiers Energy Research. – 2023. – V. 11. – DOI 10.3389/fenrg.2023.1329227.
8. The energy management strategy of a loop microgrid with wind energy prediction and energy storage system day-ahead optimization / B. Xu [et al.] // Frontiers Energy Research. – 2023. – V. 11. – DOI 10.3389/fenrg.2023.1334588.
9. A dynamic hierarchical partition method for optimal power balance of urban power system with high renewable / J. Ye [et al.] // Frontiers Energy Research. – 2024. – V. 12. – DOI 10.3389/fenrg.2024.1355606.
10. Integration of solar based charging station in power distribution network and charging scheduling of EVs / A. Shafiq [et al.] // Frontiers Energy Research. – 2024. – V. 11. – DOI 10.3389/fenrg.2023.1086793.
11. A state-of-the-art review of energy-efficient and renewable energy systems in higher education facilities / R. A. Almasri [et al.] // Frontiers in Energy Research. – 2023. – V. 11. – DOI 10.3389/fenrg.2023.1344216.
12. Subdividing end-use energy consumption based on household characteristics and climate conditions: insights from urban China / T. Wang [et al.] // Frontiers in Energy Research. – 2023. – V. 11. – DOI 10.3389/fenrg.2023.1267975.

Поступила в редакцию 19.09.2024 г., рекомендована к печати 10.10.2024 г.

ANALYSIS OF VARIOUS SCHEMES FOR THE USE OF THERMAL ENERGY: MICRO-LEVEL OF HEAT CONSUMPTION

Biriukov A.B., Garashchenko V.N.

The analysis of various schemes for the use of thermal energy by its end users and at the microlevel of heat consumption is carried out. The definition of a model of a unified micro-energy system (UMES) of energy supply to the end user has been developed, which is synthesized to represent the total energy consumption of an apartment or house, as well as any other object that is used by a person for a certain period of time. The proposed model of a unified micro-energy system includes centralized and decentralized energy supply systems as energy generation subsystems. Also, for a comprehensive consideration of the necessary and sufficient demand for thermal, electrical and other types of energy, a definition of the unit of conditional energy (U.C.E.) has been developed. The directions of using the proposed solutions are determined: comparing the actual quantitative and qualitative values of the state of the energy system with the most rational ones; determining the stages of system development; forming the composition of work and necessary equipment for reconstruction and new construction; reducing the negative impact on the environment, including from “greenhouse gases”, to the lowest possible level; increasing the efficiency of use non-renewable energy sources, as long as they are used; minimizing the total energy costs for the production of equipment for the use of renewable energy sources; improving integrated energy efficiency.

Keywords: renewable energy sources, non-renewable energy sources, cogeneration, centralized energy supply systems, decentralized energy supply systems, micro-level of heat consumption, unified micro-energy system; unit of conventional energy, integrated energy efficiency, typical heat supply schemes.

Бирюков Алексей Борисович

доктор технических наук, профессор, заведующий кафедрой технической теплофизики, проректор ФГБОУ ВО «Донецкий национальный технический университет»,
Российская Федерация, ДНР, г. Донецк.
E-mail: birukov.ttf@gmail.com

Biriukov Aleksei Borisovich

Doctor of Technical Sciences, Professor, Head of the Department of Technical Thermophysics, Vice-rector of Donetsk National Technical University,
Russian Federation, DPR, Donetsk.

Гарашенко Валерий Николаевич

младший научный сотрудник ФГБОУ ВО «Донецкий национальный технический университет»,
Российская Федерация, ДНР, г. Донецк.
E-mail: gwel@mail.ru

Garashchenko Valerii Nikolaevich

Junior Researcher of Donetsk National Technical University,
Russian Federation, DPR, Donetsk.

МЕТОДОЛОГИЯ АНАЛИЗА НЕБАЛАНСА ЭЛЕКТРИЧЕСКОЙ ЭНЕРГИИ НА ТЯГУ ПОЕЗДОВ

© 2024 *Давыдов А.И., Каштанов А.Л., Подгорная С.О.*

Приведено описание, разработанной в Омском государственном университете путей сообщения методологии анализа небаланса электрической энергии на тягу поездов на основе статистической информации автоматизированных систем. Методология может быть реализована при создании системы поддержки принятия решений в области анализа и прогнозирования расхода ТЭР на тягу поездов.

Ключевые слова: эксплуатационные показатели, небаланс электроэнергии, анализ, коэффициенты влияния, ожидаемое значение величины.

Введение. В соответствии с ГОСТ 32895-2014 небалансом электрической энергии на тягу поездов называют разницу между объемом расхода электроэнергии, относимого на тягу поездов по показаниям счетчиков тяговых подстанций (с включением соответствующей доли – от общего потребления электроэнергии – технических потерь в тяговых трансформаторах и линиях электропередач до границы балансовой принадлежности) и объемом расхода электроэнергии по показаниям счетчиков электроподвижного состава.

На практике же в настоящий момент значение небаланса электроэнергии на тягу поездов на полигоне постоянного тока определяется без учета потерь электроэнергии в понизительных трансформаторах, а на полигоне переменного тока при наличии учета электроэнергии на фидерах контактной сети тяговых подстанций небаланс определяется без потерь в тяговых трансформаторах.

Небаланс электроэнергии включает в себя составляющую реальных потерь электроэнергии – технологическую, и условную составляющую – так называемые коммерческие потери [1-3].

Небаланс включает в себя технологическую и «коммерческую» составляющие. Технологическая составляющая включает потери электроэнергии при ее транспортировке от вводов тяговых подстанций до электроподвижного состава, а также погрешности приборов учета. Так называемая «коммерческая» составляющая не связана с реальными потерями электрической энергии и обусловлена в первую очередь следующими причинами: низким качеством технических средств учета электроэнергии, методическими ошибками и погрешностями при определении интегральных значений небаланса в границах расчетного участка (рис. 1).

В последнее десятилетие на сети железных дорог проводилось большое количество организационно-технических мероприятий, направленных на повышение энергоэффективности системы тягового электроснабжения и тягового подвижного состава, в результате которых удалось добиться двукратного снижения среднего уровня небаланса электроэнергии на тягу поездов на сети железных дорог с 14,7 % в 2005 году до 9,3 % в 2023 году.

Однако, не смотря на довольно низкий среднегодовой уровень небаланса по сети железных дорог, на отдельных железных дорогах в отдельные месяцы могут происходить значительные его колебания как в сторону значительного увеличения, так и в сторону резкого снижения, вплоть до парадоксальных отрицательных значений.

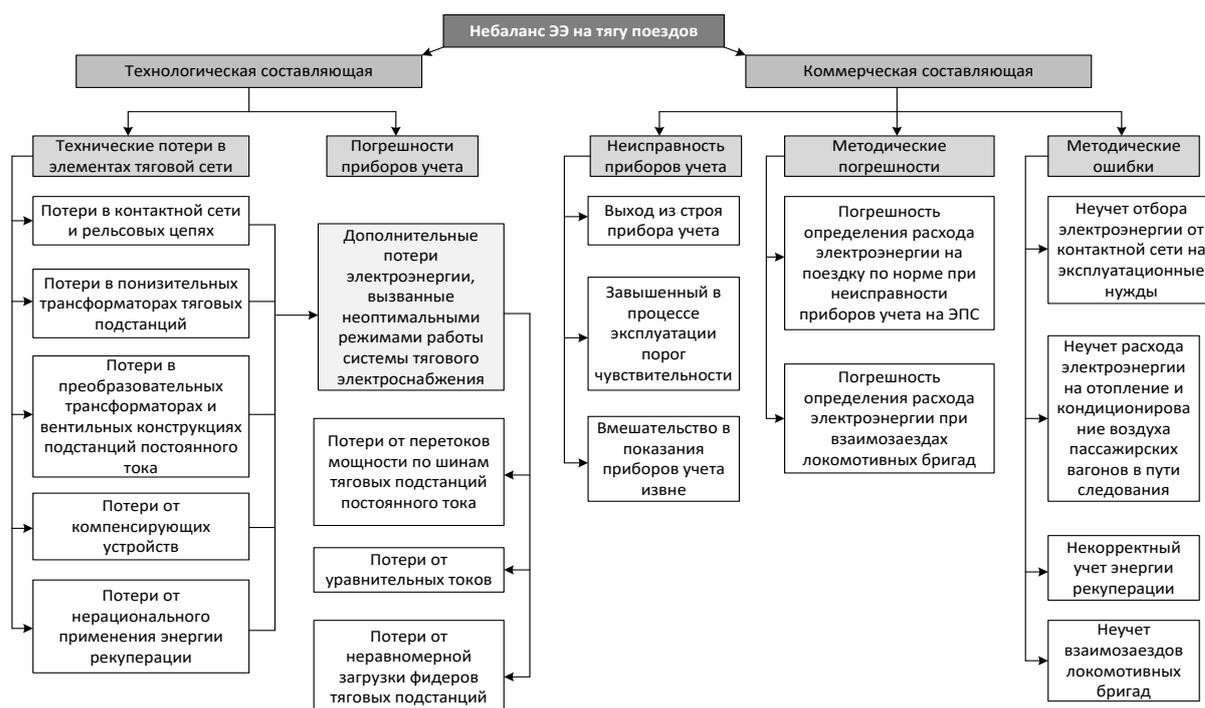


Рис. 1. Структурная схема Фурье-спектрометра с интерферометром TWINS при максимальной толщине блока Б

Такие колебания могут быть вызваны как объективными причинами, например, изменением структуры грузопотоков, состава парка подвижного состава, климатических условий и др., так и субъективными: неудовлетворительное состояние объектов системы тягового электроснабжения и электроподвижного состава, неэффективное управление режимами их работы, неисправность приборов учета электроэнергии и несанкционированное вмешательство в их работу, методические погрешности и методические ошибки при расчете уровня небаланса и др.

В этой связи необходим четкий механизм определения ожидаемых значений небаланса на тот или иной период времени для различных железных дорог, который бы учитывал все особенности этих железных дорог.

Основная часть. Анализ специфики определения ожидаемых значений небаланса позволил сформировать перечень требований к выбору метода расчета ожидаемого значения, которые заключаются в следующем:

- 1) при разработке расчетной модели участие эксперта должно быть сведено к минимуму;
- 2) метод расчета должен иметь возможность работать с зашумленными данными;
- 3) процесс расчета ожидаемых значений должен быть понятен специалисту, использующему метод;
- 4) сложность и время реализации должны быть минимальными, метод должен быть легко автоматизируемым;
- 5) метод должен поддерживать возможность корректировки модели для учета дополнительного набора факторов, оказывающих воздействие на систему.

В таблице 1 представлен сравнительный анализ методов по указанным выше критериям.

Таблица 1. Результаты сравнительного анализа методов прогнозирования

Метод	Номер требования				
	1	2	3	4	5
Экстраполяционные методы	–	+	+	+	+
Регрессионные методы	–	–	+	+	+
Экспертные методы	–	–	+	–	–
Нейронные сети	+	–	–	–	+
Нейро-нечеткие системы	+	+	–	–	+

Учитывая достоинства и недостатки рассмотренных методов и предъявляемые к ним требования, для решения задачи поиска ожидаемого значения небаланса электроэнергии на тягу поездов наиболее подходящими являются методы экстраполяции временных рядов, которые заключаются в распространении тенденций, установленных в прошлом, на будущий период, и регрессионные методы, также решающие задачу экстраполяции на основе исследования взаимного влияния показателей.

Для определения ожидаемого значения показателя были рассмотрены следующие:

- 1) абсолютная величина небаланса электрической энергии на тягу (П1);
- 2) относительная величина небаланса электрической энергии на тягу (П2);
- 3) отношение абсолютной величины небаланса электрической энергии на тягу к перевозочной работе (П3).

На следующем этапе для каждого показателя П1-П3 были построены корреляционные матрицы от следующих показателей:

- расход электроэнергии по приборам учета тяговых подстанций;
- расход по приборам учета электроподвижного состава;
- расход локомотивных бригад железной дороги на участках работы «чужой» железной дороги;
- расход локомотивных бригад «чужой» железной дороги на участках исследуемой железной дороги;
- фактический и плановый возврат электроэнергии в контактную сеть при рекуперативном торможении;
- расход электроэнергии на нужды предприятий по ремонту и техническому обслуживанию локомотивов;
- прочий расход электроэнергии из тяговой сети;
- техническая скорость движения;
- коэффициент участковой скорости;
- величина перевозочной работы по видам движения;
- резервный пробег локомотивов;
- простой локомотивов в рабочем состоянии в ожидании работы;
- нагон опоздания пассажирских поездов;
- нагрузка на ось грузового вагона состава;
- температура наружного воздуха.

В связи с большим объемом матриц приведем в качестве примера корреляционную матрицу показателей П1-П3 от величины пробега локомотивов резервом (табл. 2).

Таблица 2. Значения R^2 для регрессионных зависимостей небаланса от пробега локомотивов резервом по полигону железной дороги

Железная дорога	Полигон	Коэффициент детерминации R^2								
		2013 г.			2014 г.			2015 г.		
		П1	П2	П3	П1	П2	П3	П1	П2	П3
ЗСЖД	пост. ток	0,913	0,663	0,649	0,600	0,569	0,842	0,874	0,652	0,914
	перем. ток	0,927	0,426	0,649	0,703	0,379	0,910	0,800	0,495	0,784
ЮВЖД	пост. ток	0,787	0,712	0,820	0,786	0,446	0,927	0,868	0,627	0,925
	перем. ток	0,828	0,527	0,600	0,731	0,612	0,882	0,828	0,352	0,927
СКЖД	пост. ток	0,586	0,588	0,895	0,800	0,656	0,651	0,623	0,458	0,898
	перем. ток	0,895	0,500	0,817	0,912	0,682	0,891	0,694	0,497	0,853
ВСЖД	перем. ток	0,617	0,435	0,847	0,911	0,386	0,667	0,836	0,557	0,926

Согласно данным (табл. 2) пробег локомотивов резервом оказывает значительное влияние на величину небаланса по той же причине, что и время простоя локомотивов в рабочем состоянии в ожидании работы.

Результаты проведенных обширных исследований по выявлению факторов, оказывающих влияние на небаланс электрической энергии на тягу поездов, с помощью методов корреляционного регрессионного анализа позволили определить список указанных факторов:

- величина перевозочной работы в структурном подразделении;
- доля перевозочной работы в грузовом движении;
- техническая скорость движения в грузовом, пассажирском и хозяйственном движении;
- коэффициент участковой скорости в грузовом, пассажирском и хозяйственном движении;
- масса грузового, пассажирского и хозяйственного поезда;
- нагрузка на ось грузового вагона в грузовом и хозяйственном движении;
- доля порожнего пробега грузовых вагонов;
- время простоя локомотивов в рабочем состоянии в ожидании работы;
- пробег локомотивов резервом;
- усредненный индикатор энергоэффективности электровоза;
- температура атмосферного воздуха.

Анализ различных зависимостей показателей П1 – П3 (абсолютный небаланс, относительный небаланс и отношение величины небаланса к величине перевозочной работы) и указанных факторов показал, что в большинстве случаев максимальный коэффициент корреляции выявлен для зависимостей показателя П3. Именно его целесообразно использовать для построения регрессионных зависимостей ожидаемой величины небаланса от влияющих факторов.

Для учета влияния изменения эксплуатационных факторов на величину удельного небаланса (отношение величины небаланса и объема перевозочной работы) целесообразно составлять регрессионные зависимости:

$$\Delta w = \sum_{i=1}^n \beta_i X_i,$$

где Δw – ожидаемая величина небаланса;

β_i – коэффициенты влияния эксплуатационных факторов и показателей энергопотребления;

X – величина показателя из списка, приведенного выше, в рассматриваемом периоде.

Проведенные исследования позволили определить средние значения коэффициентов влияния, которые представлены в табл. 3.

Таблица 3. Коэффициенты влияния факторов на значение небаланса

№ п/п	Влияющий фактор	Значение коэффициента влияния (β_i)	
		Пост. ток	Перем. ток
1	Средняя масса поезда		
	в грузовом движении	0,001421	0,00237
	в пассажирском движении	-0,000630	-0,00055
2	Средняя техническая скорость		
	в грузовом движении	0,4931	0,38794
	в пассажирском движении	-0,0340	0,00875
3	Средний коэффициент участковой скорости в грузовом движении	-1,788	-2,0782
4	Удельная рекуперация	0,3496	0,17952
5	Доля работы ЭПС в грузовом виде движения	0,1066	0,08795
6	Усредненный индикатор энергоэффективности электровозов	-2,355	-1,74562
7	Перевозочная работа	$5,826 \cdot 10^{-7}$	
8	Средняя нагрузка на ось в грузовом движении	0,7677	
9	Порожний пробег вагонов	-0,1768	
10	Простой ЭПС в ожидании работы без локомотивной бригады	$9,760 \cdot 10^{-5}$	
11	Пробег ЭПС при следовании резервом	0,01037	
12	Время нагона опоздания пассажирских поездов	$8,1 \cdot 10^{-4}$	
13	Средняя температура атмосферного воздуха	-0,00584	

Выводы. Коэффициенты влияния целесообразно периодически пересчитывать с целью учета изменения факторов, которые на данном этапе невозможно учесть в силу отсутствия возможности объективной оценки значений.

Рассмотренный подход может быть использован при проведении анализа изменения небаланса электрической энергии на уровне структурного подразделения.

СПИСОК ИСТОЧНИКОВ

1. Черемисин, В. Т. Контроль достоверности учета и потерь электрической энергии в сетях железнодорожных узлов / В. Т. Черемисин, М. М. Никифоров // Экономия топливно-энергетических ресурсов на железных дорогах - текущие и перспективные задачи : Доклады и тезисы, представленные на конференцию: Сетевая научно-техническая практическая конференция, Москва, Экспериментальное кольцо ВНИИЖТ, 21–22 июня 2000 года / Министерство путей сообщения Российской Федерации, Государственное Унитарное Предприятие "Всероссийский научно-исследовательский институт железнодорожного транспорта"; Центральное правление Российского инженерно-технического Общества железнодорожников. – Москва, Экспериментальное кольцо ВНИИЖТ: Всероссийский научно-исследовательский институт железнодорожного транспорта, 2000. – С. 53-57. – EDN YMGBWH.
2. Каштанов, А. Л. Новые подходы к исследованию проблемы небаланса электроэнергии на тягу поездов / А. Л. Каштанов, М. М. Никифоров, С. Ю. Ушаков // Известия Транссиба. – 2017. – № 2(30). – С. 77-87. – EDN ZWQAGP.
3. Комяков, А. А. Оценка влияния межпоездного интервала на потери электроэнергии в тяговой сети с применением методов имитационного моделирования / А. А. Комяков, М. М. Никифоров, А. И. Шкулов // Актуальные проблемы проектирования и эксплуатации устройств электроснабжения электрического транспорта : Материалы всероссийской научно-технической конференции с международным участием, посвященной 90-летию со дня рождения д.т.н., профессора, академика Российской академии транспорта В.П. Михеева, Омск, 25–27 октября 2023 года. – Омск: Омский государственный университет путей сообщения, 2023. – С. 123-131. – EDN GPRAGH.

Поступила в редакцию 24.09.2024 г., рекомендована к печати 14.10.2024 г.

**METHODOLOGY FOR ANALYZING THE IMBALANCE OF ELECTRICAL ENERGY
FOR TRAIN TRACTION**

Davydov A.I., Kashtanov A.L., Podgornaia S.O.

A description is given of the methodology developed at the Omsk State Transport University for analyzing the imbalance of electrical energy on train traction based on statistical information from automated systems. The methodology can be implemented when creating a decision support system in the field of analysis and forecasting of fuel and energy resources consumption for train traction.

Keywords: operational indicators, power imbalance, analysis, influence coefficients, expected value.

Давыдов Алексей Игоревич

кандидат технических наук, доцент кафедры информационной безопасности ФГБОУ ВО «Омский государственный университет путей сообщения»,
Российская Федерация, г. Омск.
E-mail: davydovai@bk.ru

Davydov Aleksei Igorevich

Candidate of Technical Sciences, Associate Professor at the Department of Information Security of the Omsk State Transport University,
Russian Federation, Omsk.

Каштанов Алексей Леонидович

кандидат технических наук, директор Института автоматизации, телекоммуникаций и информационных технологий ФГБОУ ВО «Омский государственный университет путей сообщения»,
Российская Федерация, г. Омск.
E-mail: al.kashtanov@mail.ru

Kashtanov Aleksei Leonidovich

Candidate of Technical Sciences, Director of the Institute of Automation, Telecommunications and Information Technology of the Omsk State Transport University
Russian Federation, Omsk.

Подгорная Светлана Олеговна

кандидат технических наук, доцент кафедры информатики и компьютерной графики ФГБОУ ВО «Омский государственный университет путей сообщения»,
Российская Федерация, г. Омск.
E-mail: davydovai@bk.ru

Podgornaia Svetlana Olegovna

Candidate of Technical Sciences, Associate Professor at the Department of Informatics and Computer Graphics of the Omsk State Transport University,
Russian Federation, Omsk.

МОДЕЛИРОВАНИЕ ТЕМПЕРАТУРНОГО ПОЛЯ НАСАДОК РЕГЕНЕРАТИВНЫХ ГОРЕЛОК

© 2024 *Начкебия Н.С., Бирюков А.Б., Асламова Я.Ю.*

В работе представлена математическая модель для расчета температурных полей шариковой регенеративной насадки регенеративной горелки. Модель основана на элементарных балансах условных слоев насадки, а элементы насадки рассматриваются как термически массивные тела. В результате расчетов были получены данные распределения температур в теплоаккумулирующем слое насадки, и изменение во времени температур горячего и холодного теплоносителей. Полученные данные соответствуют практическому опыту применения регенеративных горелок и могут считаться достоверными.

Ключевые слова: энергоэффективность, регенеративная горелка, теплообмен, регенеративная насадка, тепловой баланс, методика расчета.

Введение. Энергоэффективность и экологичность современного производства являются неотъемлемым условием конкурентоспособности производимой продукции. Использование теплоты отходящих газов это один из наиболее эффективных способов повышения энергоэффективности. Утилизировать тепло отходящих газов можно различными способами. При выборе способа утилизации необходимо принимать во внимание температуру отходящих газов, их состав, особенности агрегатов производственного цикла и наличие места для размещения агрегата для утилизации тепла.

При высоких температурах теплоносителя рациональным является использование регенеративных теплообменников. Высокую эффективность на практике показало применение регенеративных горелок. Регенеративные горелки могут использоваться в диапазоне рабочих температур 800–1500 °С. Такие горелки позволяют утилизировать 85–90% теплоты отходящих из печи газов, обеспечивая подогрев поступающего воздуха горения до очень высоких температур, которые могут достигать величины всего на 100–150°С меньше, чем рабочая температура печи [1].

Принцип работы регенератора цикличен. Регенеративный цикл состоит из двух периодов: нагрева и охлаждения насадки. Поэтому, как правило, регенеративные насадки применяются парно. В первом периоде цикла один из регенераторов разогревается продуктами сгорания, второй охлаждается воздухом, идущим на горение. После «перекидки клапанов» происходит реверс газовых потоков, первый регенератор начинает охлаждаться, а второй нагревается отходящими газами [2].

Несмотря на существующий опыт эксплуатации таких горелочных устройств, все еще недостаточно исследованы процессы теплообмена в микрогенеративной насадке. На сегодняшний день не существует общепринятой методики расчета для регенеративных горелок. Однако, большинство исследователей сходятся во мнении, что для исследования теплотехнических свойств микрогенеративной насадки наиболее целесообразным является математическое моделирование процессов. Разными специалистами разработан ряд математических моделей этих процессов [3–6], однако во всех работах математическая модель составляется с определенной целью и не является универсальной.

Целью данной работы является создание математической модели теплообменных процессов в микрогенеративной насадке, учитывающей термическую массивность тел, образующих насадку.

Разработанная модель основывается на рассмотрении элементарных тепловых балансов условно выделенных по высоте насадки слоев теплоаккумулирующей засыпки (шариков). Также модель позволяет исследовать термическую массивность шарика и степень ее влияния на температурное поле насадки.

Важным параметром для моделирования является количество теплоты, которое вносит теплоноситель при входе в насадку:

$$Q_0 = V_0 \cdot c \cdot U_n, \text{Вт},$$

где V_0 – расход теплоносителя при НФУ, $\text{м}^3/\text{с}$;

c – теплоемкость теплоносителя, $\text{Дж}/(\text{м}^3 \cdot \text{К})$;

U_n – начальная температура теплоносителя, $^\circ\text{C}$.

Высота теплообменной насадки, занимаемую шариками (аккумулятором тепла), условно делим на равные части, шаг по высоте обозначим Δh . Далее для обозначения шага по толщине насадки используется индекс i , для обозначения шага по времени Δt – индекс k :

$$H = \Delta h \cdot n, \text{м},$$

где H – высота регенеративной насадки, м;

Δh – высота элементарного слоя, м;

n – количество слоев.

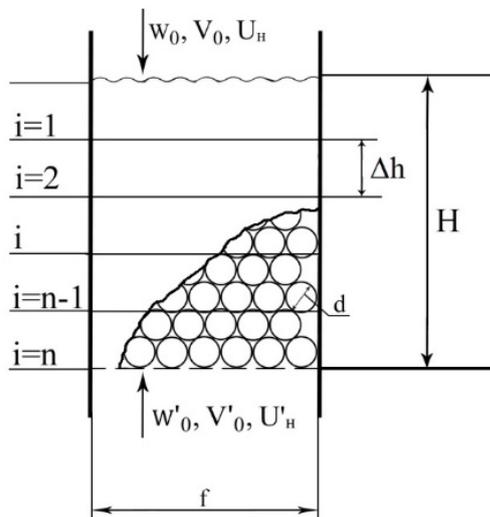


Рис. 1. Схема шариковой насадки, разделенной на элементарные слои и схематические потоки тепла

Принимаем, что теплоноситель входит в теплообменную насадку, и занимает элементарный объем $\Delta v = \Delta h \cdot f$, на время Δt . За время Δt теплоноситель отдает теплоаккумулирующей насадке количество тепла $\Delta Q_{i,k}$. Примем допущение, что при расчете тепловых потерь через ограждающие стенки регенеративной насадки, слоем металлических стенок ограждения можно пренебречь, так как теплопроводность металла достаточно высокая и толщина металлических стенок несравнимо мала при сравнении с теплоизоляционным слоем.

Расчет производится для i количества слоев. Индекс шага по времени k .

Количество тепла, передаваемое от горячего теплоносителя теплоаккумулирующим шарикам, рассчитывается по формуле:

$$\Delta Q_{i,k+1} = \alpha_{v_{i,k}} \cdot \Delta h \cdot f \cdot \Delta \tau \left(U_{i,k}^D - t_{i,k} \right) - \frac{2\pi (t_{i,k} - t_{окр})}{\frac{1}{\alpha_{v_{i,k}} r_1} + \frac{1}{\lambda} \ln \frac{r_2}{r_1} + \frac{1}{\alpha_{окр} r_2}}, \text{Дж}.$$

Количество тепла, которое шарики передают холодному теплоносителю, рассчитывается по формуле:

$$\Delta Q_{i,k+1} = \alpha_{v_{i,k}} \cdot \Delta h \cdot f \cdot \Delta \tau \left(t_{i,k} - U_{i,k}^B \right) - \frac{2\pi (t_{i,k} - t_{окр})}{\frac{1}{\alpha_{v_{i,k}} r_1} + \frac{1}{\lambda} \ln \frac{r_2}{r_1} + \frac{1}{\alpha_{окр} r_2}}, \text{Дж}.$$

где $\alpha_{v_{i,k}}$ – объемный коэффициент конвективной теплоотдачи в слое шариков. Определяется по эмпирической формуле Китаева-Фурнаса, Вт/(м²·К):

$$\alpha_{v_{i,k}} = A \frac{w_0^{0,9} (U_{i,k} + 273)^{0,3}}{d^{0,75}} \cdot 10^{1,68\omega - 3,56\omega^2}, \text{Вт} / (\text{м}^2 \cdot \text{К}),$$

где A – экспериментально определяемый коэффициент, зависящий от материала шариков $\cong 100 \dots 200$

f – поперечное сечение шариковой насадки, м²

w_0, V_0 – скорость и расход теплоносителя при НФУ, м/с, м³/с;

ω – порозность слоя засыпки;

$U_{i,k}$ – температура теплоносителя на i -том шаге по координате и k -том шаге по времени °С;

$t_{i,k}$ – температура теплоаккумулятора (шариков засыпки насадки) на i -том шаге по координате и k -том шаге по времени, °С;

$t_{окр}$ – температура окружающей среды, °С;

$\alpha_{окр}$ – коэффициент теплоотдачи окружающей среды Вт/(м²·К);

λ – коэффициент теплопроводности теплоизоляционного материала ограждений насадки Вт/(м·К);

r_1, r_2 – внутренний и внешний радиусы слоя теплоизоляции насадки, м.

Количество теплоты теплоносителя, передаваемого в следующий элементарный слой теплоаккумулятора.

$$Q_{i+1,k} = Q_{i,k} - \Delta Q_{i,k}, \text{Дж}.$$

Передача тепла от горячего теплоносителя (дыма) к теплоаккумулирующим шарикам приводит к повышению температуры шариков и охлаждению теплоносителя. Расчет температур теплоносителя производится по формуле:

$$U_{i,k}^D = U_{i,k-1}^D - \frac{\Delta Q_{i,k}}{V_0 \cdot c_{дым} \frac{1}{\varphi}}, \text{°С},$$

где $U_{i,k}^D$ – температура дыма, °С;

$c_{дым}$ – теплоемкость дыма, Дж/(м³·К).

Передача тепла от теплоаккумулирующих шариков к теплоносителю (воздуху) приводит к повышению температуры теплоносителя и охлаждению шариков. Расчет температур теплоносителя производится по формуле:

$$U_{i,k}^B = U_{i,k-1}^B - \frac{\Delta Q_{i,k}}{V_0 \cdot c_B \frac{1}{\varphi}}, \text{°С},$$

$U_{i,k}^B$ – температура воздуха, °С;

$c_{Bдым}$ – теплоемкость воздуха, Дж/(м³·К).

Исследование термической массивности шариков производится путем решения дифференциального уравнения нестационарной теплопроводности в одномерной постановке:

$$\frac{\partial t}{\partial \tau} = \frac{\lambda}{\rho \cdot c_{эф}} \left(\frac{\partial^2 t}{\partial r^2} + \frac{2}{r} \cdot \frac{\partial t}{\partial r} \right).$$

В математической модели реализовано решение численным методом по явной схеме. Для точки в центре шарика принимаются граничные условия второго рода $\frac{\partial t}{\partial r} = 0$.

Для точки на поверхности шарика задаются граничные условия третьего рода [7]:

$$-\lambda \frac{\partial t}{\partial r} = \alpha (t_{cp} - t_n).$$

Для расчета шарик условно разделили на слои, n контрольных точек находятся на расстоянии $\frac{r}{n}$ м друг от друга. Расчет по толщине шарика производился по аппроксимированным формулам:

- для поверхности шарика:

$$t_{i,k}^1 = U_{i,k}^{cp} - \left(U_{i,k}^{cp} - t_{i,k-1}^1 \right) e^{-\frac{\Delta \tau \lambda \alpha}{r_0 \rho \Delta \tau}}, \text{°С};$$

- для точек от 2 до $n-1$:

$$t_{i,k}^j = t_{i,k-1}^j + \frac{\lambda \Delta \tau}{\rho c} \left(\frac{t_{i,k-1}^{j+1} - 2t_{i,k-1}^j + t_{i,k-1}^{j-1}}{\Delta r^2} + \frac{2}{r} \cdot \frac{t_{i,k-1}^{j+1} - t_{i,k-1}^j}{\Delta r} \right), \text{°С};$$

- температуру центра считаем равной температуре предыдущей точки $n - 1$:

$$t_{i,k}^n = t_{i,k}^{n-1}, \text{°С}.$$

После расчета температур шарика в контрольных точках производится расчет среднemasсовой температуры шарика, считаем, что температура условно выделенного i -того слоя одинакова на всем объеме слоя.

Проведен расчет для исходных данных: начальная температура дыма $U_n=1000^\circ\text{C}$; начальная температура насадки $t_n=20^\circ\text{C}$; высота насадки $H=0,223$ м; диаметр шариков $d=0,02$ м; по высоте насадки выделили 5 слоев, с одинаковым шагом по высоте; шаг по времени $\Delta\tau = 1$ с.

Полученные результаты представлены ниже на рисунках 2-5.

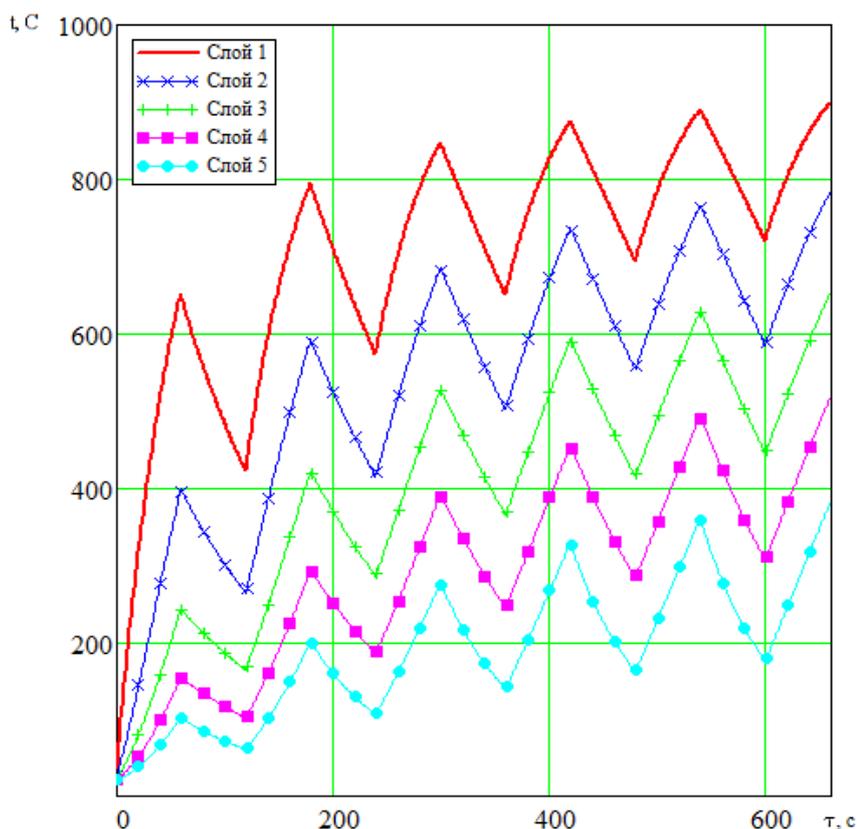


Рис. 2. График распределения температур по толщине насадки без учета термомассивности

В процессе выполнения исследования были сделаны расчеты температурного поля насадки с учетом потерь через ограждающие конструкции и без учета потерь. Основываясь на полученных данных, можно сделать вывод, что при использовании современных теплоизоляционных материалов разница, получаемая при расчетах, незначительна, при иллюстрировании результатов виде графиков не заметна, хотя имеет место в численных значениях.

Было проведено исследование термомассивности наполнителя регенеративной шариковой насадки и соответственно построены графики распределения температур в разных условно выделенных слоях насадки. На графиках представлены условно выделенный первый и последний слой насадки.

При расчете мы использовали 10 контрольных точек, равномерно распределенных по радиусу шарика. Такое количество точек, и принятое допущение, что на второй половине шарика распределение температур будет симметрично, мы считаем достаточно достоверным. На графиках отображены только 5 контрольных точек для придания наглядности графику.

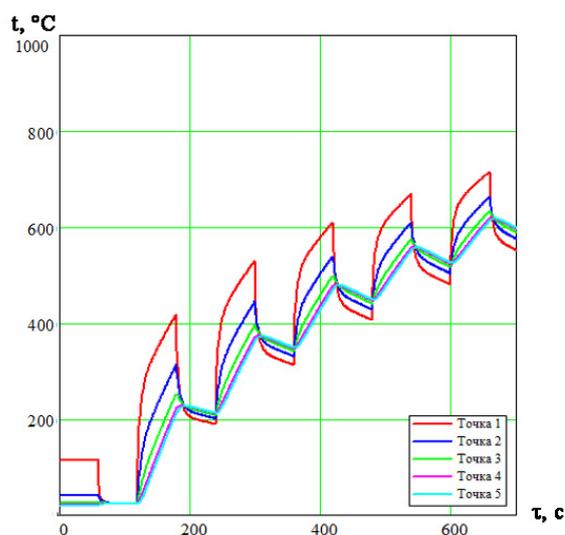


Рис. 3. Распределение температур по сечению шарика, 1 слоя насадки

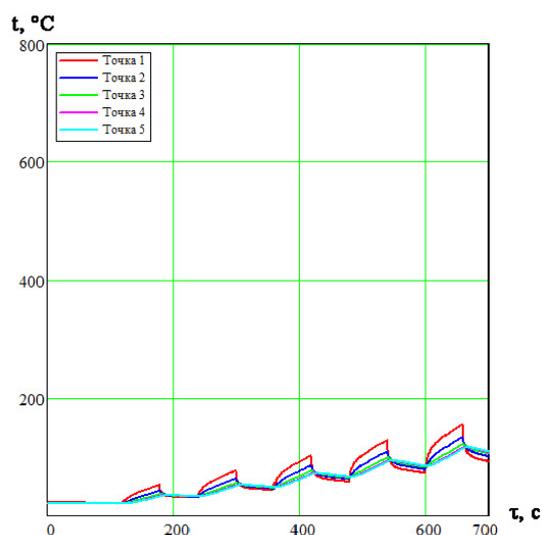


Рис. 4. Распределение температур по сечению шарика, 5 слоя насадки

Рассчитанные среднемассовые температуры для условно выделенных слоев микрорегенеративной насадки.

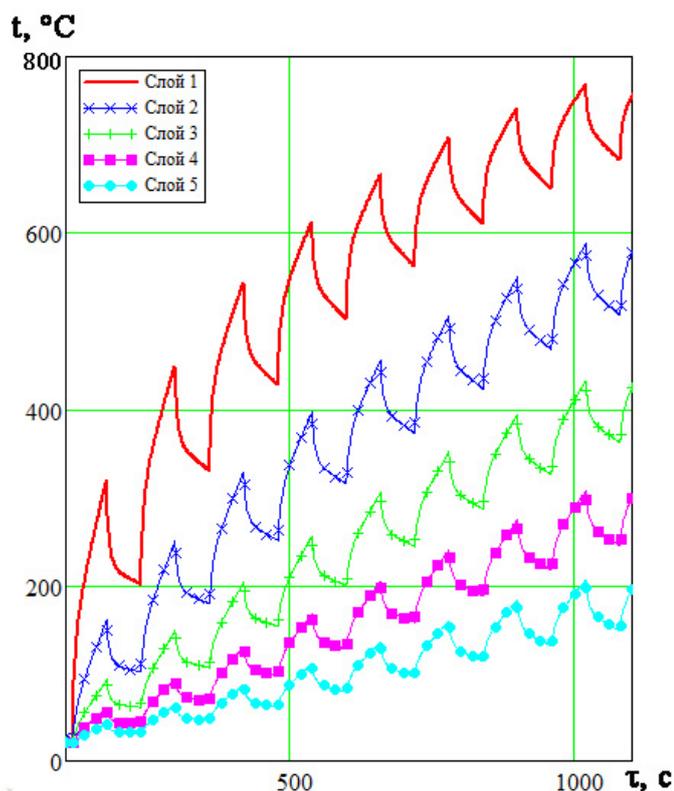


Рис. 5. График распределения среднемассовых температур по толщине насадки с учетом термомассивности

Сравнив температуры, получаемые при расчете температурного поля без учета термомассивности и температурное поле, полученное по среднемассовым значениям температуры с учетом термомассивности, видно:

- при рассмотрении графика распределения температур без учета термомассивности, изменение температуры происходит плавно и равномерно, как в случае нагрева, так и в случае охлаждения;
- при рассмотрении графика распределения температур учитывающих термомассивности, изменение температуры при нагреве происходит плавно, а при охлаждении сначала происходит достаточно резкое падение температуры шариков, однако примерно на середине цикла оно замедляется;
- при сравнении графиков на рисунках 2 и 5 стоит отметить, что нагрев теплоаккумулирующей насадки с учетом термомассивности происходит до температуры ниже по значению чем без учета;
- мы считаем, что картина с учетом термомассивности в большей степени соответствует реальности и отображает процессы приближенные к реальности.

Выводы. Проанализировав полученные в результате моделирования данные и сравнив их с данными, полученными на производстве [3, 4] можно сделать выводы:

1. Разработанная математическая модель качественно правильно описывает исследуемый процесс.
2. При применении современных теплоизоляционных материалов доля потерь в окружающую среду через ограждающие стенки теплообменной насадки не является существенной.
3. Проведённые исследования термомассивности теплоаккумулирующих шариков позволяют получить поле температур по толщине шарика.
4. Получаемая среднемассовая температура наиболее приближена к реальности. Эти данные позволяют спрогнозировать термические напряжения в теплоаккумулирующем материале.
5. Использование разработанной математической модели может быть применено при диагностировании состояния работающих микрорегенеративных горелок, а так же при проектировании новых.
6. Математическая модель, описанная в этой статье и результаты моделирования являются началом исследования, целью которого создать инженерную методику расчета шариковых микрорегенеративных насадок регенеративных горелок.

СПИСОК ИСТОЧНИКОВ

1. Бирюков, А. Б. Анализ теплотехнической эффективности печей, отапливаемых регенеративными горелками / А. Б. Бирюков, А. Н. Лебедев, П. А. Гнигитев // Сталь. – 2018. – № 3. – С. 64-67.
2. Бирюков, А. Б. Методика определения основных параметров теплообменной насадки регенеративных горелок / А. Б. Бирюков // Сталь. – 2018. – № 11. – С. 72-75.
3. Сысоева, Т. Е. Конечно-разностное решение задачи нагрева (охлаждения) неподвижного пористого слоя материала на основе уточненной физической модели движения газа / Т. Е. Сысоева, Ю. Я. Абраменков // Металлургическая теплотехника: Сб. науч. тр. НМетАУ. – Днепропетровск, 2008. – С. 272-286.
4. Reference Document on Best Available Techniques for Energy Efficiency [Электронный ресурс]. – URL: <https://eippcb.jrc.ec.europa.eu/reference/energy-efficiency> (дата обращения: 07.06.2024).
5. Бирюков, А. Б. Оценка технико-экономической целесообразности использования рекуперативных или регенеративных горелок для реконструкции печей / А. Б. Бирюков // Сталь. – 2019. – № 6. – С. 71-73.
6. Начкебия, Н. С. Анализ расчетных методов и математических моделей теплообменных процессов в насадках регенеративных горелок / Н. С. Начкебия, А. Б. Бирюков // Вестник Донецкого национального университета. Серия Г: Технические науки. – 2023. – № 1. – С. 50-56. – EDN USIRWS.
7. Бирюков, А. Б. Энергоэффективность и качество тепловой обработки материалов в печах: монография / А. Б. Бирюков. – Донецк: Ноулидж, 2012. – 247 с.

Поступила в редакцию 02.10.2024 г., рекомендована к печати 28.10.2024 г.

MODELING THE TEMPERATURE FIELD OF NOZZLES OF REGENERATIVE BURNERS

Nachkebiia N.S., Biriukov A.B., Aslamova IA.IU.

The paper presents a mathematical model for calculating the temperature fields of a ball regenerative nozzle of a regenerative burner. The model is based on elementary balances of conditional packing layers. As a result of the calculations, the temperature distributions in the heat storage layer of the packing, hot and cold coolants were obtained. The data obtained correspond to practical experience in the use of regenerative burners and can be considered reliable.

Keywords: energy efficiency, regenerative burner, heat exchange, heat balance, regenerative nozzle, calculation method.

Начкебия Наталья Сергеевна

аспирант кафедры технической теплофизики
ФГБОУ ВО «Донецкий национальный
технический университет»,
Российская Федерация, ДНР, г. Донецк,
E-mail: moon-ananaska@yandex.ru

Nachkebiia Natalia Sergeevna

Postgraduate Student at Department of Technical
Thermophysics of Donetsk National Technical
University,
Russian Federation, DPR, Donetsk.

Бирюков Алексей Борисович

доктор технических наук, профессор, заведующий
кафедрой технической теплофизики, проректор
ФГБОУ ВО «Донецкий национальный
технический университет»,
Российская Федерация, ДНР, г. Донецк.
E-mail: birukov.ttf@gmail.com

Biriukov Aleksei Borisovich

Doctor of Technical Sciences, Professor, Head of the
Department of Technical Thermophysics, Vice-rector
of Donetsk National Technical University,
Russian Federation, DPR, Donetsk.

Асламова Яна Юрьевна

кандидат технических наук, доцент кафедры
руднотермических процессов и малоотходных
технологий ФГБОУ ВО Донецкий национальный
технический университет»,
Российская Федерация, ДНР, г. Донецк,
E-mail: aslamova_yana@mail.ru

Aslamova IAna IUrevna

Candidate of Technical Sciences, Associate Professor
at Department of Ore-Thermal Processes and Low-
Waste Technologies of Donetsk National Technical
University,
Russian Federation, DPR, Donetsk.

ПЕРСПЕКТИВЫ ПРИМЕНЕНИЯ СОВРЕМЕННЫХ СИЛОВЫХ ЭЛЕКТРОННЫХ КОМПОНЕНТОВ В МОДЕРНИЗАЦИИ СТАНОЧНОГО ПАРКА РОССИЙСКОЙ ФЕДЕРАЦИИ

© 2024 *Городько В.В., Городько О.Ю., Гридин С.В.*

В статье рассматривается вектор развития силовой электроники, приведены сравнительные технические характеристики тиристорных ГТО, ГСТ и транзисторных IGBT.

Ключевые слова: электропривод постоянного тока, система управления, тиристор, силовой транзистор.

Введение. В настоящее время электроэнергия является самым удобным, универсальным видом энергии, ее можно генерировать, хранить, передавать, потреблять. Любое преобразование электрической энергии в механическую энергию вращения производится посредством электропривода, который является «сердцем» всех производственных процессов.

Требования к электроприводам и системам управления станками определяются технологией обработки, конструктивными возможностями станка и режущего инструмента. Основными технологическими требованиями являются обеспечение:

- самого широкого круга технологических режимов обработки с использованием современного режущего инструмента;

- максимальной производительности;

- наибольшей точности обработки;

- высокой чистоты обрабатываемой поверхности;

- высокой степени повторяемости (стабильности) размеров деталей в обрабатываемой партии.

Основные требования к станочному электроприводу:

- минимальные габариты электрических двигателей при высоком вращающем моменте;

- высокая максимальная скорость;

- значительная нагрузочная способность в режиме кратковременной и повторной кратковременной нагрузки;

- широкий диапазон регулирования;

- высокая стабильность характеристик и, в первую очередь, усилителей и тахогенератора;

- высокое быстродействие при апериодическом характере переходных процессов разгона и торможения;

- высокое быстродействие при наброске и сбросе нагрузки и при реверсе под нагрузкой на самых малых частотах вращения;

- равномерность движения при различной нагрузке на всех скоростях, вплоть до самых малых;

- надежность и ремонтпригодность;

- удобство конструктивной установки двигателя на станке и встройки преобразователей в шкафы и ниши станков;

- малые габаритные размеры и расход активных материалов;
- небольшой расход дефицитных материалов;
- простота наладки, ремонта и эксплуатации;
- высокая унификация узлов отдельных элементов;
- высокая экономичность и малая стоимость [1].

Удовлетворение всем этим и другим требованиям зависит от характеристик станка и режущего инструмента, мощности главного привода и электромеханических свойств приводов подачи и систем управления.

На сегодняшний день в связи с моральным и физическим износом требуется модернизация электроприводов станочного парка бывшего СССР. Необходим переход электроприводов постоянного тока от тиристорных преобразователей с аналоговым управлением на шоррег схемы и инверторы с IGBT транзисторами с соответствующим цифровым управлением.

Вектор развития силовой электроники. В 1960-70х годах основным регулируемым электроприводом в мире был преобразователь на базе тиристоров SCR, питающий двигатель постоянного тока.

Бурное развитие в 70-х, 80-х годах прошлого века цифровой микроэлектроники в странах Запада позволило увеличить на порядки вычислительную мощность новых цифровых систем управления по сравнению с аналоговыми системами, с приемлемым быстродействием и точностью, однако время реакции тиристора на сигнал управления так и оставалось зависимым от частоты сети, перехода синусоиды через ноль.

Появление за пределами СССР силового MOSFET транзистора (1978, International Rectifier Corporation) [2, 3] и IGBT транзистора (1979, General Electric) [4] привело к кардинальному изменению системы управления, на смену аналоговому управлению мощностью (управлению фазой) для тиристоров SCR пришло цифровое «бесшовное» управление мощностью, широтно-импульсное регулирование PWM для IGBT и MOSFET. В начале 90-х годов внедрение IGBT в промышленные электроприводы приобрело массовый характер. Это послужило отправной точкой к переходу от тотального распространения регулируемых электроприводов постоянного тока к регулируемым приводам как постоянного тока, на основе DC шоррег схем, так и переменного тока, инверторов – FVD VSI и FVD CSI электроприводов. В странах Запада произошло цифровое технологическое переформатирование промышленности и станкостроения.

В Советском Союзе в 70-х годах также проводились работы по синтезу МДП-транзистора с мощным биполярным транзистором, которые увенчались успехом. В 1980 г. было выдано авторское свидетельство на новый тип транзистора (аналога IGBT) – побистор [5] и начато серийное производство системы числового программного управления (СПЧУ) «Электроника-НЦ31», аналог СПЧУ компании FANUC, мирового лидера в области систем промышленной автоматизации. В стране с 1980 по 1990 гг. ежегодно производилось 150-250 тыс. металлорежущих станков [6].

Доля станков с ЧПУ от общего количества произведенных металлорежущих станков в РСФСР составляла 23% в конце 80-х годов [7]. Модернизация рассчитана именно на станки этих лет выпуска, вероятно имеющих меньший физический износ. Трагический распад СССР в начале 90-х годов и последующий экономический кризис не дал перейти отечественному станкостроению на новый технологический уровень, поэтому новейшие разработки в области силовой электроники оказались невостребованными. Конкурентоспособность страны в области станкостроения в мире была потеряна.

В 2018 г. мировой рынок силовых транзисторов составлял 16,3 млрд. долларов [8], в то же время Россия ежегодно импортировала силовые транзисторы на 50 млн. долларов США [9], т.е. Россия занимала менее 1% мирового рынка. В настоящее время ситуация с российским рынком силовых электронных компонентов несильно изменилась, т.к. на сегодняшний день 90% силовых транзисторов в России продолжает поставляться из-за рубежа [10]. Это говорит о недостаточном внимании в стране к этому очень перспективному направлению в области электроники, позволяющему усилить конкурентоспособность товаров, как в гражданской сфере, так и продукции военного назначения.

Необходимость усовершенствования данного вида электронной продукции усилилась с появлением в коммерческой продаже первой перезаряжаемой литий-ионной батареи (1991 г., Sony). С этого момента понятия «аккумуляторы» и «преобразователи» стали неразделимы.

Постоянное расширение диапазона технических характеристик силовых транзисторов, появление на коммерческом рынке, наряду с транзисторами на кремниевой основе (Si) транзисторов на основе карбида кремния (SiC) и нитрида галлия (GaN) (рис.1) способствует совершенствованию таких продуктов и систем, как:

- беспилотные системы;
- беспроводные системы связи;
- электромобили и ЖД транспорт;
- летательные аппараты (самолеты, спутники и др.);
- инверторы и мощные металлургические электроприводы;
- промышленные и коллаборативные роботы;
- системы передачи электроэнергии HVDC и FACT.

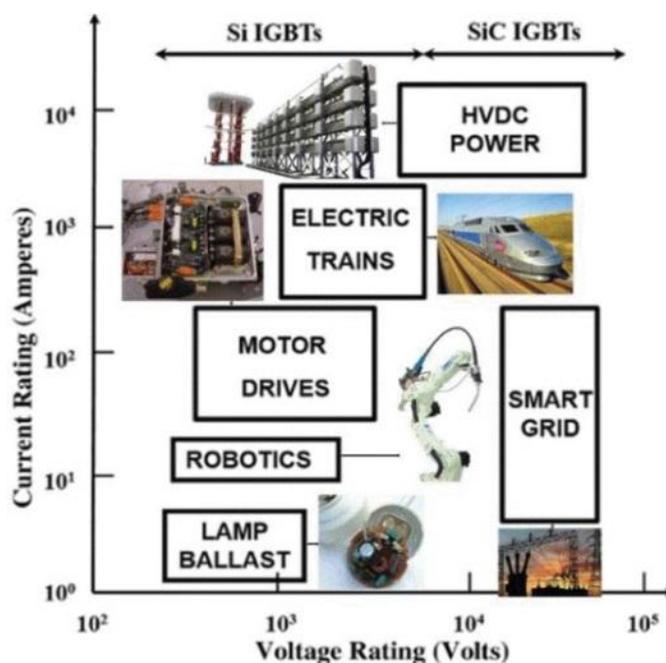


Рис. 1. Области применения силовых транзисторов [11]

В настоящее время драйверами развития силовой электроники является энергетика и автомобилестроение. Разработкой и производством данного вида электронной продукции занимаются компании ведущих технологически развитых стран мира [12, 13] (таблица 1):

- Microchip, Texas Instruments, Onsemi, Mag (США);
- Mitsubishi Electric, FANUC, Yamazaki Mazak, Okuma, Fuji Electric, Makino, Renesas Electronics (Япония);
- NUM (Франция);
- ST Microelectronics (Швейцария);
- Infineon Technologies AG, EMAG (Германия);
- NXP (Нидерланды);
- SMTCL, DMTG, QCMТТ, BOOH, Ningbo Hision Precision Industry Co. (Китай) и др.

Таблица 1. Ассортимент продукции ведущих мировых производителей станков с ЧПУ

Название фирмы-производителя	Ассортимент продукции
FANUC (Япония)	Системы ЧПУ для лазерных станков различной мощности, начиная от компактных настольных моделей и заканчивая промышленными обрабатывающими центрами, системы ЧПУ для широкого спектра промышленных машин (пресс-машины, штамповочные машины, намоточные машины и др.), сервоприводы с двигателями переменного тока с большим крутящим моментом до 18000 Н·м и мощностями до 500 кВт, шпиндельные двигатели переменного тока с номинальным моментом до 2000 Н·м и мощностями до 200 кВт.
Yamazaki Mazak (Япония)	Многоцелевые станки (от 7,5 до 22 кВт), 5-тиосевые обрабатывающие центры (от 15 до 22 кВт), токарные центры с ЧПУ (от 7,5 до 15 кВт), вертикальные и горизонтальные обрабатывающие центры (от 22 до 45 кВт), лазерные обрабатывающие станки (от 2 до 15 кВт)
EMAG (Германия)	Токарные станки (от 14 до 38 кВт), шлифовальные станки (от 1,1 до 15 кВт), обрабатывающие центры и фрезерные станки (до 15 кВт), зуборезные станки, обрабатывающие центры для соединительных втулок, станки для лазерной очистки
NUM Франция)	Сервоприводы, системы управления с ЧПУ, серводвигатели от 0,318 до 160 Н·м, шпиндельные двигатели мощностью до 55 кВт, специальные двигатели для комплектов (шпиндельные двигатели с жидкостным охлаждением, серводвигатели с жидкостным охлаждением, асинхронные и синхронные шпиндельные двигатели (мотор-шпиндель))

Развитие силовых транзисторов постоянно идет по двум направлениям:

- увеличение плотности мощности, т.е. передачи максимальной мощности при минимальном объёме «ключа»;
- увеличение эффективности преобразования энергии, т.е. уменьшение электрических потерь в установившемся и переходных режимах коммутации, ведущих к нагреву транзистора.

Уже создано семь поколений IGBT транзисторов различных типоразмеров, уровней коммутируемых мощностей и назначения.

Преимущество IGBT транзисторов перед тиристорами:

- принудительная коммутация (SCR с естественной коммутацией);
- более высокая скорость переключения (диапазон рабочих частот до 25 кГц);
- более низкое напряжение насыщения (таблица 2);
- простая и малая мощность схемы управления;
- меньшее время переключения (таблица 2);
- меньшие требования к снабберным цепям;
- мощность сигнала, коммутируемая транзисторами, сопоставима с мощностью сигнала для тиристорov (таблица 1);
- значительно меньшие габариты.

Недостаток IGBT транзисторов – меньшая стойкость к токам короткого замыкания по сравнению с тиристорами (табл. 2). В итоге, применение в силовой преобразовательной технике IGBT вместо тиристорных дало:

- значительное снижение массы и габаритов оборудования;
- повышение точности обработки сигнала управления;
- создание совместно с аккумуляторами систем с различной степенью автономности.

Таблица 2. Сравнение технических характеристик IGBT, GTO, GCT производства Mitsubishi Electric

Тип	Ипр.	Вобр.	Vsat	Икз.	Икз./Ипр.	ткз.	твкл.	тоткл.
	кА	кВ	В	кА		мкс	мкс	мкс
Тиристоры GTO								
FG6000AU-120D	6	6	6	40	6,5	8300	10	30
FG1000BV-90DA	1	4,5	4	8,4	8,4	8300	10	20
Тиристоры GCT								
FGC4000BX-90DS	4	4,5	6	25	5,5	8300	3	3
Тиристоры IGBT								
CM1000HG-130XA	1	6,5	3,25	4,8	4,8	10	1,2	7,7
CM2400HC-34N	2,4	1,7	2,4	14	5,8	10	1,5	3

Однако разработчики до сих пор проявляют интерес к тиристорам. Сыграла главную роль в модернизации силовой электроники в 60-х годах, тиристоры теперь применяются в основном только в области высоких напряжений и токов.

На данный момент, наряду с SCR, выпускаются запираемые IGCT, GCT, GTO с рабочей частотой не более 200-1000Гц [14].

Сравнительный анализ силовых электронных компонентов даже на базе кремния, выпускаемых компаниями Mitsubishi Electric (Япония), АО «Ангстрем» и АО «Протон-Электротекс» (Россия) в 2024 г (рис. 2, рис. 3) показывает отставание России в разработке и коммерческом выпуске данного вида продукции. Сохраняется негативная зависимость от зарубежных технологий в области электроники.

Современные системы управления электроприводами делятся на две составляющие:

- системы управления циклами,
- системы управления случайностями.

Системы управления циклами основаны на ПИД регулировании, когда параметры системы определены и изменяются по определенным законам, и любое несанкционированное «выпадение» из цикла воспринимается системой как аварийная ситуация. Данная система реализуется в аппаратном и программном сегментах в робототехнике, металлургии, станкостроении и др.

Системы управления случайностями основаны на обработке данных и обучении для выработки оптимального управления, основанного на прогнозировании нелинейных процессов и случайностей. Эти передовые методы управления включают в себя:

- управление на основе искусственного интеллекта (ИИ);
- модельно-прогностическое управление (МРС);
- управление скользящим режимом на основе отслеживания N-infinity;
- адаптивное управление.

Данные методы реализуются в программном обеспечении систем управления электроприводами, в транспортной и военной сферах.

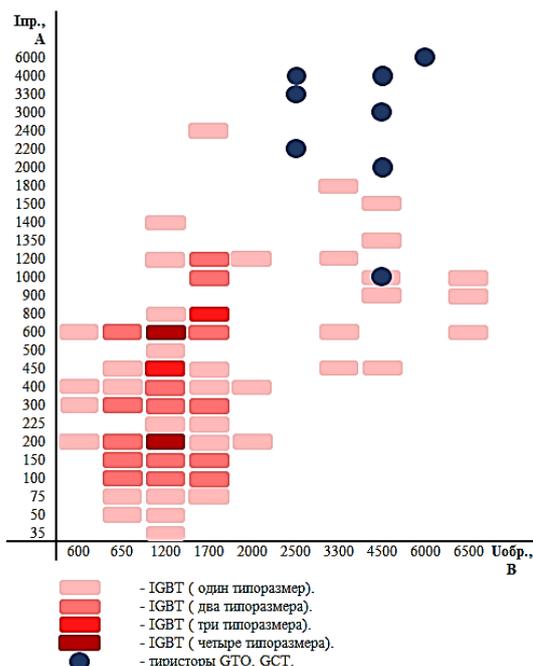


Рис. 2. $U_{обр}(I_{пр})$ IGBT, GTO, GCT производства Mitsubishi Electric в 2024 г.

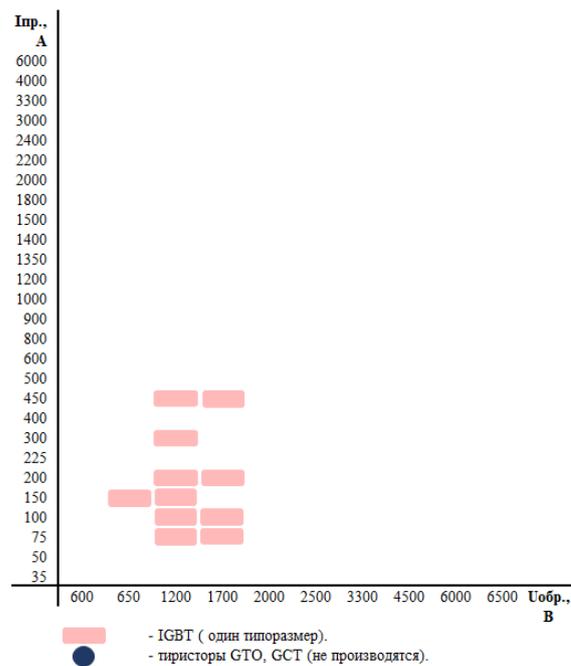


Рис. 3. $U_{обр}(I_{пр})$ IGBT, GTO, GCT производства АО «Ангстрем» и АО «Протон-Электротекс» в 2024 г.

Результаты исследования. В ГБУ «НИИВЭ» в рамках выполнения проекта «Разработка микроконтроллерных платформ для модернизации систем металлорежущего оборудования» идет испытание собственного драйвера управления IGBT модулями.

В основе проекта лежат положения о том, что технический уровень машиностроительного оборудования определяется следующими основными показателями: гибкостью, производительностью, точностью, надежностью, удельной металлоемкостью и удельным энергопотреблением. Развитие в машиностроении прогрессивных технических средств, обеспечивающих значительное сокращение вмешательства обслуживающего персонала в процесс функционирования оборудования, требует создания и внедрения принципиально новых систем автоматизированного электропривода. Современные электроприводы базируются на широком применении специализированных электрических машин, силовой полупроводниковой техники, средств микроэлектроники.

За последние годы произошли качественные изменения в номенклатуре и техническом уровне электроприводов, применяемых в станкостроении. Зарубежные компании производят и внедряют в машиностроительном оборудовании качественно новые электроприводы, построенные на базе серводвигателей постоянного и переменного тока мощностью до 100 кВт. Они компактны, содержат в своём корпусе двигатель, датчики, редуктор и тормоз. Однако их стоимость очень высока. Мощности модельного ряда российских серводвигателей не может покрыть нужды отечественного станочного парка.

Электропривод, состоящий из отдельных узлов и элементов, позволяет проводить ремонт и последующую модернизацию в условиях ограниченного финансирования. Поэтому на данный момент, в условиях санкционного давления, является актуальной разработка нового поколения преобразовательной техники для поддержания конкурентно способности существующего станочного парка.

США, Евросоюз, Япония, Китай, стремятся первыми войти и стать лидерами в следующей индустриальной эпохе – Индустрии 6.0, которая будет характеризоваться:

- автономным производством, созданием интеллектуальных, интегрируемых заводов с минимальным вмешательством человека;
- ускорением разработок в области высоких технологий, основанных на автономных лабораториях;
- системами идентификации в режиме реального времени с использованием больших баз данных и искусственного интеллекта;
- применением роботов с прямым взаимодействием с человеком, наряду с роботами, с запрещенной рабочей зоной, применяемых в промышленности.

Одним из основных драйверов развития новой технологической революции будет повсеместное внедрение автономности, ИИ и робототехники на базе тотальной зависимости от источников электроэнергии.

Все эти новшества будут касаться в первую очередь военных технологий, а не только гражданской сферы жизни человека.

Уже сейчас беспилотные летательные аппараты (БПЛА) «дроны», применяемые в зоне проведения СВО, в подавляющем большинстве включают в себя автономный электропривод на MOSFET транзисторах с PWM регулированием.

Также, в связи широкой цифровизацией наземных боевых платформ, применением станций РЭБ, систем управления боем, радаров, камер, зарядных станций для дронов растёт потребность в бортовой электроэнергии.

Решить эту задачу возможно за счет внедрения гибридного электропривода (HEV) на боевые машины, который увеличивает электрический запас хода боевой машины и обеспечивает ей свойства малозаметности в режиме готовности при останове и движения при наблюдении, путем значительного снижения акустического и инфракрасного излучений. Дополнительно силовой агрегат также может «экспортировать» электроэнергию, что означает, что он может работать как генератор для различных задач, которые могут возникнуть как в боевых действиях, так и в гуманитарных миссиях. Экономия потребления топлива ведёт к сокращению количества пополнения топливом, а значит, к увеличению безопасности логистики топлива к передовым позициям в зоне боевых действий.

В настоящее время проектами применения электропривода HEV в наземных боевых машинах уже занимаются такие компании как BAE system для Bradley (США), GM для Hummer, Striker и Abrams (США) [15, 16].

Принцип построения схемотехнических решений тяговых инверторов HEV в основном аналогичен образцам, применяемым в промышленности и станкостроении, инверторы комплектуются, в зависимости от напряжения АКБ, силовыми IGBT или MOSFET SiC.

Эти данные показывают важность развития электроприводов нового поколения не только в станкостроении и других гражданских отраслях, но и перспективы их применения в военной сфере [17].

Происходит усиление конкуренции между странами в области создания и развития новых технологий. «Россия – уникальная цивилизация», как справедливо отметил президент Путин В.В., поэтому она обязана принять вызов для сохранения своего суверенитета путём восстановления технологической независимости и, как следствие, военного паритета со странами Запада.

Выводы. Таким образом, решить проблему восстановления технологического паритета в области электроприводов и станкостроения с вышеуказанными странами возможно, работая в двух направлениях одновременно:

- развивая системы управления циклами;
- развивая системы управления случайностями.

В целом, получение компетенций в области управления IGBT и создание надежного силового модуля, драйвера совместно с IGBT модулем позволит проектировать электроприводы как переменного, так и постоянного тока на новой элементной базе.

СПИСОК ИСТОЧНИКОВ

1. Модернизация электропривода подачи станка 16K20T1 [Электронный ресурс] / Библиофонд : [Сайт] - URL: <https://www.bibliofond.ru/view.aspx?id=669606> (дата обращения 10.09.2024).
2. The industrial electronics handbook [Электронный ресурс]. – URL: https://archive.org/details/industrialelectr000unse_m6c8 (дата обращения 10.09.2024).
3. Силовые MOSFET: расширяет возможности [Электронный ресурс]. – URL: https://russianelectronics.ru/files/55167/EK2011_02_73-78.pdf?ysclid=m0usnga38y711572276 (дата обращения 10.09.2024).
4. Устройство IGBT: физика, конструкция и применение биполярного транзистора с изолированным затвором [Электронный ресурс]. – URL: <https://www.overdrive.com/media/9148684/the-igbt-device> (дата обращения 10.09.2024).
5. Дьяконов, В. Побистор или IGBT и имитационное моделирование устройств на них / В. Дьяконов // Силовая Электроника. – 2010. – №5. – С.24-32.
6. Народное хозяйство в СССР в 1990 г.: статистический ежегодник / Госкомстат СССР. – М.: Финансы и статистика, 1991. – 752 с.
7. Станковый хребет экономики. Как вернуть утраченные позиции на рынке станко- и подшипникоостроения [Электронный ресурс]. – URL: <https://www.sonar2050.org/storage/files/Доклады/Лизан/Станковый%20хребет%20экономики.pdf?ysclid=m0uuliuip0469230582> (дата обращения 10.09.2024).
8. Продажи силовых транзисторов по всему миру с 2008 по 2024 год [Электронный ресурс]. – URL: <https://www.statista.com/statistics/575536/worldwide-power-transistor-sales/> (дата обращения 10.09.2024).
9. Микроэлектроника: Обзор рынка силовых транзисторов от J&P [Электронный ресурс]. – URL: <http://www.mforum.ru/news/article/119678.htm?ysclid=lzv125wqa3786531092> (дата обращения 10.09.2024).
10. Силовая ЭКБ: линейка полностью отечественных IGBT-модулей [Электронный ресурс]. – URL: <https://power-e.ru/components/power-angstrem/> (дата обращения 10.09.2024).
11. Thermal Management on IGBT Power Electronic Devices and Modules [Электронный ресурс]. – URL: <https://ieeexplore.ieee.org/document/8258946> (дата обращения 10.09.2024).
12. Силовой «Элемент» [Электронный ресурс]. – URL: <https://www.kommersant.ru/doc/6596730> (дата обращения 10.09.2024).
13. Топ-10 производителей станков с ЧПУ в 2024 году [Электронный ресурс]. – URL: <https://artazono.com/ru/лучшие-производители-станков-с-чпу/?ysclid=m2virbg1eu72732497> (дата обращения 10.09.2024).
14. Gate turn-off thyristors (GTO) [Электронный ресурс]. – URL: <https://www.hitachienergy.com/products-and-solutions/semiconductors/gate-turn-off-thyristors-gto> (дата обращения 10.09.2024).
15. Electric Drives for Combat Vehicles [Электронный ресурс]. – URL: https://www.defstrat.com/magazine_articles/electric-drives-for-combat-vehicles/ (дата обращения 10.09.2024).
16. Next-Generation Abrams Tank & Stryker To Be Hybrids [Электронный ресурс]. – URL: <https://cleantechnica.com/2022/10/07/next-generation-abrams-tank-stryker-to-be-hybrids/> (дата обращения 25.09.2024).
17. Введенская, М. И. Интенсификация конструирования электродвигателей в среде САПР КОМПАС / М. И. Введенская, В. В. Варакута, С. В. Гридин // Вестник Донецкого национального университета. Серия Г: Технические науки. – 2024. – № 1. – С. 59-68. – DOI 10.5281/zenodo.10974640. – EDN VFWAQV.

Поступила в редакцию 08.11.2024 г., рекомендована к печати 22.11.2024 г.

**PROSPECTS FOR THE USE OF MODERN POWER ELECTRONIC COMPONENTS IN THE
MODERNIZATION OF THE MACHINE TOOL FLEET OF THE RUSSIAN FEDERATION**

Gorodko V.V., Gorodko O.IU., Gridin S.V.

The article considers the vector of development of power electronics is described, comparative technical characteristics of GTO, GCT thyristors and IGBT transistors.

Keywords: DC electric drive, control system, thyristor, power transistor.

Городько Виталий Викторович

ведущий инженер-конструктор ГБУ «Научно-исследовательский, проектно-конструкторский и технологический институт взрывозащищенного и рудничного электрооборудования»,
Российская Федерация, ДНР, г. Донецк.
E-mail: vitaliy.automat2010@ya.ru

Gorodko Vitalii Viktorovich

Leading Design Engineer of Scientific Research, Design and Technological Institute of Explosion-proof and Mining Electrical Equipment,
Russian Federation, DPR, Donetsk.

Городько Ольга Юрьевна

заведующий сектором ГБУ «Научно-исследовательский, проектно-конструкторский и технологический институт взрывозащищенного и рудничного электрооборудования»,
Российская Федерация, ДНР, г. Донецк.
E-mail: gorodolya1@rambler.ru

Gorodko Olga IUrevna

Head of the Sector of Scientific Research, Design and Technological Institute of Explosion-proof and Mining Electrical Equipment,
Russian Federation, DPR, Donetsk.

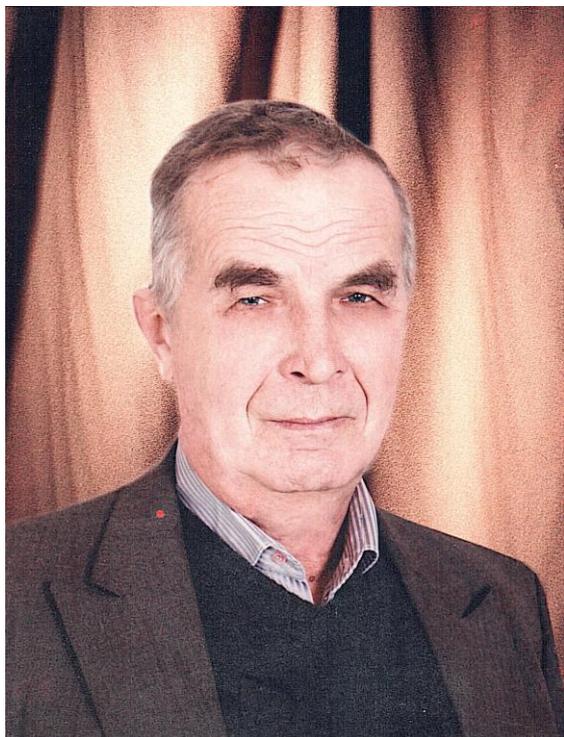
Гридин Сергей Васильевич

кандидат технических наук, доцент, доцент кафедры промышленной теплоэнергетики ФГБОУ ВО «Донецкий национальный технический университет»,
Российская Федерация, ДНР, г. Донецк.
E-mail: gridin_@mail.ru

Gridin Sergei Vasilevich

Candidate of Technical Sciences, Associate Professor, Associate Professor at Department of Industrial Heat and Power Engineering of Donetsk National Technical University,
Russian Federation, DPR, Donetsk.

ПАМЯТИ ВЛАДИМИРА ИВАНОВИЧА ТИМЧЕНКО



14 ноября 2024 года скоропостижно скончался член редакционной коллегии журнала, доцент кафедры радиофизики и инфокоммуникационных технологий Донецкого государственного университета **Владимир Иванович Тимченко**.

Тимченко Владимир Иванович родился 31 июля 1946 года в с. Есеновичи Тверской области. Среднее образование получил в СШ №22 г. Донецка, которую закончил в 1964 г. С 1964 по

1966 учился в Днепропетровском государственном университете, затем перевелся на Физический факультет Донецкого государственного университета, который закончил в 1969 по специальности «Физика».

В 1969-1970 гг. **Тимченко В.И.** служил в рядах Советской Армии. После чего, с декабря 1970 г. принят на должность ассистента кафедры физики низких температур физического факультета ДонГУ. С 1973 г. обучался в аспирантуре по специальности «Физика твердого тела». С 1975 г. принят старшим преподавателем кафедры физики низких температур.

В 1981 г. **Владимир Иванович** защитил диссертацию на соискание ученой степени кандидата физико-математических наук по специальности «Физика полупроводников и диэлектриков». В 1986 г. избран на должность доцента кафедры физики твердого тела физического факультета ДонГУ, по которой в 1988 году **Владимиру Ивановичу** присваивается ученое звание доцента.

В 1988 г. **Тимченко В.И.** переводится в Макеевский инженерно-строительный институт, где занимает должность доцента кафедры физики, а с 1995 по 1998 гг. обучается в докторантуре. С 2000 г. переведен на кафедру электротехники и автоматики МИСИ на должность доцента с выполнением обязанностей заведующего кафедрой, на которой проработал более 17 лет.

С 2017 г. **Тимченко В.И.** переведен на должность доцента кафедры радиофизики и инфокоммуникационных технологий Донецкого национального университета, на которой и проработал до конца своих дней...

За период работы на кафедре РФиИКТ **Владимир Иванович** внес весомый вклад в развитие таких научных направлений как: диагностика и контроль свойств материалов методами акустической дефекто- и спектроскопии; взаимодействие упругих волн с нарушениями среды распространения; автоматизированные системы научных исследований параметров поверхностных акустических волн; основы измерений в экспериментальной и прикладной физике; оптические информационные технологии. Также **Владимир Иванович** стоял у истоков открытия «Серии Г» журнала «Вестник Донецкого национального университета» и регулярно участвовал в рецензировании статей.

Владимир Иванович пользовался большим авторитетом среди коллег и был примером для молодежи, был простым в общении и отзывчивым человеком, всегда умел поддержать в трудные моменты, оказать любую помощь, вселить оптимизм и замотивировать на новые свершения.

Выражаем глубокие соболезнования родным и близким, разделяем с ними боль и горечь невосполнимой утраты. Светлая память о **Владимире Ивановиче Тимченко** навсегда сохранится в сердцах коллег, друзей и близких.

Редакционная коллегия журнала и
коллектив кафедры РФиИКТ

ПРАВИЛА ДЛЯ АВТОРОВ

1. Редакция журнала «Вестник Донецкого национального университета. Серия Г: Технические науки» принимает к рассмотрению на предмет опубликования в журнале ранее не опубликованные и не представленные для одновременной публикации в других изданиях научные статьи по тематике журнала.
2. Авторы статей несут персональную ответственность за наличие элементов плагиата в текстах статей, за полноту и достоверность изложенных фактов и положений, за факт их публикации. Допускается публикация не более двух статей от одного автора в одном номере и наличие не более пяти авторов одной статьи. Один из авторов должен иметь ученую степень. Гонорар не выплачивается и плата за публикацию статей не взимается.
3. Авторами представляется рукопись статьи в одном экземпляре (на русском или английском языке), напечатанном с одной стороны листа бумаги формата А4 (экземпляр подписывается авторами). Страницы рукописи должны быть последовательно пронумерованы. Объем рукописи, как правило, от 5 до 15 страниц (для обзорной статьи – до 25), включая рисунки, таблицы, список источников, сведения об авторах. Вместе с рукописью на адрес электронной почты vestnik.g@mail.ru высылается электронный вариант с полным текстом статьи в формате doc или docx. При отсутствии возможности предоставления в редколлегию рукописи, на электронную почту редакции высылается полный текст статьи в формате pdf. Статьи, не соответствующие Правилам для авторов или тематике журнала, к рассмотрению не принимаются. Все статьи, принятые к рассмотрению, подвергаются рецензированию и в случае положительной рецензии – научному и техническому редактированию.
4. Решение о публикации, публикации после доработки/переработки или отклонении статьи принимается редакционной коллегией журнала. Если статья рекомендуется к публикации после небольшой авторской доработки, редакция направляет авторам статьи замечания (без указания рецензента). Статья, направленная авторам на доработку, должна быть возвращена в исправленном виде в максимально короткий срок. К доработанной статье необходимо предоставить письмо от авторов, содержащее ответы на все замечания и поясняющие все изменения, сделанные в статье. В случае отклонения статьи редакция направляет авторам рецензию (без указания рецензента) или выдержки из нее, либо аргументированное письмо редактора. Редакция не вступает в дискуссию с авторами отклоненных статей, за исключением случаев явного недоразумения.
5. Обязательные требования к оформлению статей:
 - **Размер бумаги:** А4.
 - **Ориентация:** книжная.
 - **Поля:** зеркальные, верхнее – 30 мм, нижнее – 30 мм, внутри – 30 мм, снаружи – 20 мм.
 - **Положение верхнего и нижнего колонтитула относительно края:** 2 см.
 - **Междустрочный интервал:** одинарный.
 - **Абзацный отступ:** 1 см.
 - **Основной текст статьи:** шрифт Times New Roman, размер 12 пт., с выравниванием по ширине, без автоматической расстановки переносов.
 - **Аннотация, список источников, сведения об авторах:** шрифт Times New Roman, размер 10 пт., с выравниванием по ширине, без автоматической расстановки переносов.
 - **Подписи, заглавие и содержание таблиц:** шрифт Times New Roman, размер 10 пт., с выравниванием по центру, без автоматической расстановки переносов.

6. Текст статьи должен соответствовать следующей структурной схеме:

- **Первая строка:** индекс УДК – выравнивание по левому краю, без абзацного отступа.
- **Через строку:** **НАЗВАНИЕ** статьи – полужирный, выравнивание по центру, прописными буквами, без переноса слов, без абзацного отступа.
- **Через строку:** копирайт © и год – полужирный, выравнивание по левому краю, без точки после года, через 3 пробела *фамилия(и) и инициалы автора(ов)* – полужирный, курсив, выравнивание по левому краю, без абзацного отступа.
- **Через две строки:** аннотация объемом до 100 слов, кратко отражающая предмет статьи, примененные методы исследований и основные результаты, полученные авторами.
- **На следующей строке:** словосочетание *Ключевые слова:* – полужирный, курсив. Далее приводятся ключевые слова количеством до 10 слов, отделяемых друг от друга запятой. Блок аннотация и ключевые слова рекомендуется отделять от основного текста с помощью инструмента «Границы».
- **Через две строки:** основной текст статьи, который содержит введение, основную часть и выводы. Во **введении** рекомендуется излагать постановку проблемы в общем виде и связь с важнейшими научными и практическими задачами, анализ последних исследований и публикаций, в которых начато решение данной проблемы и на которые опирается автор, выделение нерешенных ранее частей общей проблемы, которым посвящена статья, формулировку целей статьи. В **основной части** рекомендуется излагать основные материалы исследования с полным обоснованием полученных научных результатов. Основную часть рекомендуется структурировать на разделы с названиями, отражающими их содержание (например: постановка задачи; методы ее решения; анализ полученных результатов). В **выводах** рекомендуется кратко и четко подводить итоги основных результатов, полученных авторами и описывать перспективы дальнейших исследований в данном направлении.
- **Через строку:** словосочетание **СПИСОК ИСТОЧНИКОВ** размером 10 пт., полужирный, выравнивание по левому краю, прописными буквами, без абзацного отступа. Далее через строку с абзацным выступом 0,5 см приводится нумерованный перечень используемых источников общим списком в порядке упоминания в тексте на языке оригинала в соответствии с ГОСТ Р 7.0.100–2018 «Библиографическая запись. Библиографическое описание. Общие требования и правила составления». Ссылка на источник приводится в квадратных скобках. Разрешается ссылаться только на опубликованные работы. Необходимо наличие источников последних 5-10 лет по исследуемому вопросу.
- **Через строку:** надпись «*Поступила в редакцию xx.xx.20xx г., рекомендована к печати xx.xx.20xx г.*» размером 10 пт., курсив, выравнивание по правому краю.
- **Через строку:** приводится **НАЗВАНИЕ** статьи, *фамилия(и) и инициалы автора(ов)*, аннотация и ключевые слова (*Keywords*) на английском языке (перевод) размером 10 пт.
- **Через строку:** сведения об авторах размером 10 пт., в таблице с параметром «Нет границ», содержащей два столбца (для сведений на русском и английском языках):
- **Фамилия, имя, отчество;**
- ученая степень, ученое звание, должность, с указанием организации (в соответствии с уставом), которую представляет автор;
- страна, город;
- адрес электронной почты.

7. Образец оформления статьи:

УДК 681.5:621.377

ЭЛЕМЕНТЫ УСТРОЙСТВ ВЫЧИСЛИТЕЛЬНОЙ ТЕХНИКИ ДЛЯ АСНИ КОНТРОЛЯ РАДИООБСТАНОВКИ

© 2024 Третьяков И.А., Данилов В.В.

Приведены описания элементов устройств вычислительной техники для создания специализированных автоматизированных систем научных исследований контроля радиобстановки на основе эхо-эффекта. Приведен механизм формирования фононного эха.

Ключевые слова: АСНИ, радиобстановка, эхо-эффект, фононное эхо, обработка радиосигналов.

Введение. Важнейшим элементом современной концепции [1-3] ...

Механизм формирования сигналов. Электрическая компонента первого [4] ...

Выводы. Таким образом, рассмотренные элементы позволяют ...

СПИСОК ИСТОЧНИКОВ

1. Третьяков, И. А. Элементы устройств вычислительной техники для АСНИ контроля радиобстановки на основе эхо-эффекта / И. А. Третьяков, В. В. Данилов, С. В. Борщевский // Вестник Донецкого национального университета. Серия Г: Технические науки. – 2023. – № 1. – С. 34-39. – EDN CADMAM.
2. Разработка автономной автоматизированной системы непрерывного контроля состояния строительного объекта / В. Ю. Подлесный, В. И. Тимченко, И. А. Третьяков [и др.] // Вестник Донецкого национального университета. Серия Г: Технические науки. – 2023. – № 2. – С. 51-59. – EDN PHCERN.
3. Автоматизированные системы научных исследований [Электронный ресурс] / Studbooks : [Сайт]. – URL: <https://studbooks.net/1386372/> (дата обращения 25.01.2024).

Поступила в редакцию xx.xx.2024 г., рекомендована к печати xx.xx.2024 г.

ELEMENTS OF COMPUTING DEVICES FOR ASRS CONTROL OF RADIO PLACEMENT

Tretiakov I.A., Danilov V.V.

The descriptions of the elements of computer technology devices for the creation of specialized automated systems for scientific research of radio control based on the echo effect are given. The mechanism of phonon echo formation is given.

Keywords: ASRS, radio substitution, echo effect, phonon echo, radio signal processing.

Третьяков Игорь Александрович

кандидат технических наук, заместитель декана по научной работе, доцент кафедры радиопроизводства и инфокоммуникационных технологий ФГБОУ ВО «Донецкий государственный университет», Российская Федерация, ДНР, г. Донецк.
E-mail: i.tretiakov@mail.ru

Tretiakov Igor Aleksandrovich

Candidate of Technical Sciences, Deputy Dean for Research, Associate Professor at Department of Radiophysics and Infocommunication Technologies of Donetsk State University, Russian Federation, DPR, Donetsk.

Данилов Владимир Васильевич

доктор технических наук, профессор, заведующий кафедрой радиопроизводства и инфокоммуникационных технологий ФГБОУ ВО «Донецкий государственный университет», Российская Федерация, ДНР, г. Донецк.
E-mail: ut5iv@mail.ru

Danilov Vladimir Vasilevich

Doctor of Technical Sciences, Full Professor, Head of the Department of Radiophysics and Infocommunication Technologies of Donetsk State University, Russian Federation, DPR, Donetsk.

8. Рисунки и таблицы располагаются по тексту строго в пределах печатного поля книжной ориентации страниц после их первого упоминания в тексте статьи. Вся текстовая информация на рисунках должна быть четкой и разборчивой и не иметь лишних деталей. Каждый рисунок должен иметь подрисуночную подпись (не совмещенную с рисунком) в формате: «Рис. №. Название». Каждая таблица должна иметь заглавие в формате: «Таблица №. Название». Во избежание смещений рисунков рекомендуется размещать их в таблицах с двумя строками с параметром «Нет границ» (в первой строке размещается рисунок, во второй - подрисуночная подпись). Все рисунки и таблицы должны быть последовательно пронумерованы арабскими цифрами, допускается применение букв при нумерации рисунков, состоящих из нескольких частей. Сканированные рисунки не допускаются. Образец оформления рисунка и таблицы:

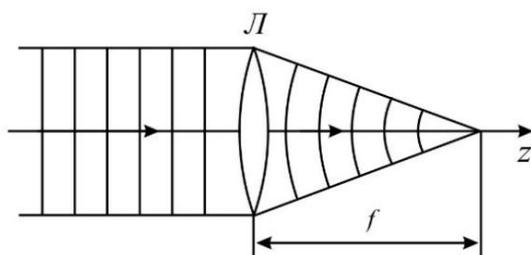


Рис. 1. Действие линзы на плоскую волну

Таблица 1. Номиналы элементов схемы УМ класса F₃

Элемент	Номинал	Единица измерения
C _{out}	25	пФ
C ₁	10	нФ
C ₂	100	мкФ
L ₁	150	мкГн
L ₂	460	нГн
L ₃	765	нГн
R ₁	0,39	МОм

9. Формулы выравниваются по центру и имеют, при необходимости, сквозную нумерацию по правому краю. Для выравнивания формул рекомендуется использовать таблицы с параметром «Нет границ». После формул ставится тот знак препинания, который необходим исходя из построения фразы. Нумеровать следует только те формулы, на которые имеются ссылки в тексте. Ссылки на формулы приводятся в круглых скобках. Каждая формула отделяется от основного текста интервалом в 6 пт. перед и после. Формулы, их компоненты и все переменные в тексте и отдельно в строках набираются только с помощью редакторов формул Microsoft Equation 3.0 или MathType 5.0 и выше с параметрами:

- **текст и переменная:** наклонный;
- **матрица-вектор:** полужирный наклонный;
- **размер:** обычный - 12 пт., крупный индекс - 9 пт., мелкий индекс - 7 пт., крупный символ - 18 пт., мелкий символ - 12 пт.

Образец оформления формулы:

$$y_t = c_0 + \sum_{i=1}^n c_i y_{t-i} + b \xi_t \quad (1)$$

10. Материалы предоставляются по адресу:

283001, Российская Федерация, Донецкая Народная Республика, г. Донецк, пр. Театральный, д. 13, Физико-технический факультет.

Главный редактор: Данилов Владимир Васильевич, каб. 203.

Зам. главного редактора: Третьяков Игорь Александрович, каб. 210.

Ответственный секретарь: Пометун Екатерина Дмитриевна, каб. 001.

Контактный тел.: +7 (856) 302 07 81

Е-mail: vestnik.g@mail.ru

Научное издание

**Вестник Донецкого национального университета.
Серия Г: Технические науки**

– 2024. – № 4.

Научное и техническое редактирование: И.А. Третьяков, Е.Д. Пометун

Адрес редакции:

ФГБОУ ВО «Донецкий государственный университет»,
283001, Российская Федерация, ДНР, г. Донецк, пр. Театральный, д. 13.

Тел.: +7 (856) 302 07 81

E-mail: vestnik.g@mail.ru

URL: <http://donnu.ru/vestnikG>

Подписано в печать 03.12.2024 г.

Формат 60×84/8. Бумага офсетная.

Печать – цифровая. Усл. печ. л. 29,76.

Тираж 100 экз.

Бесплатно.

Издательство ФГБОУ ВО «Донецкий государственный университет»
283001, Российская Федерация, ДНР, г. Донецк, ул. Университетская, д. 24.

Тел.: +7 (856) 302 92 27.