ISSN: 2224-5227 (Print)
ISSN: 2518-1483 (Online)

ACADEMIC SCIENTIFIC JOURNAL OF COMPUTER SCIENCE

№4
2025

ISSN 2518-1726 (Online),
ISSN 1991-346X (Print)

CENTRAL ASIAN ACADEMIC
RESEARCH CENTER

ACADEMIC SCIENTIFIC JOURNAL OF COMPUTER SCIENCE

4 (356)

OCTOBER – DECEMBER 2025

PUBLISHED SINCE JANUARY 1963
PUBLISHED 4 TIMES A YEAR

ALMATY, NAS RK

CHIEF EDITOR:

MUTANOV Galimkair Mutanovich, doctor of technical sciences, professor, academician of NAS RK, acting General Director of the Institute of Information and Computing Technologies CS MES RK (Almaty, Kazakhstan), https://www.scopus.com/authid/detail.uri?authorId=6506682964, https://www.webofscience.com/wos/author/record/1423665

EDITORIAL BOARD:

KALIMOLDAYEV Maksat Nuradilovich, (Deputy Editor-in-Chief), Doctor of Physical and Mathematical Sciences, Professor, Academician of NAS RK, Advisor to the General Director of the Institute of Information and Computing Technologies of the CS MES RK, Head of the Laboratory (Almaty, Kazakhstan), https://www.scopus.com/authid/detail.uri?authorId=56153126500, https://www.webofscience.com/wos/author/record/2428551

MAMYRBAEV Orken Zhumazhanovich, (Academic Secretary), PhD in Information Systems, Deputy Director for Science of the Institute of Information and Computing Technologies CS MES RK (Almaty, Kazakhstan), https://www.scopus.com/authid/detail.uri?authorId=55967630400, https://www.webofscience.com/wos/author/record/1774027

BAIGUNCHEKOV Zhumadil Zhanabaevich, Doctor of Technical Sciences, Professor, Academician of NAS RK, Institute of Cybernetics and Information Technologies, Department of Applied Mechanics and Engineering Graphics, Satbayev University (Almaty, Kazakhstan), https://www.scopus.com/authid/detail.uri?authorId=6506823633, https://www.webofscience.com/wos/author/record/1923423

WOICIK Waldemar, Doctor of Technical Sciences (Phys.-Math.), Professor of the Lublin University of Technology (Lublin, Poland), https://www.scopus.com/authid/detail.uri?authorId=7005121594, https://www.webofscience.com/wos/author/record/678586

SMOLARJ Andrej, Associate Professor Faculty of Electronics, Lublin polytechnic university (Lublin, Poland), https://www.scopus.com/authid/detail.uri?authorId=56249263000, https://www.webofscience.com/wos/author/record/1268523

KEILAN Alimkhan, Doctor of Technical Sciences, Professor (Doctor of science (Japan)), chief researcher of Institute of Information and Computational Technologies CS MES RK (Almaty, Kazakhstan), https://www.scopus.com/authid/detail.uri?authorId=8701101900, https://www.webofscience.com/wos/author/record/1436451

KHAIROVA Nina, Doctor of Technical Sciences, Professor, Chief Researcher of the Institute of Information and Computational Technologies CS MES RK (Almaty, Kazakhstan), https://www.scopus.com/authid/detail.uri?authorId=37461441200, https://www.webofscience.com/wos/author/record/1768515

OTMAN Mohamed, PhD, Professor of Computer Science Department of Communication Technology and Networks, Putra University Malaysia (Selangor, Malaysia), https://www.scopus.com/authid/detail.uri?authorId=56036884700, https://www.webofscience.com/wos/author/record/747649

NYSANBAYEVA Saule Yerkebulanovna, Doctor of Technical Sciences, Associate Professor, Senior Researcher of the Institute of Information and Computing Technologies CS MES RK (Almaty, Kazakhstan), https://www.scopus.com/authid/detail.uri?authorId=55453992600, https://www.webofscience.com/wos/author/record/3802041

BIYASHEV Rustam Gakashevich, doctor of technical sciences, professor, Deputy Director of the Institute for Informatics and Management Problems, Head of the Information Security Laboratory (Kazakhstan), https://www.scopus.com/authid/detail.uri?authorId=6603642864, https://www.webofscience.com/wos/author/record/3802016

KAPALOVA Nursulu Aldazharovna, Candidate of Technical Sciences, Head of the Laboratory cybersecurity, Institute of Information and Computing Technologies CS MES RK (Almaty, Kazakhstan), https://www.scopus.com/authid/detail.uri?authorId=57191242124,

KOVALYOV Alexander Mikhailovich, Doctor of Physical and Mathematical Sciences, Academician of the National Academy of Sciences of Ukraine, Institute of Applied Mathematics and Mechanics (Donetsk, Ukraine), https://www.scopus.com/authid/detail.uri?authorId=7202799321, https://www.webofscience.com/wos/author/record/38481396

MIKHALEVICH Alexander Alexandrovich, Doctor of Technical Sciences, Professor, Academician of the National Academy of Sciences of Belarus (Minsk, Belarus), https://www.scopus.com/authid/detail.uri?authorId=7004159952, https://www.webofscience.com/wos/author/record/46249977

TIGHINEANU Ion Mihailovich, Doctor of Physical and Mathematical Sciences, Academician, President of the Academy of Sciences of Moldova, Technical University of Moldova (Chisinau, Moldova), https://www.scopus.com/authid/detail.uri?authorId=7006315935, https://www.webofscience.com/wos/author/record/524462

Academic Scientific Journal of Computer Science
ISSN 2518-1726 (Online),
ISSN 1991-346X (Print)

Owner: «Central Asian Academic Research Center» LLP (Almaty).

Certificate **№ KZ77VPY00121154** on the re-registration of the periodical printed and online publication of the information agency, issued on **05.06.2025** by the Republican State Institution «Information Committee» of the Ministry of Culture and Information of the Republic of Kazakhstan

Subject area*: information and communication technologies.*

Currently: *included in the list of journals recommended by the CCSES MSHE RK in the direction of «Information and communication technologies».*

Periodicity: *4 times a year.*

http://www.physico-mathematical.kz/index.php/en/

© «Central Asian Academic Research Center» LLP, 2025

БАС РЕДАКТОР:

МҰТАНОВ Ғалымқайыр Мұтанұлы, техника ғылымдарының докторы, профессор, ҚР ҰҒА академигі, ҚР ҒЖБМ ҒК «Ақпараттық және есептеу технологиялары институты» бас директорының м.а. (Алматы, Қазақстан), https://www.scopus.com/authid/detail.uri?authorId=6506682964, https://www.webofscience.com/wos/author/record/1423665

РЕДАКЦИЯ АЛҚАСЫ:

ҚАЛИМОЛДАЕВ Мақсат Нұрәділұлы, (бас редактордың орынбасары), физика-математика ғылымдарының докторы, профессор, ҚР ҰҒА академигі, ҚР ҒЖБМ ҒК «Ақпараттық және есептеу технологиялары институты» бас директорының кеңесшісі, зертхана меңгерушісі (Алматы, Қазақстан), https://www.scopus.com/authid/detail.uri?authorId=56153126500, https://www.webofscience.com/wos/author/record/2428551

МАМЫРБАЕВ Өркен Жұмажанұлы (ғалым хатшы), Ақпараттық жүйелер саласындағы техника ғылымдарының (PhD) докторы, ҚР ҒЖБМ ҒК «Ақпараттық және есептеу технологиялары институты» директорының ғылым жөніндегі орынбасары (Алматы, Қазақстан), https://www.scopus.com/authid/detail.uri?authorId=55967630400, https://www.webofscience.com/wos/author/record/1774027

БАЙГУНЧЕКОВ Жұмаділ Жаңабайұлы, техника ғылымдарының докторы, профессор, ҚР ҰҒА академигі, Кибернетика және ақпараттық технологиялар институты, Қолданбалы механика және инженерлік графика кафедрасы, Сәтбаев университеті (Алматы, Қазақстан), https://www.scopus.com/authid/detail.uri?authorId=6506823633, https://www.webofscience.com/wos/author/record/1923423

ВОЙЧИК Вальдемар, техника ғылымдарының докторы (физ-мат), Люблин технологиялық университетінің профессоры (Люблин, Польша), https://www.scopus.com/authid/detail.uri?authorId=7005121594, https://www.webofscience.com/wos/author/record/678586

СМОЛАРЖ Анджей, Люблин политехникалық университетінің электроника факультетінің доценті (Люблин, Польша), https://www.scopus.com/authid/detail.uri?authorId=56249263000, https://www.webofscience.com/wos/author/record/1268523

КЕЙЛАН Әлімхан, техника ғылымдарының докторы, профессор (ғылым докторы (Жапония)), ҚР ҒЖБМ ҒК «Ақпараттық және есептеу технологиялары институтының» бас ғылыми қызметкері (Алматы, Қазақстан), https://www.scopus.com/authid/detail.uri?authorId=8701101900, https://www.webofscience.com/wos/author/record/1436451

ХАЙРОВА Нина, техника ғылымдарының докторы, профессор, ҚР ҒЖБМ ҒК «Ақпараттық және есептеу технологиялары институтының» бас ғылыми қызметкері (Алматы, Қазақстан), https://www.scopus.com/authid/detail.uri?authorId=37461441200, https://www.webofscience.com/wos/author/record/1768515

ОТМАН Мохаммед, PhD, Информатика, Коммуникациялық технологиялар және желілер кафедрасының профессоры, Путра университеті Малайзия (Селангор, Малайзия), https://www.scopus.com/authid/detail.uri?authorId=56036884700, https://www.webofscience.com/wos/author/record/747649

НЫСАНБАЕВА Сауле Еркебұланқызы, техника ғылымдарының докторы, доцент, ҚР ҒЖБМ ҒК «Ақпараттық және есептеу технологиялары институтының» аға ғылыми қызметкері (Алматы, Қазақстан), https://www.scopus.com/authid/detail.uri?authorId=55453992600, https://www.webofscience.com/wos/author/record/3802041

БИЯШЕВ Рустам Гакашевич, техника ғылымдарының докторы, профессор, Информатика және басқару мәселелері институты директорының орынбасары, Ақпараттық қауіпсіздік зертханасының меңгерушісі (Қазақстан), https://www.scopus.com/authid/detail.uri?authorId=6603642864, https://www.webofscience.com/wos/author/record/3802016

КАПАЛОВА Нұрсұлу Алдажарқызы, техника ғылымдарының кандидаты, ҚР ҒЖБМ ҒК «Ақпараттық және есептеу технологиялары институты», Киберқауіпсіздік зертханасының меңгерушісі (Алматы, Қазақстан), https://www.scopus.com/authid/detail.uri?authorId=57191242124,

КОВАЛЕВ Александр Михайлович, физика-математика ғылымдарының докторы, Украина Ұлттық Ғылым академиясының академигі, Қолданбалы математика және механика институты (Донецк, Украина), https://www.scopus.com/authid/detail.uri?authorId=7202799321, https://www.webofscience.com/wos/author/record/38481396

МИХАЛЕВИЧ Александр Александрович, техника ғылымдарының докторы, профессор, Беларусь Ұлттық Ғылым академиясының академигі (Минск, Беларусь), https://www.scopus.com/authid/detail.uri?authorId=7004159952, https://www.webofscience.com/wos/author/record/46249977

ТИГИНЯНУ Ион Михайлович, физика-математика ғылымдарының докторы, академик, Молдова Ғылым Академиясының президенті, Молдова техникалық университеті (Кишинев, Молдова), https://www.scopus.com/authid/detail.uri?authorId=7006315935, https://www.webofscience.com/wos/author/record/524462

Academic Scientific Journal of Computer Science
ISSN 2518-1726 (Online),
ISSN 1991-346X (Print)

Меншіктеуші: «Орталық Азия академиялық ғылыми орталығы» ЖШС (Алматы).

Ақпарат агенттігінің мерзімді баспасөз басылымын, ақпарат агенттігін және желілік басылымды қайта есепке қою туралы ҚР Мәдениет және Ақпарат министрлігі «Ақпарат комитеті» Республикалық мемлекеттік мекемесі **05.06.2025** ж. берген **№ KZ77VPY00121154** Куәлік.

Тақырыптық бағыты: *ақпараттық-коммуникациялық технологиялар*

Қазіргі уақытта: *«ақпараттық-коммуникациялық технологиялар» бағыты бойынша ҚР БҒМ БҒСБК ұсынған журналдар тізіміне енді.*

Мерзімділігі: *жылына 4 рет.*

http://www.physico-mathematical.kz/index.php/en/

© «Орталық Азия академиялық ғылыми орталығы» ЖШС, 2025

ГЛАВНЫЙ РЕДАКТОР:

МУТАНОВ Галимкаир Мутанович, доктор технических наук, профессор, академик НАН РК, и.о. генерального директора «Института информационных и вычислительных технологий» КН МНВО РК (Алматы, Казахстан), https://www.scopus.com/authid/detail.uri?authorId=6506682964, https://www.webofscience.com/wos/author/record/142 3665

Редакционная коллегия:

КАЛИМОЛДАЕВ Максат Нурадилович, (заместитель главного редактора), доктор физико-математических наук, профессор, академик НАН РК, советник генерального директора «Института информационных и вычислительных технологий» КН МНВО РК, заведующий лабораторией (Алматы, Казахстан), https://www.scopus.com/authid/detail.uri?authorId=56153126500, https://www.webofscience.com/wos/author/record/2428551

МАМЫРБАЕВ Оркен Жумажанович, (ученый секретарь), доктор философии (PhD) по специальности «Информационные системы», заместитель директора по науке РГП «Институт информационных и вычислительных технологий» Комитета науки МНВО РК (Алматы, Казахстан), https://www.scopus.com/authid/detail.uri?authorId=55967630400, https://www.webofscience.com/wos/author/record/1774027

БАЙГУНЧЕКОВ Жумадил Жанабаевич, доктор технических наук, профессор, академик НАН РК, Институт кибернетики и информационных технологий, кафедра прикладной механики и инженерной графики, Университет Сатпаева (Алматы, Казахстан), https://www.scopus.com/authid/detail.uri?authorId=6506823633, https://www.webofscience.com/wos/author/record/1923423

ВОЙЧИК Вальдемар, доктор технических наук (физ.-мат.), профессор Люблинского технологического университета (Люблин, Польша), https://www.scopus.com/authid/detail.uri?authorId=7005121594, https://www.webofscience.com/wos/author/record/678586

СМОЛАРЖ Анджей, доцент факультета электроники Люблинского политехнического университета (Люблин, Польша), https://www.scopus.com/authid/detail.uri?authorId=56249263000, https://www.webofscience.com/wos/author/record/1268523

КЕЙЛАН Алимхан, доктор технических наук, профессор (Doctor of science (Japan)), главный научный сотрудник РГП «Института информационных и вычислительных технологий» КН МНВО РК (Алматы, Казахстан), https://www.scopus.com/authid/detail.uri?authorId=8701101900, https://www.webofscience.com/wos/author/record/1436451

ХАЙРОВА Нина, доктор технических наук, профессор, главный научный сотрудник РГП «Института информационных и вычислительных технологий» КН МНВО РК (Алматы, Казахстан), https://www.scopus.com/authid/detail.uri?authorId=37461441200, https://www.webofscience.com/wos/author/record/1768515

ОТМАН Мохамед, доктор философии, профессор компьютерных наук, Департамент коммуникационных технологий и сетей, Университет Путра Малайзия (Селангор, Малайзия), https://www.scopus.com/authid/detail.uri?authorId=56036884700, https://www.webofscience.com/wos/author/record/747649

НЫСАНБАЕВА Сауле Еркебулановна, доктор технических наук, доцент, старший научный сотрудник РГП «Института информационных и вычислительных технологий» КН МНВО РК (Алматы, Казахстан), https://www.scopus.com/authid/detail.uri?authorId=55453992600, https://www.webofscience.com/wos/author/record/3802041

БИЯШЕВ Рустам Гакашевич, доктор технических наук, профессор, заместитель директора Института проблем информатики и управления, заведующий лабораторией информационной безопасности (Казахстан), https://www.scopus.com/authid/detail.uri?authorId=6603642864,https://www.webofscience.com/wos/author/record/3802016

КАПАЛОВА Нурсулу Алдажаровна, кандидат технических наук, заведующий лабораторией кибербезопасности РГП «Института информационных и вычислительных технологий» КН МНВО РК (Алматы, Казахстан), https://www.scopus.com/authid/detail.uri?authorId=57191242124,

КОВАЛЕВ Александр Михайлович, доктор физико-математических наук, академик НАН Украины, Институт прикладной математики и механики (Донецк, Украина), https://www.scopus.com/authid/detail.uri?authorId=7202799321, https://www.webofscience.com/wos/author/record/38481396

МИХАЛЕВИЧ Александр Александрович, доктор технических наук, профессор, академик НАН Беларуси (Минск, Беларусь), https://www.scopus.com/authid/detail.uri?authorId=7004159952, https://www.webofscience.com/wos/author/record/46249977

ТИГИНЯНУ Ион Михайлович, доктор физико-математических наук, академик, президент Академии наук Молдовы, Технический университет Молдовы (Кишинев, Молдова), https://www.scopus.com/authid/detail.uri?authorId=7006315935, https://www.webofscience.com/wos/author/record/524462

Academic Scientific Journal of Computer Science
ISSN 2518-1726 (Online),
ISSN 1991-346X (Print)

Собственник: *ТОО «Центрально-азиатский академический научный центр» (г. Алматы).*

Свидетельство о постановке на переучет периодического печатного издания, информационного агентства и сетевого издания **№ KZ77VPY00121154.** Дата выдачи **05.06.2025**

Тематическая направленность: *информационно-коммуникационные технологии.*

В настоящее время: *вошел в список журналов, рекомендованных КОКСНВО МНВО РК по направлению «информационно-коммуникационные технологии».*

Периодичность: *4 раза в год.*

http://www.physico-mathematical.kz/index.php/en/

© ТОО «Центрально-азиатский академический научный центр», 2025

CONTENTS

B. Assanova, Zh. Moldasheva, A.T. Kishubaeva
Decision support system structure and blocks for selecting efficient delayed coking modes.....11

Zh.T. Abildayeva, R.K. Uskenbayeva, G.S. Beketova, N.B. Konyrbaev, S.B. Seydazimov
Multi-criterion optimization of advertising budget allocation in the agro-industrial complex based on NSGA-III algorithm.....26

A.O. Aliyeva, B.S. Omarov, R.B. Abdrakhmanov, D.R. Sultan, A.B. Toktarova
Neural network model for automatic detection of Kazakh-language hate speech.....40

O. Auyelbekov, E. Bostanov, S. Sapakova, L. Tukenova, A. Kozhagul
Modeling and analysis of a generator with permanent and variable magnets.....55

G. Autova, G. Nurtayeva, E. Zulbukharova, G. Yeleussizova, R. Zhumabekova
Theoretical foundations of interdisciplinary integration of physics and computer science.....73

A.Zh. Akhmetova, M.A. Kantureyeva, A.A. Abisheva, A. Aubakirova, A.A. Shekerbek
Analysis of the social network user's enviroment.....89

A.Sh. Barakova, K.S. Shadinova, A.S. Orynbaeva, G. Sugurzhanova
Design of a model for protecting a website's authentication data and content based on blockchain technology.....102

A.N. Zhidebayeva, G.U. Madaliyeva, B.O. Tastanbekova, S.S. Karzhaubekova, G.S. Shaimerdenova
Deep neural network Conv-LSTM for ECG-based cardiac disorder identification.....122

N.M. Zhunissov, A.B. Aben, A.B. Amanzholova
The fraud detection model in text messages.....138

A. Issakhov, A. Alzhanov, A. Akhmedov, A. Amanzholov, T. Murat
Numerical simulation of thermohydrodynamics during heated water discharge into Lake Balkhash.....152

Z. Kaderkeyeva, B. Razakhova, G. Bekmanova, A. Nazyrova, M. Zhasuzakova
Q-Bilim: an intelligent system for assessing learning outcomes based on competencies.....171

N. Karymsakova, A. Boltaboyeva, D. Turmakhanbet, M. Maulenbekov, T. Abdirova
Unsupervised learning for the identification of critical conditions in renewable energy production.....184

A.Kulakayeva, E.Daineko, B. Medetov, A. Nurlankyzy
Evaluation of the effectiveness of modern neural network architectures for VAD under low snr ratio conditions.....203

B. Orazbayev, A. Zhumadillayeva, K. Orazbayeva, R. Yessirkessinov, Zh. Tuleuov
Development of models of sulfur production processes based on artificial neural networks and simulation.....216

L. Rzayeva, A. Ryzhova, M. Zhaparkhanova, A. Myrzatay, Zh. Kozhakhmet
A new LSTM-based web application for automated password strength evaluation.....234

D. Sagidoldin, A. Zhetpisbayeva, B. Zhumazhanov, B. Zhumazhanov
Increasing the reliability of data transmission from small spacecraft using SDR equipment.....259

A.N. Seraly, A.D. Mekhtiyev, G.Z. Ziyatbekova, K.B. Begalieva, R.A. Mekhtiyev
Development of hardware for monitoring optical parameters.....274

A.A. Taurbekova, M.V. Markosyan
Development and implementation of a computational model of magmatic processes in the bowls of the Earth and on its surface.....288

K. Chezhimbayeva, A. Mukhamejanova, Y. Garmashova
Fuzzy-logic-based expert system for predicting QoS in 5G networks.....306

МАЗМҰНЫ

Б.У. Асанова, Ж.Ж. Молдашева, А. Кишубаева
Баяу кокстеу қондырғысы үшін тиімді жұмыс режимдерін таңдауға шешім қолдау жүйесі құрылымы.....11

Ж.Т. Әбілдаева, Р.К. Ускенбаева, Г.С. Бекетова, Н.Б. Қоңырбаев, С.Б. Сейдазимов
NSGA-III алгоритмі негізінде агроөнеркәсіптік кешендегі жарнамалық бюджетті бөлуді көп критериалды оңтайландыру.....26

А.О. Әлиева, Б.С. Омаров, Р.Б. Абдрахманов, Д.Р. Султан, А.Б. Тоқтарова
Қазақ тіліндегі дискриминацияны автоматты анықтауға арналған нейрондық желілік моделі.....40

О. Әуелбеков, Е. Бостанов, С. Сапақова, Л. Түкенова, А. Қожағұл
Тұрақты және айнымалы магниттері бар генераторды модельдеу және талдау.....55

Г.М. Аутова, Г.К. Нуртаева, Э.М. Зульбухарова, Г.С. Елеусизова, Р.Р. Жұмабекова
Физика мен информатика пәндерінің пәнаралық интеграциясының теориялық негіздері.....73

А.Ж. Ахметова, М.А. Кантуреева, А.А. Абишева, А. Аубакирова, А.А. Шекербек
Әлеуметтік желі қолданушыларының ортасын талдау.....89

А.Ш. Баракова, К.С. Шадинова, А.С. Орынбаева, Г. Сугуржанова
Блокчейн технологиясы негізінде веб сайттың аутентификациялық деректері мен өнімін қорғау моделін құрастыру.....102

А.Н. Жидебаева, Г.У. Мадалиева, Б.О. Тастанбекова, С.С. Қаржаубекова, Г.С. Шаймерденова
Жүрек ауруларын анықтауда Conv-LSTM архитектурасына негізделген терең нейрондық желі.....122

Н.М. Жунисов, А.Б. Абен, Ә.Б. Аманжолова
Мәтіндік хабарламалардағы алаяқтықты анықтау моделі.....138

А.А. Исахов, А. Альжанов, А. Ахмедов, А. Аманжолов, Т. Мурат
Балқаш көліне жылы су ағызу кезіндегі термогидродинамиканы сандық модельдеу.....152

З.К. Кадеркеева, Б.Ш. Разахова, Г.Т. Бекманова, А.Е. Назырова, М.Ж. Жасұзақова
Q-Bilim: құзыреттерге негізделген оқу нәтижелерін бағалауға арналған интеллектуалды жүйе.....171

Н. Карымсакова, А. Болтабоева, Д. Тұрмаханбет, М. Мауленбеков, Т. Абдирова
Жаңартылатын энергия өндірісіндегі критикалық режимдерді анықтауға арналған мұғалімсіз оқыту.....184

А. Кулакаева, Е. Дайнеко, Б. Медетов, А. Нурланкызы
Сигнал/шуыл қатынасы төмен жағдайларда заманауи нейрондық желілік VAD архитектураларының тиімділігін бағалау.....203

Б. Оразбаев, А. Жумадиллаева, К. Оразбаева, Р. Есиркесинов, Ж. Тулеуов
Күкірт өндіру процесстерінің модельдерін жасанды нейрондық желілер негізінде әзірлеу және модельдеу.....216

Л. Рзаева, А. Рыжова, М. Жапарханова, А. Мырзатай, Ж. Кожахмет,
Құпиясөздің беріктігін автоматты бағалауға арналған LSTM негізіндегі жаңа веб-қосымша.....234

Д.Т. Сагидолдин, А.Т. Жетписбаева, Б.Р. Жумажанов, Б.С. Жумажанов
SDR жабдықтарын пайдалану арқылы, шағын ғарыш аппараттарынан деректерді берудің сенімділігін арттыру.....259

А.Н. Сералы, А.Д. Мехтиев, Г.З. Зиятбекова, К.Б. Бегалиева, Р.А. Мехтиев
Оптикалық параметрлерді бақылауға арналған аппараттық құрылғыны әзірлеу.....274

А.А. Таурбекова, М.В. Маркосян
Жер көзіндегі және оның бетіндегі магматтық процестердің есептік моделін әзірлеу және енгізу.....288

К.С. Чежимбаева, А. Мухамеджанова, Ю. Гармашова
Айқын емес логика негізінде 5G желілеріндегі QoS болжау эксперттік жүйесі.....306

СОДЕРЖАНИЕ

Б.У. Асанова, Ж.Ж. Молдашева, А. Кишубаева
Структура и функциональные блоки системы поддержки решений для выбора режимов замедленного коксования.....11

Ж.Т. Абилдаева, Р.К. Ускенбаева, Г.С. Бекетова, Н.Б. Конырбаев, С.Б. Сейдазимов
Многокритериальная оптимизация распределения рекламного бюджета в апк на основе алгоритма NSGA-III.....26

А.О. Алиева, Б.С. Омаров, Р.Б. Абдрахманов, Д.Р. Султан, А.Б. Токтарова
Нейросетевая модель для автоматического обнаружения дискриминации в казахском языке.....40

О. Ауельбеков, Е. Бостанов, С. Сапакова, Л. Тукенова, А. Кожагул
Моделирование и анализ генератора с постоянными и переменными магнитами.....55

Г.М. Аутова, Г.К. Нуртаева, Э.М. Зульбухарова, Г.С. Елеусизова, Р.Р. Жумабекова
Теоретические основы междисциплинарной интеграции физики и информатики.....73

А.Ж. Ахметова, М.А. Кантуреева, А.А. Абишева, А. Аубакирова, А.А. Шекербек
Анализ окружения ползователей социальной сети.....89

А.Ш. Баракова, К.С. Шадинова, А.С. Орынбаева, Г. Сугуржанова
Разработка модели защиты аутентификационных данных и контента веб-сайта на основе технологии блокчейн.....102

А.Н. Жидебаева, Г.У. Мадалиева, Б.О. Тастанбекова, С.С. Каржаубекова, Г.С. Шаймерденова
Глубокая нейронная сеть на основе архитектуры Conv-LSTM для выявления сердечных заболеваний.....122

Н.М. Жунисов, А.Б. Абен, А.Б. Аманжолова
Модель обнаружения мошенничества в текстовых сообщениях.....138

А.А. Исахов, А. Альжанов, А. Ахмедов, А. Аманжолов, Т. Мурат
Численное моделирование термогидродинамики при сбросе подогретых вод в озеро Балхаш.....152

З.К. Кадеркеева, Б.Ш. Разахова, Г.Т. Бекманова, А.Е. Назырова, М.Ж. Жасузакова
Q-Bilim: интеллектуальная система оценки результатов обучения на основе компетенций.....171

Н. Карымсакова, А. Болтабоева, Д. Тұрмаханбет, М. Мауленбеков, Т. Абдирова
Обучение без учителя для выявления критических режимов в производстве возобновляемой энергии.....184

А. Кулакаева, Е. Дайнеко, Б. Медетов, А. Нурланкызы
Оценка эффективности современных нейросетевых архитектур VAD при низком отношении сигнал/шум.....203

Б. Оразбаев, А. Жумадиллаева, К. Оразбаева, Р. Есиркесинов, Ж. Тулеуов
Разработка моделей процессов производства серы на основе искусственных нейронных сетей и моделирование.....216

Л. Рзаева, А. Рыжова, М. Жапарханова, А. Мырзатай, Ж. Кожахмет
Новое веб-приложение на основе LSTM для автоматизированной оценки надежности паролей.....234

Д.Т. Сагидолдин, А.Т. Жетписбаева, Б.Р. Жумажанов, Б.С. Жумажанов
Повышение надёжности передачи данных с малых космических аппаратов с применением SDR оборудования.....259

А.Н. Сералы, А.Д. Мехтиев, Г.З. Зиятбекова, К.Б. Бегалиева, Р.А. Мехтиев
Разработка аппаратного средства для контроля оптических параметров.....274

А.А. Таурбекова, М.В. Маркосян, Н.Т. Карымсакова
Разработка и реализация вычислительной модели магматических процессов в недрах земли и на её поверхности.....288

К.С. Чежимбаева, А. Мухамеджанова, Ю. Гармашова
Экспертная система прогнозирования QoS в 5G-сетях на основе нечеткой логики.....306

ACADEMIC SCIENTIFIC JOURNAL OF COMPUTER SCIENCE
ISSN 1991-346X
Volume 4. Namber 356 (2025). 11–25

https://doi.org/10.32014/2025.2518-1726.380

МРНТИ 27.47.19

© B. Assanova[1*], Zh. Moldasheva[1], A.T. Kishubaeva[2], 2025.
[1]Kh. Dosmukhamedov Atyrau University, Atyrau, Kazakhstan;
[2]Toraighyrov University, Pavlodar, Kazakhstan.
E-mail: baha1981_13@mail.ru

DECISION SUPPORT SYSTEM STRUCTURE AND BLOCKS FOR SELECTING EFFICIENT DELAYED COKING MODES

Assanova Baktygul — PhD, associate professor of Kh. Dosmukhamedov Atyrau University, Atyrau, Kazakhstan,
E-mail: baha1981_13@mail.ru, ORCID ID: https://orcid.org/0000-0003-1029-6266;
Moldasheva Zhadra — PhD, associate professor of Kh. Dosmukhamedov Atyrau University, Atyrau, Kazakhstan,
E-mail: zhadira1985@mail.ru, ORCID ID: https://orcid.org/0000-0002-0559-3410;
Kishubayeva Aizhan — doctoral student Toraigyrov University, Pavlodar, Kazakhstan,
E-mail: Altynai_999@mail.ru orcid id https://orcid.org/0009-0004-1125-8696.

Abstract. This work is aimed, using the example of a delayed coking unit, to study the problems of effective control of the operating modes of complex chemical-technological systems, characterized by a shortage and vagueness of initial information, and to effectively solve them using intelligent systems. Recently, due to the complication of production technological systems and the processes occurring in them, the shortage and vagueness of the initial information that reliably describes them, effective management of the operating modes of such objects has become an urgent scientific and practical task. The paper examines the problems of developing and using intelligent decision support systems for effectively solving decision-making problems to control the operating modes of the object of study in a fuzzy environment. As a result of the study, the structure of an intelligent decision support system is proposed, which allows choosing effective operating modes of complex technological objects characterized by vagueness, such as a delayed coking unit, and the functional blocks of such systems and their systemic functioning are described. The proposed structure of an intelligent system compared to similar systems is that its structure includes: a package of object models and heuristic decision-making methods for choosing an effective mode of operation in a fuzzy environment, based on the knowledge, experience and intuition of decision-makers and expert specialists; intelligent user interface.

Keywords: decision-making tasks, intellectualized system, decision maker, delayed coking unit, fuzzy information, functional blocks, user interface

Funding *This work was supported the Science Committee of the Ministry of Science and Higer Education of the Republic of Kazakhstan (Grant No. fund this research AP19679897-Development of an intelligent decision-making system for controlling the petroleum coke production process).*

© Б.У. Асанова [1*], Ж.Ж. Молдашева[1], А. Кишубаева[2], 2025.
[1] Х. Досмухамедов атындағы Атырау университеті, Атырау, Қазақстан;
[2]Торайғыров университеті, Павлодар, Қазақстан
E-mail: baha1981_13@mail.ru

БАЯУ КОКСТЕУ ҚОНДЫРҒЫСЫ ҮШІН ТИІМДІ ЖҰМЫС РЕЖИМДЕРІН ТАҢДАУҒА ШЕШІМ ҚОЛДАУ ЖҮЙЕСІ ҚҰРЫЛЫМЫ

Асанова Бактыгул — PhD, Х. Досмұхамедов атындағы Атырау университетінің қауымдастырылған профессоры, Атырау, Қазақстан,
E-mail: baha1981_13@mail.ru, ORCID ID: https://orcid.org/0000-0003-1029-6266;
Молдашева Жадра — PhD, Х. Досмұхамедов атындағы Атырау университетінің қауымдастырылған профессоры, Атырау, Қазақстан,
E-mail: zhadira1985@mail.ru, ORCID ID: https://orcid.org/0000-0002-0559-3410;
Кишубаева Айжан — Торайғыров Университетінің докторанты, Павлодар, Қазақстан,
E-mail: Altynai_999@mail.ru, orcid id https://orcid.org/0009-0004-1125-8696.

Аннотация. Бұл зерттеу баяу кокстеу қондырғысы мысалында бастапқы ақпараттың тапшылығы және айқынсыздығымен сипатталатын күрделі химиялық-технологиялық жүйелердің жұмыс режимдерін тиімді басқару мәселелерін зерттеп, оларды интеллектуалдандырылған жүйелер көмегімкен тиімді шешуге арналған. Соңғы уақыттарды өндістік технологиялық жүйелер мен оларда жүретін процесстердің күрделенуіне, оларды дұрыс сипаттайтын ақпараттардың тапшылығы, ал қолжетімді ақпараттың айқынсыздығына байланысты мұндай нысандар жұмыссын тиімді басқаруаса өзекті ғылыми-практикалық мәселе болып табылады. Жұмыста зерттеу нысанының жұмыс режимдерін басқару бойынша шешім қабылдау есептерін ақпарат айқынсыздығы жағдайында тиімді шешу үшін интеллектуалдандырылған шешім қабылдауды қолдау жүйелерін әзірлеп, қолдану сұрақтары зерттелген. Зертеу нәтижесінде баяу кокстеу қондырғысы сияқты айқынсыздықпен сипатталатын күрделі технологиялық нысандардың тиімді жұмыс режимдерін таңдауға мүмкіндік беретін интеллектуалдандарылған шешім қабылдауды қолдау жүйесінің құрылымы жасақталып, ондай жүйелердің құрамындағы функционалдық блоктары мен олардың жүйелі жұмысын сипатталған. Құрылымы ұсынылған интеллектуалдандырылған жүйенің ұқсас жүйелерден артықшылығы: оның құрамына шешім қабылдаушы тұлға, эксперт-мамандар білімдері, тәжірибелері мен интуициясы негізінде құрылған модельдер пакеті мен басқару нысанының тиімді жұмыс режимін

айқын емес ортада таңдау бойынша шешім қабылдау эвристикалық тәсілдері және интеллектуалдандырылған пайдаланушы интерфейсі ендірілген.

Түйін сөздер: шешім қабылдау есебі, интеллектуалдандырылған жүйе, шешім қабылдаушы тұлға, баяу кокстеу қондырғысы, айқын емес ақпарат, функционалды боктар, пайдаланушы интерфейсі

Қаржыландыру. *Зерттеу Қазақстан Республикасы Ғылым және жоғары білімминистрлігініңҒылымкомитетіқаржыландырады(грант№AP19679897 «Мұнай коксын өндіру процесін басқару үшін интеллектуалдандырылған шешім қабылдау жүйесін әзірлеу» гранттық жобасы аясында жасалды.*

© Б. Асанова[1*], Ж. Молдашева[1], А. Кишубаева[2], 2025.
[1]Атырауский Университет имени Х. Досмухамедова, Атырау, Казахстан;
[2]Торайгыров университет, Павлодар, Казахстан.
E-mail: baha1981_13@mail.ru

СТРУКТУРА И ФУНКЦИОНАЛЬНЫЕ БЛОКИ СИСТЕМЫ ПОДДЕРЖКИ РЕШЕНИЙ ДЛЯ ВЫБОРА РЕЖИМОВ ЗАМЕДЛЕННОГО КОКСОВАНИЯ

Асанова Бактыгул — PhD, ассоциированный профессор, Атырауского университета имени Х. Досмухамедова, Атырау, Казахстан,
E-mail: baha1981_13@mail.ru, ORCID ID: https://orcid.org/0000-0003-1029-6266;
Молдашева Жадра — PhD, ассоциированный профессор, Атырауского университета имени Х. Досмухамедова, Атырау, Казахстан,
E-mail: zhadira1985@mail.ru, ORCID ID: https://orcid.org/0000-0002-0559-3410;
Кишубаева Айжан — докторант Университета имени С. Торайгырова, Павлодар, Казахстан,
E-mail: Altynai_999@mail.ru, orcid id https://orcid.org/0009-0004-1125-8696.

Аннотация. Данная работа направлено на примере установки замедленного коксования исследованию проблем эффективного управления режимами работы сложных химико-технологических систем, характеризующихся дефицитом и нечеткостью исходной информации, и эффективному решению их с помощью интеллектуализированных систем. В последнее время в связью усложнением производственных технологических систем и процессов протекающих в них, дефицитом и нечеткостью достоверно их описывающей исходной информации, эффективное управление режимами работы таких объектов стала актуальной научно-практической задачей. В работе исследуются проблемы разработки и использования интеллектуальных систем поддержки принятия решений для эффективного решения задач принятия решений по управлению режимами работы объекта исследования в нечеткой среде. В результате исследования предложена структура интеллектуальной системы поддержки принятия решений, позволяющая выбрать эффективные режимы работы сложных технологических объектов, характеризующихся

нечеткостью, таких как установка замедленного коксования, и описаны функциональные блоки таких систем и системное их функционирование. Предлагаемая структура интеллектуальной системы перед аналогичными системами в том, что в состав ее структуры включены: пакет моделей объекта и эвристические методы принятия решений по выбору эффективного режима работы его в нечеткой среде, основанных на знаниях, опыте и интуиции ЛПР, экспертов-специалистов; интеллектуальный пользовательский интерфейс.

Ключевые слова: задачи принятия решений, интеллектуализироваанная система, лицо принимающее решение, установка замедленного коксования, нечеткая информация, функциональные блоки, интерфейс пользователя

Финансирование. *Исследование финансируется Комитетом науки Министерства науки и высшего образования Республики Казахстан (грант AP19679897 "Разработка интеллектуализированной системы принятия решений для управления процессом производства нефтяного кокса")*

Кіріспе. Нашар формализацияланатын күрделі химиялық-технологиялық жүйелер (ХТЖ) қолданылатын мұнай өңдеу және басқа өндірістертердің заманауи даму кезеңінде олардың жұмыс режимдерін басқаруда шешім қабылдау есептерін тиімді шешу ұшін компьютерлік шешім қабылдауды қолдау жүйелерін қолдану аса өзекті ғылыми-практикалық мәселеге жатады. Барлық ХТЖ күрделі басқару объектілері болып табылады, себебі олар бір-біріне әсер ететін параметрлердің үлкен санымен сипатталатын өзара байланысты көптеген элементтерден (технологиялық агрегаттардан) тұрады (Lisicyn et al., 2021). Қазіргі уақытта мұндай күрделі ХТЖ технологиялық параметрлерін оптимизациялаудың, олардың жұмыс режимдерін тиімді басқарудың ең тиімді технологиясына заманауи математикалық тәсілдер мен компьютерлерге негізделген түрлі автоматтандырылған, мысалы шешім қабылдауды қолдау жүйелері (ШҚҚЖ) жатады. Аталған ХТЖ бір біріне әсер ететін көптеген параметрлермен және өз ара байланысқан агрегаттардан тұратындықтан, сондай-ақ оладардың кейбір параметрлері, өндіріс көрсеткішері өлшенбейтіндіктен олардың көпшілігі айқынсыздықпен сипатталады (Czabanski et all., 2017). Мұндай айқынсыздықпен сипатталатын ХТЖ жұмыс режимдерін басқаруда қосымша шешім қабылдауды тұлғалардың (ШҚТ), эксперттердің табиғи тілде тұжырымдалған тәжірибесі, білімі мен интуициясы болып табылатын айқын емес ақпарат қолданылады. Сондықтан айқынсыздықпен сипатталатын ХТЖ жұмыс режимдерін басқаруда шешім қабылдау есептерін тиімді шеші үшін аталған айқын емес ақпаратты тиімлі формализациялап, қолдануға мүмкіндік беретін интеллектуалдандырылған шешім қабылдауды қолдау жүйелерін (ИШҚҚЖ) құру қажеттігі туындайды (Orazbayev et all., 2023).

Бұл жұмыста ИШҚҚЖ деп математикалық қамтуына нашар формализацияланатын ХТЖ-дің мдельдер пакетін, айқысыздықта шешім

қабылдау эвристикалық тәсілдерін қамтитын және білімдер, деректер базасын, логикалық қорытындылау блогын мен интеллектуалды пайдаланушы интерфейсі бар заманауи компьютерлер негізіндегі ақпараттық-есептеу жүйесін атаймыз. Мұндай интеллектуалдандырылған жүйе күрделі, нашар формализацияланатын нысандарды басқаруда адам-оператордың қатысуымен жұмыс жасайды және шешім қабылдау процесінде ШҚТ-ға интеллектуалды қолдау көрсетуге қабілетті болуы керек. Бұл жұмыста баяу кокстеу қондырғысы (БКҚ) жұмыс режимдерін басқаруда шешім қабылдауды қолдауға арналған ИШҚҚЖ-сін құру мәселелері зерттеліп, ол жүйенің құрылымы әзірленіп, негізгі функционалдық блоктары сипатталады.

Соңғы уақытта математикалық тәсілдердің соңғы жетістіктері мен заманауи ақпараттық-коммуникациялық технологияларды пайдалана отырып, шешім қабылдауды қолдау жүйелері негізінде ХТЖ-ді отималды басқару арқылы өндірістердің экономикалық тиімділігін мен экологиялық қауіпсіздігін арттыру мәселелерін зерттеу белсенді түрде жүргізілуде (Zaskanov et all., 2013; Vadinsky et all., 2018). Мысалы (Zaskanov et all., 2013; Prokopenko et all., 2017) жұмыстарының авторлары шешім қабылдауды қолдау жүйелерінің негіздері мен оларды құру мәселелері мен тәсілдемелерін зерттеген. Келесі (Gil'boa et all., 2017) монографиясында шешім қабылдау теориясы, тәсілдері жааәне оларды практикада шешім қабыылдау процесстерінде қолдану сұрақтары мазмұндалған. Ал (Gelovani et all., 2008; Vadinsky et all., 2018) зерттеу жұмыстарында табиғи ортан күйі жайындағы ақпаратты қолдануға негізделген интеллектуалды шешім қабылдау жүйесі қарастырып, интеллектуалды жүйелердің интеллектісі деңгеййін бағалау тәсілдеріне шолу жасалған.

Алайда практикада көптеген өндірістік нысандар, солардың ішінде мұнай өңдеу ХТЖ-ін эксплуатациялау барысында қолжетімді бастапқы ақпараттыі айқынсыздығымен сипатталатын түрлі өндірістік жағдайлар туындайды. Сәйкесінде мұндай жағдайларда нысандарды дәстүрлі тәсілдерге негізделген шешім қабылдау және басқару есептеріен формализациялау және шешу қажетті нәтиже бермейді немесе мүлдем мүмкін емес. Сонымен қатар, мұндай сандық сипатталуы қиын ХТЖ жұмыс режимдерін басқару бойынша шешім қабылдау есептерін математикалық тұжырымдау және шешу олардың күрделілігіне, көпкритерийлігіне байланысты қиындай түседі. Ал ол нысандарды басқаруда ескерілітін экономикалық және экологиялық тиімді шешімдер аймағында қайшы келетіні шешім қабылдау есебін тіпті күрделендіреді. ИШҚҚЖ әзірлеуде занауи тәсілдер, мысалы, имитациялық модельдеу; білім қорларында білімді іздеу; айқын емес жиындар, эволюциялық есептеулер тәсілдері; ситуациялық талдау; жасанды нейрондық желілер және т.б. жасанды интеллект технологиялары тәсілдері қолданылады (Osipov et all., 2015; Hamilton et al., 2018).

Сонымен зерттеу тақырыбы бойынша жарияланған еңбектерді талдау

нәтижесінде айқынсыздықпен сипатталатын күрделі ХТЖ-дің тиімді жұмыс режимдерін таңдау үшін шешім қабылдауды қолдау жүйесін жасанды интеллект тәсілдері негізінде әзірлеу бүгінгі таңда аса өзекті ғылыми-практикалық мәселе болып табылады. Бұл жұмыста зерттеу нысаны болып өндірілетін өнімдерінің кей сапа көрсеткіштері (кокс ұшаңдығы, күлділігі т.б.) айқынсыздықпен сипатталатын Атырау мұнай өңдеу зауындағы қолданыстағы БКҚ кокстеу реакторлары алынған. Зерттеу пәні ретінде жасанды интеллект тәсілдері қарастырылған. Зерттеу барысында эксперттік және айқын емес жиындар теориясы тәсілдері, шешім қабылдау есептерін шешуде эвристикалық тәсілдер және жүйелік талдау методологиясы қолданылады. Жұмыстың зерттеу мақсаты: БКҚ тиімді жұмыс режимдерін таңдауға мүмкіндік беретін интеллектуалдандырылған ШҚҚЖ құрылымын құру және ондай жүйелердің функционалдық блоктары мен олардың жүйелі жұмысын сипаттау. Алынған зерттеу нәтижелері шешім қабылдау теориясы мен тәсілдері үшін теориялық және практикалық маңызы зор. Себебі алынған нәтижелер шешілетін өндірістік есептердің ауқымын кеңейтеді, өндірістік жағдайлар мен міндеттерді неғұрлым адекватты сипаттауға және айқын емес ортада ИШҚҚЖ негізінде тиімді шешуге мүмкіндік береді.

Материалдар мен әдістер. Зерттеу барысында қажетті материалдар ретінде зерттеу нысанының, яғни кокстеу реакторларының жұмыс регламенті мен олардың өнімдерінің көлемі мен сапасын сипаттайтын авторлар (Assanova et all., 2023) жұмысында ұсынған модельдер кешені қолданылады. Сонымен қатар материалдар ретінде шешім қабылдау есептерін шешу эвристикалық тәсілдері мен интеллектуалдандандырылған жүйелерді құру процедуралары да пайдаланылады.

Зерттеу материалдары, атап айтқанда кокстеу реакторлары өнімдері (мұнай коксы, мұнай өнімдері булары) жайында мәліметтер сандық (өлшеу құралдары арқылы өлшенген) форматта қолданылады. Ал негізгі өнім кокстың сапасын, яғни ұшаңдығы мен күлділігін бағалайтын ақпарат сапалық (эксперт-мамандардың табиғи тілінде сипаталған, яңни айқын емес) тұрғыда болады.

Тұжырымдалған мақсатқа жету, яғни БКҚ тиімді жұмыс режимдерін таңдау интеллектуалдандырылған ШҚҚЖ құрылымын құрып, функционалдық блоктары жүйелі жұмысын сипаттау, үшін зерттеу бағытындағы жұмыстарды талдау, жалпылау, жетілдіру және зерттеу нысанының ерекшеліктерін ескере отырып ИШҚҚЖ құрылымы ұсынылады. Содан кейін ұсынылған ИШҚҚЖ жүйесінің функционалдық блоктарын өз ара байланыстары мен жүйелі жұмысы сипатталады.

Кокстеу ракторларында өндірілетін мұнай коксы мен мұнай өнімдері буларв көлемдерін анықтау, олардың реакторлардың кіріс, режимдік параметрлеріне байцланысын идентификациялау эксперименталдық-статистикалық тәсілдер (Douglas et all., 2021; Zhao et all., 2018) негзінде жүзеге асырылады. Ал зерттеу

нысанының негізгі өнімі мұнай коксының, айқынсыздықпен сиатталатын сапа көрсеткіштері жайында ақпаратты жинау, формализациялау және өңдеу үшін эксперттік бағалау және айқын емес теориясы тәсілдері (Jorgensen et all., 2019; Orazbayev et all., 2018) қолданылады. Сонымен қатар бұл зерттеуде ИШҚҚЖ блокторы жұмыстарын жүйелендіру және түрлі сипаттағы ақпараттарды модельдеу мен шешңм қабылдацуда пайдалануға мүмкіндік беретін жүйелік талдау тәсілдемесі (Huang et all., 2019; Ivo et all., 2021) қолданылады.

Нәтижелер мен талқылау. Практикада БКҚ сияқты күрделі, айқынсыздықпен сипаттаолатын ХТЖ жұмыс режимдекрін тиімді басқару үшін ШҚТ үнемі нақты туындаған өндірістік жағдай талаптарына сәйкес адекватты шешімдердер қабылдауы тиіс. Түрлі факторларға байланымты болатын жіне ақпарат тапшылығы және йқынсыздығымен сипатталатын мұндай шешімдерді ғылыми негізделген тәсілдер мен арнаулы компьютерлік жүйелер көмегінсіз мұндай шешімдарды қабылдау мүкін емес. Бастапқы ақпараттың жетіспеушілігі және айқынсыздығын ескере отырып, ХТЖ басқару бойынша тиімді шешімдер қабылдауға арналған заманауи математикалық, эвристикалық тәсілдер негізінде құрылатын ИШҚҚЖ ШҚТ-ға шешімдерді дайындау және қабылдау процесінде көмектесетін қуатты құрал болып табылады. Сондықтан бұл жұмыстың негізгі нәтижелері БКҚ жұмыс режимдерін айқынсыздықта басқару үшін құрвлатын ИШҚҚЖ құрылымын әзірле және және оның негізгі функционалдық блоктары мен олардың жүйелі жұмысын сипаттауға бағытталған.

Ұсынылған құрылымдағы ИШҚҚЖ шешім қабылдаушылардың талдау, шығармашылық ойлау және олардың өз біліміне, тәжірибесіне және интуициясына негізделген ұтымды шешім қабылдау қабілеттері мен қазіргі заманғы компьютерлердің үлкен көлемдегі ақпаратты жылдам өңдеуге және күрделі есептеулерді тез мүмкіндіктерін біріктіреді. Сондай құрылатын ИШҚҚЖ-де кибернетиканың негізін қалаушы Н.Винердің «Адамға адамға лайық функцияларды, ал машинаға машинаға тән функцияларды беру» және «ашық жүйе» принципрері сақталатын болады. Яғни адам, біздің жағдайды ШҚТ және компьютердің функцияларын ұтымды ажырату және жүйені дамытуға, жаңа жағдайларға бейімдеуге жүзеге асыру маңызы орын алады. Сондықтан шешңм қабылдаву процесінде ШҚТ мен компьютер (жүйе) арасында тиімді диалогты ұйымдастыру үшін қажет ыңғайлы «интеллектуалды» интерфейс қарастырылады.

– ШҚТ-лардың компьютерлерден артықшылығы -қолайлы программалау ортасы негізінде шешім қабылдау есептерін шешу үшін құрастырылған модельдер мен әзірленген алгоритмдерді программалық жүзеге асыру (6-кезеңде (6-ші кезеңде);

– альтернативті таңдау кезінде қабылданған пікірлер негізінде таңдалған шешімді түсіндіру блогы әзірлеу (7-ші кезеңде);

– жағдайлар мен шарттар өзгерген кезде құрастырылған модельдердің

параметрлерін идентификациялау және баптау үшін параметрлік идентификациялау тәсілдеріне негізделген программалық пакет түрінде модель параметрлерінің идентификаторын әзірлену (8-ші кезең);

– логикалық программалау тілдеріне (мысалы, Prolog) негізделген, ШҚТ-ның жүйемен, мысалы, табиғи немесе кәсіби тілдерде ыңғайлы және таныс жұмыс режимін қамтамасыз етуге арналған «интеллектуалдандырылған» интерфейсті әзірлеу (9-ші кезең).

Жоғарыда ұсынылған ИШҚҚЖ-рін құру әдістемесінің жоғарыда сипатталған кезеңдері мен шешім қабылдау қолдау жүйесінің құрылымы мен құрамына қойылатын талаптарды талдау негізінде БКҚ-нің жұмыс режимдерін басқаруға арналған ИШҚҚЖ-нің келесі құрылымы ұсынылады (сурет 1). Ұсынылған ИШҚҚЖ құрылымында ол жүйенің функционалды блоктары арасындағы байланысты көрсететілген.

Сурет 1 – БКҚ жұмыс режимдерін басқару бойынша шешім қабылдауды қолдау жүйесінің құрылымы.

Ұсынылған ИШҚҚЖ құрылымы ашық жүйе принципі бойынша құрылған, яғни қажет болған жағдайда олардың бұл жүйе құрамындағы функциионалдық блоктарды жетілдіруге және басқа блоктарды енгізуге болады.

БКҚ жұмыс режимдерін басқаруға арналған құрылымы әзірленген ИШҚҚЖ-нің негізгі функционалды блоктарының негізгі функциялары мен сипаттамасын қарастырайық.

1) *ИШҚҚЖ пайдалушысы* – бұл нысан джұмысын басқаратын оператор-технолог, яғни ШҚТ. Пайлдаланушы, ИШҚҚЖ қолдауымен басқарылатыен нысан жұмысының ең жақсы режимін таңдайды, яғни мүмкін шешімдер жиынынан барлық шектеулердің шарттары орындалуын ескере отырып таңдалған критерийлердің оотималды (тиімді, ымыралы) мәндерін

қамтамасыз ететін шешім нұсқасын қабылдайды. ШҚТ ең жақсы шешімді қалыптасқан өндірістік жағдайға, өндіріс жоспарына, шикізаттың құрамына, нарықтағы өнім сапасына қойылатын талаптарды және т.б. ескере отырып қабылдайды. Сонымен қатар, бұл жағдайлар мен талаптар өзгеруі мүмкін, бұл ШҚТ тарапынан шешім қабылдау процесінің нақты критерийлері мен енгізілген шектеулер мен басқа да параметрлердің маңыздылығын ескеретін салмақтау коэффициенттерінің өзгеруіне әкеледі.

Шешімдерді дайындау мен қабылдаудың қарастырылған мәселелерін шешу үшін пайдаланушы ИШҚҚЖ-нің келесі блокарын қолданады: БКҚ технологиялық агргаттарының математикалық модельдері пакетін; шешім қабылдау есептерін шешу тәсілдері, соның ішінде айқынсыздықта да жұмыс жасайтын эвристикалық алгоритмдерін; сондай-ақ, қажет болған жағдайда, деректер және білімдер базасы блогы және алынған шешімді түсіндіру блогын.

Басқару ныаны параметрлері мен жұмыс жағдайы өзгерген жағдайларда, жаға жағдайға адаптациялау процесінде модельдердің параметрлерін қайа идентификациялау, баптау және білім қорын толтыру мен толықтыру кезінде ШҚТ эксперт ретінде қатысады. Ұсынылған құрылымнан көрініп тұрғандай, ШҚТ-ның аталған барлық функциялары пайдаланушы интерфейс арқылы жүзеге асырылады.

2) *«Пайдаланушы интерфейсі»* блогы. ИШҚҚЖ-нің бұл функционалдық блогы пайдалнушыға, яғни ШҚТ-ға шешімдерді дайындау және қабылдау процесінде жүйемен жұмыс жасауға, ақпарат алмасуға ыңғайлы, интерактивті режимді қамтамасыз ететін программалар жиынтығы болып табылады. Сонымен қатар бұл блок жүйенің бірқатар басқа функцияларын жүзеге асыру кезінде (мысалы, пайдаланушыға таныс форматта нәтижелерді көрсету, талдау жүргізу, нысанның интерактивті схемасымен жұмыс жасау, шешімдерді таңдау процесін визуалды бақылау және т.б.) қолданылады. ШҚТ мен жүйе арсындағы ақпарат алмасуды, диалогты ыңғайлы ету үшін ИШҚҚЖ «интеллектісін» арттыру үшін ақпарат алмасу процесі пайдаланушыға таныс табиғи немесе кәсіби тілде болуы керек.

Кәсіби тіл ретінде кәдімгі адам тілінің тым үлкен бөлігі болып табылмайтын, шешілетін мәселелердің мәнін мамандықтың пәндік саласындағы мамандардың өзара түсінуін жеңілдететін белгілі бір қасиеттері бар табиғиға жақын тілді түсінеміз. Кәсіби тілде әр ұғым мен сөз біржақты түсінілуі керек, яғни мұндай тілде синонимдер болмауы тиіс. Сонымен қатар, табиғи тілдегі сөз тіркестері мен сөйлемдермен салыстырғанда, кәсіби тілдегі сөз тіркестері мен сөйлемдер неғұрлым қатаң ережелерді қолдана отырып құрылады. Кәсіби тілді құру оңай болмаса да, мұндай «жеңілдетілген» тіл ШҚТ мен ИШҚҚЖ арасындағы тиімдірек жұмыс режиміне мүмкіндік береді және адам мен компьютер арасындағы ыңғайлы диалогты ұйымдастырудың тиімді құралы болып табылады. Сонымен бұл пайдаланушы блогы ИШҚҚЖ-

нің маңызды блоктарының бірі болып табылады, интерфейс пайдаланушы мен жүйе арасындағы байланыстың ыңғайлылығын, қарапайымдылығын және тиімділігін қамтамасыз ететіндіктен, ол ИШҚҚЖ-нін интеллект деңгейін анықтайды. Сонымен қатар ИШҚҚЖ-нің интеллект деңгейі білім базасында формализацияланған және сақталған білімнің көлемі мен мазмұнына, модельдерді құру кезінде қолданылатын эксперттердің білім деңгейіне және шешім қабылдау есептерін шешудің эвристикалық алгоритмдеріне де байланысты екенін атап өткен дұрыс.

Пайдаланушы, яғни ШҚТ мен компьютерлер арасында ақпарат алмасуға мүмкіндік беретін граикалық интерфейсті әзірлеу кезінде адамның компьютермен өз ара әрекетін ұйымдастырудың келесі негізгі талаптары ескерілді (Larichev et all., 2019).

– ИШҚҚЖ пайдаланушысы болып табылатын ШҚТ-дан ақпарат алу әдісінің дұрыстығы. Бұл талаптың сақталуын қамтамасыз ету үшін алдымен режимдері басқарылатын нысанның ШҚТ-мен талқылау және келісу ұсынылады, осылайша о5ан неғұрлым түсінікті және таныс тіл, мысалы, кәсіби бағытталған тіл, ақпарат қабылдау және шығару үшін қолданылады;

– БКҚ жұмыс режимін тиімді басқару үшін ИШҚҚЖ пайдаланушысы мен компьютер арасындағы диалогты және ақпарат алмасуды ұйымдастырудың ыңғайлылығы мен жайлылығы. Бұл талап мыналарды қамтиды: есептеу процедурасының тиімділігін қамтамасыз ету және оларды ШҚТ-ға тез және ыңғайлы түрде ұсынуға мүмкіндік беретін қажетті параметрлер мен өндірістік көрсеткіштерді анықтау; диалог процесінің әр қадамында қажетті түсініктемелерді, кеңестер мен ақпаратты алу мүмкіндігін қамтамасыз ету;

– ШҚТ мен компьютер арасындағы икемді диалог принципін жүзеге асыру. Бұл принцип диалог процесінде қай тараптың жетекші екенін дұрыс анықтауды талап етеді. Мәселені шешу сатысына байланысты кейбір жағдайларда компьютерге шешім қабылдау процесінің жалпы схемасын енгізу арқылы диалогтың жетекші тарапы ретінде компьютерді анықтап, шешім қабылдаушыны ақпарат көзі ретінде анықтаған жөн. Шешім қабылдау процесінің басқа жағдайларында немесе кезеңдерінде, керісінше, жетекші мәселені шешудің барысын анықтайтын және нәтижелерін бағалайтын ШҚТ-ны анықтап, компьютерді есептеу құралы ретінде анықтаған жөн.

ШҚТ мен компьютер арасындағы диалогты ұйымдастыру және жүзеге асыру тәртібі сәйкес ішкі мәзірлер мен әрекеттері бар мәзірге негізделген. Осылайша, әрбір мәзір өз кезегінде әртүрлі әрекеттерді жүзеге асыратын бірнеше ішкі элементтерді қамтиды. Мәзір жүйесін әзірлеу кезінде оның мазмұны мен нысандары ШҚТ-ның, яғни БКҚ ақпараттық жүйесін пайдаланушылар болып табылатын операторлары мен технологтарының талаптары мен тілектерін ескере отырып анықталады.

3) *«БКҚ – басқару нысаны математикалық модельдері» блогы.* Бұл функционалдық блок басқару нысанынның жеке аграгаттарының құрылған

модельдері мен олардың технологиялық схемаға сійкес бір пакетке біріктірілген құрылымдардың жиынтығы болып табылады. Модельдер пакетінде әр түрлі типтегі модельдер болуы мүмкін, мысалы, статистикалық, айқын емес, біріктірілген және олар бірге басқару нысанының жұмысын жүйелі модельдеуге және оптимизациялауға мүмкіндік беруі керек. Модельдер пакеті басқару нысаны жұмыс режимдерін басқару үшін қолданылатын кіріс және жұмыс параметрлерінің түрлі мәндері үшін БКҚ жұмыс сапасын бағалайтын критерийлердің мәндерін есептеу (анықтау) үшін қолданылады.

БКҚ жеке агрегаттарының математикалық модельдері модельдерді құруға қажетті қолжетімді ақпараттың сипатына байланысты әр түрлі тәсілдердің қолдану арқылы, декмпозиция принципіне сәйкес әзірленеді. Бұл принципкке сәйкес модельдер алдымен қолжетімді ақпаратқа байланысты түрлі тәсілдермен құрастырылады, содан кейін БКҚ жеке арегаттарының әзірленген түрлі типтегі модельдері кокстеу технологиялық процестің жүру схемасына сәйкес бір жүйеге, яғни модельдер пакетіне біріктіріледі.

4) «*Шешім қабылдау есептерін шешу тәсілдері, эвристикалық тәсілдер»* блогы. ИШҚҚЖ-нің бұл функционалды блогы басқару нысанының ең тиімді жұмыс режимін таңдауда шешім қабылдау есептерін шешу кезінде қолданылады және әдетте, айқын емес ортада шешім қабылдау есептерін шешудің эвристикалық алгоритмдерін қамтиды. Шешім қабылдау есептерін шешу алгоритмдері, модельдер пакетін, декректер мен білімдер базасын және логикалық қорытынды ережелерін қолдану арқылы тиімді шешімдердің жиынын анықтайды. Содан кейін ол тиімді (Парето) жиынынан ішінен таңдалған критерийлер бойынша ең жақсысын, яғни нысанның туындаған жағдайларға байланысты ең тиімдсін ШҚТ таңдайды. Сонымен соңғы, ең тиімді шешімді қабылдау пайдаланушы-ШҚТ құзіретінде болады.

5) *«Қабылдауға ұсынылған шешімді түсіндіру» блогы.* Әдетте бұл блоктың құрамында *«Логикалық қорытындалау»* функциясын жүзеге асыратын ішкі блок, немесе жеке блок қосымша өолданылады. Бұл функционалдық блок қабылданған шешімдерді қорытындылау және түсіндіру стратегиясын жүзеге асыруға арналған. Шешім қабылдау есептерін шешу алгоритмдерін және модельдер пакетін, білімдер және деректер базасы блогын пайдалана отырып, аталған блок шешім қабылдаушыларға басқарылатын нысанның жұмыс режимдерін тиімді басқару бойынша ұсыныстарды тұжырымдауға мүмкіндік береді.

Логикалық қорытындылау және таңдалған шешімді түсіндіру процедуралары шартты қорытынды жасаудың логикалық (композициялық) ережелерін пайдалана отырып, айқын емес жиындар бойынша операцияларды орындау және альтернативті таңдауларда қабылданған барлық ойлар мен шешімденрді тіркеу негізінде жүзеге асырылады. Сондай-ақ, қажет болған жағдайда, ШҚТ шығарылған ұсыныстар бойынша түсініктеме ала алады. Мұндай түсініктемелер жүргізілген логикалық қорытындының компьютерде

сақталған ізі негізінде алынады. Түсініктемелерді құру процесін жүзеге асыру үшін құрылымы мен функционалдық блоктра ұсынылған ИШҚҚЖ блогы «Шешім қабылдау есептерін шешу тәсілдері, эвристикалық тәсілдер», «Білімдер және деректер базасы» және «Пайдаланушы интерфейсі» блоктарымен байланысады.

6) *«Білімдер және деректер базасы»* блогы пайдаланушы – ШҚТ- ның, пән бойынша эксперттердің және БКҚ туралы білімдер мен деректер базасын құруға және оларды өңдеп, шешім қабылдау процесінде қолдануға арналған.

Бұл блокта сақталған ақпаратты келесі функцияларды орындауда пайдаланылады:

– «Шешім қабылдау есептерін шешу тәсілдері, эвристикалық тәсілдер» блогында альтекрнативті шешімдерді талдау және таңдауда;

– логикалық қорытындылау және алынған шешімді түсіндіру процесінде;

– «Модель параметрлерінің идентификаторы» блогында модельдерді жаңа жағдайларға адавптациялау, баптау үшін, сондай-ақ нысанның жұмысы туралы әртүрлі есептерді құрастыру кезінде.

Бұл жағдайда білім компьютерде сақталған, белгілі бір ережелерге сәйкес формализацияланған, ШҚТ, эксперт-мамандардан алынған және формализацияланған ресімделетін ақпаратты білдіреді. Интеллектуалды жүйелерде ШҚТ, эксперттерден алынған және олардың білімі, тәжірибесі мен интуициясы формализацияланған түрде көрсететін осындай ақпарат негізінде білімдер базасы құрылады. Компьютер, білім базасындағы білімге негізделген, содан кейін белгілі бір алгоритмдерді пайдалана отырып, қорытынды жасай алады.

7) *«Модельдер параметрлерінің идентификаторы» блогы,* модельдердің адекваттығын тексеруге және қажет болған жағдайда параметрлерді қайта идентификациялауға, баптьауға арналған. Өндірісте мұндай смодельдердің параметрлерін баптау, қайта идентификациялау қажеттіліктері түрлі себептерге (физикалық тұрғыдан ескіру, шикізат қасиеттерінің өзгеруі т.б.) басқару нысанының күйі, сипаты өзгергенде немесе тағы басқа өзгеріс тероге байланысты туындайды. Бұл блок параметрлік идентификациялау тәсілдерін жүзеге асыратын программалық кешен болып табылады.

Құрылымы ұсынылған БКҚ жұмыс режимдерін басқару бойынша тиімді шешім қабылдауды қолдау интеллектуалдандырылған жүйенің интеллекуалдық деңгейін арттыру үшін келесі әс-әрекеттер негізінде жүзеге асырылады:

– толық және қарама–қайшылығы жоқ білімдер базасын құру;

– тиімді шешімді таңдау кезінде құрылған білімдер базасын және ШҚТ-ның білімін, тәжңрибесін және интуициясын «интеллектісін, шығармашылық қабілетін барынша толық пайдалану;

– айқын емес ортада тиімді шешімді қабылдауға мүмкіндік беретін эвристикалық алгоритмдерді құру және қолдану;

- басқару нысанының, яғни БКҚ-ның модель құруға қажетті кейбір бастапқы декретер мен ақпараттардың күрделі өлшенуі немесе өлшенбеуі жағдайларында қосымша айқын емес ақпараттарды пайдалану арқылы нысанның модельдерін құру;

– пайдаланушы, яғни ШҚТ-ға жүйені шешім дайындау жіне қабылдау кезінде қолдануды оған түсінікті және ыңғайлы түрде жүзеге асыратынғ яғни ақпарат алмасуды табиғи не кәсіби тілде жүргізуге қабілетті интеллектуалдандырылған пайдалушы интерфейсін құру.

Ұсынылған БКҚ жұмыс режимдерін басқару бойынша тиімді шешімді таңдауды жүзеге асыратын ИШҚҚЖ «Пайдаланушы (ШҚТ) - компьютер (жүйе)» арасындағы ақпарат алмасуды шешім даярлау және қабылдау процесінде қамтамасыз ететін диалогты, итеративті режимінде жұмыс жасайды. Диалог процедурасы ақпаратты енгізу/шығару құрылғысы (пернетақта, тінтуір, монитор, дисплей) арқылы жүзеге асырылады. Автоматты ақпаратты өлшеу және жазу құрылғыларымен арнаулы нысанмен байланыс интерфейстері арқылы байланыстыруға болады. Кейбір маңызды ақпарат пен өндіріс бойынша есеп беру материалдарын қағазда алу қажет болса, принтерлер пайдаланылады. Архивтік деректер, режим парақтары, есептер және т.б. сияқты үлкен көлемді деректер қатты дискілерде және басқа сыртқы жады құралдарында сақталады. Ақпаратты компьютерлерден басқа компьютерлерге немесе құрылғыларға тасымалдау және тасымалдау желі арқылы немесе сыртқы сақтау құралдарын пайдалану арқылы жүзеге асырылады.

Сонымен түрлі өндірістік нысандардың тиімді жұмыс режимдерін таңдауда бойынша шешім қабылдау есептерін шешуде қолданатын ИШҚҚАЖ тиімділігі білімдерді формализациялау, құрылымдау, құрылған модельдер және шешім қабылдау есептерін шешу эвртистикалық алгоритмдері сапаларымен, сондай-ақ пайдаланушы интерфейсі ыңғайлылығымен анықталатыны туралы қорытында жасауға болады.

Қорытынды. БКҚ оптималды жұмыс режимдерін таңдау бойынша шешім даярлау және қабылдау процесстеріне қолдау көрсету ИШҚҚЖ құрылымы жасақталып, оның негізгі функционалдық блоктарының қызметтері мен сипаттамалары берілген.

Жұмыста тұжырымдалған зертеу мақсаты, яғни БКҚ тиімді жұмыс режимдерін таңдауға мүмкіндік беретін ИШҚҚЖ құрылымын құру және ондай жүйелердің функционалдық блоктары мен олардың жүйелі жұмысын сипаттау, қол жеткізіліп, жүзеге асырылған. Ал аталған мақсатқа жету үшін қойылған зерттеу мәндеттері толығымен шешілген. Жұмыста зерттеу Зерттеу барысында алынған нәтижелер БКҚ сияқты күрделі, нашар формализацияланатын, яғни айқынсыздықпен сипатталатын технологиялық жүйелердің жұмыс режимдерін тиімді басқару мақсатында шешімдер қабылдауда қолданылуы мүмкін.

Құрылымы ұсынылған ИШҚҚАЖ-нің ұқсас жүйелерден артықшылығы: ИШҚҚАЖ құрамына ШҚТ, эксперттер білімдері, тәжірибелері негізінде құрылған модельдер жүйесі мен өз-ара байланысқан технологиялық агрегатар кешенінің оптималды жұмыс режимін айқын емес ортада таңдау бойынша шешім қабылдау эвристикалық тәсілдері, және интеллектуалдандырылған пайдаланушы интерфейсі енетіндегінде.

БКҚ жұмыс режимдерін оптимизациялау үшін ұсынылған ИШҚҚАЖ құрылымы (сурет 1), жоғары деңгейде автоматтандырылған нысандардың, сондай-ақ автоматтандыру деңгейлері төмен нысандардың жұмыс режимдерін оптимизациялауға мүмкіндік береді. Жүйенің мұндай құрылымы математикалық қамтуды жеңіл кеңейтуге және өндірістік процессті жүйелік талдау бойынша жаңа функцияларды қосуға да ыңғайлы болып табылады.

Бұл жұмыста зерттеу барысында келесі негізгі нәтижелер алынған:

- қарастырылатын интеллектуалдандырылған шешім қабылдауды қолдау жүйелерінің басқа компьютерлің ақпараттық жүйелерден айырмашылықтары, артықшылықтары мен мүмкіндіктері тұжырымдалған;

-ИШҚҚЖ мүмкіндіктері мен артықшылықтарын қамтамасыз етуге қажетті ИШҚҚЖ құрамынына енетін негізгі функционалдық блоктард мен мұндай жүйелерді құрудың негізгі тізбекті қадамдары мен сипатталған;

-ИШҚҚЖ-ін құру процедурасының жүзеге асырудың негізгі кезеңдері ұсынылып, сипатталған;

– баяу кокстеу қондырғысы мысалында айқынсықдықпен сипатталатын күрделі ХТЖ-дің жұмыс режимдерін басқару бойынша шешім қабылдауды қолдау жүйесінің құрылымы ұсынылып, оның құрамындағы негізгі функциоеналдық блоктардың қызметтері сипатталған;

– құрылымы ұсынылған ХТЖ-дің жұмыс режимдерін басқару бойынша тиімді шешім қабылдауды қолдау интеллектуалдандырылған жүйесінң интеллекуалдық деңгейін арттыру жолдары мен әдістері қарастырылған.

Зерттеу барысынд алынғани нәтижелер баяу кокстеу қондырғысы жұмс режимдерін басқару процесінді тиімді шешімдерді қолдау интеллектуалдандырыған жүйесін құруда қолданылуда. Сонымен қатар алынған нәтижелер болашақта ақпарат жетіспеушілігі мен айқынсыздығымен сипатталатын басқа ХТЖ-дің жұмыс режимдерін басқару бойынша ИШҚҚЖ-ін құруда қажетті бейімдеу мен жетілдіру арқылы қолданылуы мүмкін.

References

Lisicyn N.V., Viktorov V.K. (2021) Himiko-tekhnologicheskie sistemy: optimizaciya i resursosberezhenie. — SPb: Mendeleev. — 2-izd. — 347 p. (in Eng.)

Czabanski R., Jezewski M., Jacek Leski J. (2017) Introduction to Fuzzy Systems. In book: Theory and Applications of Ordered Fuzzy Numbers. https://doi.org/10.1007/978-3-319-59614-3_2 (in Eng.)

Orazbayev B., Ospanov Y., Makhatova V., Salybek L. Abdugulova Z., Kulmagambetova Z. Suleimenova S., Orazbayeva K. (2023). Methods of Fuzzy Multi-Criteria Decision Making for Controlling the Operating Modes of the Stabilization Column of the Primary Oil-Refining Unit. Mathematics 11, 2820. https://doi.org/10.3390/math11132820 (in Eng.)

Zaskanov V.G., Ivanov D.Yu., Grishanov G.M. (2013) Sistemy podderzhki prinyatiya reshenii. –Samara: SGAKU im. S.P. Koroleva (nats. issled. un-t). — 175 p. (in Eng.)

Prokopenko N.Yu. (2017) Sistemy podderzhki prinyatiya reshenii. — N. Novgorod: NNGASU. — 188 p. (in Eng.)

Gil'boa I. (2017). Kak prinyat' nailuchshee reshenie? Teoriya prinyatiya reshenii na praktike. — M.: ID «Delo» RANKhiGS, — 288 c. (in Eng.)

Gelovani V.A., Bashlykov A.A., Britkov V.B., Vyazilov E.D. (2008) Intellektual'nye sistemy podderzhki prinyatiya reshenii v neshtatnykh situatsiyakh s ispol'zovaniem informatsii o sostoyanii prirodnoi sredy; Editorial URSS. — M.: — 304 p. (in Eng.)

Vadinsky O. (2018). An overview of approaches evaluating intelligence of artificial systems. Acta informatica pragensia. — No 7-1. – P. 74-103. https://elibrary.ru/item.asp?id=35423152 (in Eng.)

Osipov G.S. (2015) Metody iskusstvennogo intellekta. — M.: Fizmat. — 295 s. (in Eng.)

Orazbayev B.B., Serimbetov B.A. (2019). Metody i sistemy iskusstvennogo intellekta v ekonomike. — Almaty: Izd. Evero, — 308 p. (in Eng.)

Makarov I.M., Lokhin V.M., Man'ko S V., Romanov M.P. (2012) Iskusstvennyi intellekt i intellektual'nye sistemy upravleniya, M.: Nauka. — 336 p. (in Eng.)

Hamilton S.H., Jakeman A.J., Norton J.P. (2018) Artificial Intelligence techniques: An introduction to their use for modelling environmental systems. Mathematics and Computers in Simulation. — Vol. 78. — No 2. — P. 379–400. https://doi.org/10.1016/j.matcom.2008.01.028 (in Eng.)

Assanova B., Orazbayev B., Shangitova Z, Moldasheva Z., Orazbayeva K., Kozhakhmetova D. (2023) Development of Coke Chambers Models of Delayed Coking Unit under uncertain initial information. 7th International Symposium on Innovative Approaches in Smart Technologies (ISAS), Istanbul, Turkiye. — P. 1-5 https://doi.org/10.1109/ISAS60782.2023.10391778 (in Eng.)

Douglas A.M. and Danny A.M. (2021) Statistical Methods in Experimental Pathology: A Review and Primer. The American Journal of Pathology. — Vol. 191. — No 5. — P.784–794 https://doi.org/10.1016/j.ajpath.2021.02.009 . (in Eng.)

Karmanov F.I. Ostreykovskiy V.A. (2019) Statisticheskiye metody obrabotki eksperimental'nykh dannykh s ispol'zovaniyem paketa MathCad. — M.: Radio i svyaz'. — 187 p. ISBN: 978-5-905554-96-4. (in Eng.)

Zhao Zh.W., Wang D.H. (2018) Statistical inference for generalized random coefficient autoregressive model // Mathematical and Computer Modelling, — Vol.56. — No 3. — P.152-166. (in Eng.)

Jorgensen M. A (2019) Review of Studies on Expert Estimation of Software Development Effort. Journal of Systems and Software. — V.70. — P.37-60. (in Eng.)

Foth T., Efstathiou N., Vanderspank W.B. (2016) The use of Delphi and nominal group technique in nursing education: a review. Int. Jurnal Nurs Stud, 2019. —Vol. 60. — P. 112–120. https://doi.org/10.1016/j.ijnurstu..04.015 (in Eng.)

Divina T.V., Petrakova Ye.A., Vishnevskiy M.S. Osnovnyye metody analiza ekspertnykh otsenok. Journal of Economy and Business. — Vol. 7. — P. 42-44. (in Eng.)

Ryzhov A.P. (2021) Fuzzy set theory and its applications. -Moscow: Moscow State University. 2nd edition. — 285 p. ISBN: 978-3-540-70777-6. (in Eng.)

ACADEMIC SCIENTIFIC JOURNAL OF COMPUTER SCIENCE
ISSN 1991-346X
Volume 4. Namber 356 (2025). 26–39

https://doi.org/10.32014/2025.2518-1726.381

ӘОЖ 004.942. 001.57

© *Zh.T. Abildayeva[1], R.K. Uskenbayeva[1], G.S. Beketova[2], N.B. Konyrbaev[3], S.B. Seydazimov[1,4], 2025.

[1]K.I. Satpayev Kazakh National Research Technical University, Almaty, Kazakhstan;
[2]G. Daukeev Almaty University of Energy and Telecommunications, Almaty, Kazakhstan;
[3]Kyzylorda University named after Korkyt Ata, Kyzylorda, Kazakhstan;
[4]Kazakh National Women's Pedagogical University, Almaty, Kazakhstan.
E-mail: Zhansaya_A@mail.ru

MULTI-CRITERION OPTIMIZATION OF ADVERTISING BUDGET ALLOCATION IN THE AGRO-INDUSTRIAL COMPLEX BASED ON NSGA-III ALGORITHM

Abildaeva Zhansaya — doctoral student, K.I. Satpaev Kazakh National Research Technical University, Almaty, Kazakhstan,
E-mail: Zhansaya_A@mail.ru, https://orcid.org/0000-0002-2637-0443;
Uskenbaeva Raisa — Professor, PhD. K.I. Satpaev Kazakh National Research Technical University, Almaty, Kazakhstan,
E-mail: r.k.uskenbaeva@satbaev.university, https://orcid.org/0000-0002-8499-2101;
Beketova Gulzhanat — PhD. G. Daukeev Almaty University of Energy and Telecommunications, Almaty, Kazakhstan,
E-mail: G.Beketova@aues.kz, https://orcid.org/0000-0001-7160-1514;
Konyrbayev Nurbek — Professor, PhD, Korkyt Ata Kyzylorda University, Kyzylorda, Kazakhstan,
E-mail: nurbek@korkyt.kz, https://orcid.org/0000-0002-8788-4149;
Seydazimov Syrym — G. Daukeev Almaty University of Energy and Telecommunications; Kazakh National Women's Pedagogical University, Almaty, Kazakhstan,
E-mail: syreken.ss@gmail.com, https://orcid.org/0000-0002-2955-8388.

Abstract. This study explores the application of the NSGA-III multi-criteria evolutionary algorithm to optimize advertising budget allocation in the agro-industrial complex (AIC) of Kazakhstan. Given increasing market competition, rapid digitalization, and evolving consumer behavior, AIC enterprises must modernize their advertising processes. Key challenges include improving the efficiency of advertising campaigns, expanding accurate target audience coverage, and reducing marketing costs. In this context, the use of NSGA-III for multi-criteria optimization provides an effective tool for balancing conflicting objectives and supporting decision-making under uncertainty—an essential requirement for the agricultural

sector. A mathematical model of advertising budget distribution was developed, incorporating three objective functions: efficiency, cost, and coverage. The NSGA-III algorithm was employed to generate a set of Pareto-optimal solutions, whose quality was evaluated using hypervolume (HV = 0.876), diversity, and distribution uniformity metrics. The simulation confirmed the existence of trade-offs: solutions maximizing efficiency and coverage also tended to increase total costs. Another significant finding was the strong performance of digital advertising channels. Optimal strategies allocated up to 39.8% of the marketing budget to digital media, demonstrating their high effectiveness in the modern agricultural market. Overall, the research indicates that NSGA-III is a powerful and reliable tool for addressing complex multi-criteria optimization tasks in agromarketing. The results may serve as a basis for building data-driven marketing strategies and improving managerial decision-making within agricultural enterprises.

Keywords: agro-industrial complex, multi-criteria optimization, NSGA-III, marketing strategies, Pareto front, digitalization

© *Ж.Т. Әбілдаева[1], Р.К. Ускенбаева[1], Г.С. Бекетова[2], Н.Б. Қоңырбаев[3], С.Б. Сейдазимов[1,4], 2025.

[1]Қ.И. Сәтбаев атындағы Қазақ ұлттық зерттеу техникалық университеті, Алматы, Қазақстан;

[2]Ғ. Дәукеев атындағы Алматы энергетика және телекоммуникация университеті, Алматы, Қазақстан;

[3]Қорқыт Ата атындағы Қызылорда университеті, Қызылорда, Қазақстан;

[4]Қазақ ұлттық қыздар педагогикалық университеті, Алматы, Қазақстан.

E-mail: Zhansaya_A@mail.ru

NSGA-III АЛГОРИТМІ НЕГІЗІНДЕ АГРОӨНЕРКӘСІПТІК КЕШЕНДЕГІ ЖАРНАМАЛЫҚ БЮДЖЕТТІ БӨЛУДІ КӨП КРИТЕРИАЛДЫ ОҢТАЙЛАНДЫРУ

Әбілдаева Жансая — докторант, Қ.И. Сәтбаев атындағы Қазақ ұлттық зерттеу техникалық университеті, Алматы, Қазақстан,
E-mail: Zhansaya_A@mail.ru, https://orcid.org/0000-0002-2637-0443;
Ускенбаева Раиса — профессор, Phd, Қ.И. Сәтбаев атындағы Қазақ ұлттық зерттеу техникалық университеті, Алматы, Қазақстан,
E-mail: r.k.uskenbaeva@satbaev.university, https://orcid.org/0000-0002-8499-2101;
Бекетова Гүлжанат — Phd. Ғ. Дәукеев атындағы Алматы энергетика және телекоммуникация университеті, Алматы, Қазақстан,
E-mail: G.Beketova@aues.kz, https://orcid.org/0000-0001-7160-1514;
Қоңырбаев Нұрбек — профессор, Phd, Қорқыт Ата Атындағы Қызылорда Университеті, Қызылорда, Қазақстан,
E-mail: nurbek@korkyt.kz, https://orcid.org/0000-0002-8788-4149;
Сейдазимов Сырым — Ғ. Дәукеев атындағы Алматы энергетика және телекоммуникация университеті, Қазақ ұлттық қыздар педагогикалық университеті, Алматы, Қазақстан,
E-mail: syreken.ss@gmail.com, https://orcid.org/0000-0002-2955-8388.

Аннотация. Бұл зерттеуде Қазақстанның агроөнеркәсіптік кешенінде (АӨК) жарнама бюджетін бөлуді оңтайландыру үшін NSGA-III көп критерийлі эволюциялық алгоритмін қолдану қарастырылады. Қазіргі жағдайда АӨК кәсіпорындары нарықтағы жоғары бәсекелестікке, цифрландырудың артуына және тұтынушылардың мінез-құлқының өзгеруіне байланысты жарнамалық бизнес процесстері де бейімделуі тиіс. Осыған байланысты жарнамалық компаниялардың тиімділігін арттыру, мақсатты аудиторияны дұрыс қамтуды кеңейту және шығындарды оңтайландыру сияқты міндеттері аса өзекті болып отыр. Осы зерттеу контекстінде ауылшаруашылық маркетингінде көп критериалды оңтайландыру үшін NSGA-III эволюциялық алгоритмін пайдалану қарастырылады. Бұл тәсіл бізге әртүрлі факторларды теңестіруге және белгісіздік жағдайында шешім қабылдауға байланысты мәселелерді тиімді шешуге мүмкіндік береді, бұл Қазақстанның ауыл шаруашылығы бизнесі үшін өте маңызды екені сөзсіз. Зерттеу барысында жарнамалық бюджетті бөлудің математикалық моделі құрылып, тиімділік, шығындар және қамту деңгейі сияқты негізгі мақсатты функциялар анықталды. NSGA-III алгоритмін қолдану арқылы Парето-оңтайлы шешімдер жиыны алынып, олардың сапасы гиперкөлем (HV = 0,876), әртүрлілік және таралу біркелкілігі сияқты метрикалар арқылы бағаланды. Модельдеу нәтижелері критерийлер арасындағы айқын ымыраны көрсетті: жарнама тиімділігі мен қамту максималданған сайын шығындардың тең дәрежеде өсуі байқалды. Сонымен қатар, сандық жарнама арналарының басымдығы анықталды — оптималды шешімдерде оларға бюджет үлесінің 39,8%-ға дейін бөлінуі олардың жоғары тиімділігін дәлелдейді. Алынған нәтижелер NSGA-III алгоритмінің агромаркетингтегі күрделі көп критерийлі міндеттерді шешуде жоғары әлеуетке ие екенін көрсетеді. Зерттеу нәтижелері АӨК кәсіпорындарының маркетингтік стратегияларын құру мен басқару шешімдерін қолдау үшін қолданылуы мүмкін.

Түйін сөздер: агроөнеркәсіп кешені, көп критериалды оңтайландыру, NSGA-III, маркетингтік стратегиялар, Парето фронты, цифрландыру

© *Ж.Т. Абилдаева[1], Р.К. Ускенбаева[1], Г.С. Бекетова[2], Н.Б. Конырбаев[3], С.Б. Сейдазимов[1,4], 2025.
[1]Казахский национальный исследовательский технический университет им. К.И. Сатпаева, Алматы, Казахстан;
[2]Алматинский университет энергетики и связи им. Г. Даукеева, Алматы, Казахстан;
[3]Кызылординский университет имени Коркыт Ата, Кызылорда, Казахстан;
[4]Казахский национальный женский педагогический университет
E-mail: Zhansaya_A@mail.ru

МНОГОКРИТЕРИАЛЬНАЯ ОПТИМИЗАЦИЯ РАСПРЕДЕЛЕНИЯ РЕКЛАМНОГО БЮДЖЕТА В АПК НА ОСНОВЕ АЛГОРИТМА NSGA-III.

Абилдаева Жансая — докторант, К.И. Казахский национальный исследовательский технический университет им. Сатпаева, Алматы, Казахстан,
E-mail: Zhansaya_A@mail.ru, https://orcid.org/0000-0002-2637-0443;
Ускенбаева Раиса — профессор, PhD, Казахский национальный исследовательский технический университет им. К.И. Сатпаева, Алматы, Казахстан,
E-mail: r.k.uskenbaeva@satbaev.university, https://orcid.org/0000-0002-8499-2101;
Бекетова Гульжанат — PhD, Алматинский университет энергетики и связи им. Г. Даукеева, Алматы, Казахстан,
E-mail: G.Beketova@aues.kz, https://orcid.org/0000-0001-7160-1514;
Конырбаев Нурбек — профессор, PhD, Кызылординский университет имени Коркыт Ата, Кызылорда, Казахстан,
E-mail: nurbek@korkyt.kz, https://orcid.org/0000-0002-8788-4149;
Сейдазимов Сырым — Алматинский университет энергетики и связи им. Г. Даукеева; Казахский национальный Женский педагогический университет, Алматы, Казахстан,
E-mail: syreken.ss@gmail.com, https://orcid.org/0000-0002-2955-8388.

Аннотация. В исследовании рассматривается применение многокритериального эволюционного алгоритма NSGA-III для оптимизации распределения рекламного бюджета в агропромышленном комплексе (АПК) Казахстана. В условиях высокой конкуренции, усиления цифровизации и изменения потребительского поведения предприятия АПК вынуждены адаптировать свои рекламные бизнес-процессы. Такие задачи, как повышение эффективности рекламных кампаний, расширение корректного охвата целевой аудитории и оптимизация затрат, приобретают особую актуальность. В ходе исследования создана математическая модель распределения бюджета, определены основные целевые функции (эффективность, затраты, охват). С помощью NSGA-III получен набор Парето-оптимальных решений, оценённых по гиперобъёму (HV = 0,876), разнообразию и равномерности распределения. Моделирование показало наличие чётких компромиссов между критериями: максимизация охвата и эффективности сопровождается увеличением

затрат. Также подтверждено доминирование цифровых рекламных каналов — им может выделяться до 39,8% бюджета, что демонстрирует их высокую эффективность. Полученные результаты свидетельствуют о большом потенциале NSGA-III для решения многокритериальных задач агромаркетинга. Они могут использоваться для разработки маркетинговых стратегий и поддержки управленческих решений аграрных предприятий.

Ключевые слова: агропромышленный комплекс, многокритериальная оптимизация, NSGA-III, маркетинговые стратегии, фронт Парето, цифровизация

1. Кіріспе.

Ауыл шаруашылығы әлемнің кез келген елінің экономикасында орасан зор рөл атқарады, ал Қазақстан (немесе ҚР) бұдан тыс емес, өйткені агроөнеркәсіптік кешен (АӨК) азық-түлік қауіпсіздігінің негізгі көзі және маңызды экспорттық сектор болып табылады. Дегенмен, соңғы жылдары Қазақстан Республикасының агроөнеркәсіп кешені бәсекелестіктің күшеюі, өндіріс тиімділігін арттыру және өнімді бөлуді оңтайландыру қажеттілігі сияқты бірқатар міндеттерге тап болды (Шуленбаева т.б., 2011; Омарханова т.б., 2022; Мадиярова т.б., 2020; Сыздықов т.б., 2020). Сондықтан Қазақстанның агроөнеркәсіп кешеніндегі маркетинг пен жарнаманың цифрлық әдістерін математикалық және кибернетикалық әдістерді қолдана отырып зерттеудің өзектілігі бірқатар негізгі факторларға байланысты. Атап айтқанда, жаһандану және экономиканы цифрландыру жағдайында Қазақстан ауыл шаруашылығы әдістері мен процестерін жаңғырту қажеттілігіне тап болып отыр. Ал машиналық оқыту (ML) және NSGA-II, NSGA-III сияқты көп критериалды оңтайландырудың эволюциялық есептеу алгоритмдері сияқты жаңа технологияларды пайдалану маркетингтік процестерді басқаруды жақсартуға, бәсекеге қабілеттілікті арттыруға және агроөнеркәсіптік кешенде ресурстарды бөлуді оңтайландыруға бірегей мүмкіндіктер береді. Біздің ойымызша, Қазақстанның агроөнеркәсіп кешеніндегі маркетинг пен жарнаманы цифрландыру тұтынушылармен өзара іс-қимылды жақсартып қана қоймай, нарықтық үрдістерді дәлірек болжауды қамтамасыз етеді, сайып келгенде, тауарлар мен қызметтердің сапасын арттырады (Шуленбаева т.б., 2011; Блинова, 2015). Ауыл шаруашылығы маркетингіндегі цифрлық әдістерді зерттеу Қазақстанның тұрақты және бәсекеге қабілетті аграрлық секторын қалыптастыру үшін өзекті деп есептейміз, бұл тез өзгеретін нарықтық шындық жағдайында өте маңызды.

Агромаркетинг – бұл агроөнеркәсіптік сектордағы жекелеген кәсіпорындар деңгейінде де, мемлекеттік макродеңгейде де нарықтық процестерді реттейтін және талдау жасайтын ауыл шаруашылығы нарығындағы қызметті басқару жүйесі. Бірқатар зерттеушілердің (Першуков т.б., 2019; Пономаренко т.б., 2016) пікірінше, агромаркетинг нарықтың әртүрлі қатысушылары

арасында өзара тиімді алмасу механизмдерін ұсынуға, сонымен қатар тұтынушылардың қажеттіліктерін анықтауға және оларды қанағаттандыру жолдарын іздеуге жағдай жасауға қабілетті. Негізінде агромаркетинг ауыл шаруашылығы нарығын талдауды, оның динамикасын болжауды, өнімге сұранысты қалыптастыруды және олардың таралуын ынталандыруды, сайып келгенде, ауыл шаруашылығы өнімдерін өндіруге, өткізуге және нарықта жылжытуға қатысты негізделген шешімдер қабылдауға мүмкіндік беруді қамтиды, сондай-ақ агроөнеркәсіптік кешендегі (бұдан әрі Агроөнеркәсіптік кешендегі) нарықтық жағдайды бағалау және болжау мүмкіндіктерін береді.

Қазіргі заманғы ауыл шаруашылығы кәсіпорындары, әдетте, өзін-өзі реттемейді (Гарридо т.б., 2000; Козловский т.б., 2019), сондықтан олар ресурстармен және ақпараттармен алмасып, сыртқы маркетингтік ортамен үнемі әрекеттесіп отырады. Олардың жұмыс істеуі мен өмір сүруі сыртқы ортаның әсеріне байланысты, ал ауылшаруашылық кәсіпорындары өз қызметін сақтау үшін маркетингтік ортаның өзгерістеріне бір мезгілде өз мүмкіндіктері шегінде әсер ете отырып, бейімделуі керек. Ауыл шаруашылығы маркетингінің теориясы мен практикасы (Шуленбаева т.б., 2011; Мадиярова т.б., 2020) соңғы жылдары қарқынды дамып келеді, бұл ең алдымен ауыл шаруашылығы және азық-түлік нарығындағы бәсекелестіктің күшеюімен, сондай-ақ Қазақстан Республикасы халқының өсуімен және өмір сүру деңгейінің жақсаруымен, әртүрлі салалар үшін ауыл шаруашылығы шикізатының маңыздылығының артуы, ғылым мен өндірістік күштердің дамуымен байланысты. Қазақстандағы ауыл шаруашылығы кәсіпорындарының жұмыс істеуінің қазіргі тәжірибесі өндірісті ұйымдастырудың тиімділігі, өнімнің тұтынушылардың күтулеріне сәйкестігі, сондай-ақ кәсіпорында өнімді бөлу мен өткізудің сауатты және жақсы үйлестірілген жүйесі ауыл шаруашылығы тауарын өндіруші мен оның қызметкерлерінің қаржылық жағдайына тікелей әсер ететінін растайды (Блинова, 2015).

Осы зерттеу контекстінде ауылшаруашылық маркетингінде көп критериалды оңтайландыру үшін NSGA-III эволюциялық алгоритмін пайдалану қарастырылады. Бұл тәсіл бізге әртүрлі факторларды теңестіруге және белгісіздік жағдайында шешім қабылдауға байланысты мәселелерді тиімді шешуге мүмкіндік береді, бұл Қазақстанның ауыл шаруашылығы бизнесі үшін өте маңызды екені сөзсіз.

Зерттеу материалдары.

Соңғы онжылдықтарда маркетингтік стратегияларды, оның ішінде агроөнеркәсіптік кешен үшін оңтайландыру мәселелері зерттеушілердің айтарлықтай назарын аударды, бұл бәсекелестіктің күшеюі және менеджмент жүйелерінің сапасы мен бейімделуіне талаптардың өсуі жағдайында маркетингті басқару тиімділігін арттыру қажеттілігімен байланысты. Ғылыми әдебиеттерде осы саладағы зерттеудің бірнеше негізгі бағыттарын бөліп көрсетуге болады. Негізгі бағыттардың бірі маркетингтік стратегияларды

оңтайландырудың математикалық әдістерін қолданумен байланысты (Першуков т.б., 2019; Пономаренко т.б., 2016; Гарридо т.б., 2000; Козловский т.б., 2019). Мысалы, Першуков т.б., (2019) және Карнадуд О.С. (2019) белгісіздік жағдайында математикалық тәсілдердің практикалық қолданылуын көрсететін бюджеттік ресурс шектеулерін ескере отырып, маркетингтік стратегияларды оңтайландырудың математикалық үлгілерін ұсынады. Зерттеудің тағы бір маңызды бағыты - экономикалық және математикалық модельдерге негізделген деректерді талдау және болжау әдістерін пайдалану. Сонымен, Блинова Ю.Ю. (2015) агроөнеркәсіп кешеніндегі нарықтық тенденцияларды болжау дәлдігін жақсарту үшін экономикалық және математикалық модельдеу әдістерін қолдануды зерттейді, бұл неғұрлым негізделген басқару шешімдерін қабылдауға мүмкіндік береді.

Маркетингтік стратегияларды тарату арналарымен біріктіруге ерекше көңіл бөлінеді. Мысалы, зерттеу авторлары (Ковальчук т.б., 2019) персоналды талдау және өнімді тарату арналарын бағалау негізінде маркетингті басқару стратегияларын таңдау модельдерін ұсынды, бұл авторлардың пікірінше, маркетингтік стратегияның көп компонентті сипатын және оның кәсіпорынның қаржылық нәтижелеріне әсерін ескеруге мүмкіндік береді. Яғни, мәні бойынша көп критерийлі мәселе қойылуда.

Сондай-ақ, ауылшаруашылық маркетингінде шешім қабылдауды қолдау үшін заманауи машиналық оқыту технологияларын қолданудың өзектілігін атап өткен жөн. Qiu J. (2021) және Hung-Yi T. (2024) сияқты зерттеулер ауыл шаруашылығы кәсіпорындарын басқаруда интеллектуалды жүйелерді қолдану перспективаларын ашатын үлкен деректерді талдау негізінде маркетингтік стратегияларды әзірлеу үшін нейрондық желілер мен оңтайландыру алгоритмдерін пайдалануды көрсетеді.

Дегенмен, қазіргі уақытта агроөнеркәсіптік кешендегі маркетингтік стратегияларды оңтайландыру міндеттері үшін арнайы NSGA-III қолдану бойынша ауқымды зерттеулер жүргізілмегенін атап өткен жөн. NSGA-III алгоритмі қалалық агломерациялардағы қалдықтарды басқару (Qu et al., 2019) және сауда стратегияларына негізделген көп агенттік жүйелерде тапсырмаларды бөлу (Yang et al., 2025) сияқты басқа қолданбаларда өзінің тиімділігін көрсеткенімен (Yang et al., 2025), оны қолдану белгілі бір контексте мәдени зерттеулерді жасайды. біздің зерттеуімізде толтыру орынды болатын ғылыми олқылықтар.

Алдыңғы зерттеулерді талдауды қорытындылай келе, ауылшаруашылық маркетингін көп критериалды оңтайландыруда эволюциялық алгоритмдер мен ML әдістерін қолдану, біздің ойымызша, стратегиялық жоспарлаудың тиімділігін және тез өзгеретін нарық жағдайларына бейімделуін айтарлықтай арттыратынын атап өтеміз. Сондықтан бұл зерттеу аграрлық сектордағы кәсіпорындар үшін оңтайлы маркетингтік стратегияларды және жарнаманы тұжырымдау үшін эволюциялық модельдеу әдістерін, атап айтқанда NSGA-

III-ді одан әрі әзірлеуге және қолдануға бағытталған.

Зерттеудің мақсаты мен міндеттері.

Қазақстанның агроөнеркәсіп кешеніндегі маркетингтік стратегияларды оңтайландырудың NSGA-III алгоритмі негізінде көп критериалды оңтайландыру әдісін әзірлеу.

Зерттеу мақсатына жету үшін келесі міндеттер шешілді.

Агроөнеркәсіптік кешендегі маркетингтік бюджетті бөлу проблемасының мақсатты функциялары мен шектеулерін формализациялау.

Агроөнеркәсіптік маркетингтің ерекшеліктерін ескере отырып, мәселені шешу үшін NSGA-III алгоритмін бейімдеу.

Гиперкөлем (HV), диверсификация және таралу біркелкілігі (Spacing) метрикасы арқылы ерітінділердің сапасын бағалаумен есептеу экспериментін жүргізу. Сондай-ақ алынған Парето-оңтайлы шешімдерді талдау, оның ішінде бюджеттерді жылжыту арналары бойынша бөлу және критерийлер арасындағы ымыраға келу.

Әдістер мен модельдер.

Агроөнеркәсіптік кешеннің маркетингтік саясатын көп критериалды оңтайландыру барысында қарастырылатын критерийлерді ресімдейік.

1. Жарнамалық науқанның тиімділігі (f_1)– өнімді өткізу көлемінің ұлғаюы (мысалы, ауылшаруашылық кәсіпорнының табысының немесе кірісінің артуы).

2. Маркетинг және жарнама шығындары(f_2)–жарнамалық науқандарға және ауыл шаруашылығы өнімдерін жылжытуға жұмсалған жалпы шығындар.

3. Мақсатты аудиторияны қамту деңгейі(f_3)– жарнамалық науқан арқылы жеткен мақсатты аудиторияның үлесі.

4. Жарнамалық стратегияның тұрақтылығы(f_4)–ұзақ уақыт кезеңіндегі әсердің тұрақтылығы (мысалы, маусымдық ауытқулардың төмендеуі).

5. Сату арналарымен интеграция(f_5)–маркетинг күштері мен тарату арналары арасындағы үйлестіру деңгейі (мысалы, онлайн платформалар немесе ауыл шаруашылығы кәсіпорындарының серіктес желілері арқылы тапсырыстар санының артуы).

Критерийлердің әрқайсысын мақсаттық функция түрінде талдап, формалдап көрейік.

Содан кейін [2, 3, 5, 6, 9, 11] зерттеулерге сүйене отырып, біз осы формада мақсаттық функцияларды жаза аламыз.

$$f_1 = \frac{\Delta P}{R} \times 100\%, \qquad (1)$$

ΔP – ауылшаруашылық кәсіпорнының өніміне жарнамалық науқан жүргізгеннен кейінгі пайданың артуы.

R – жарнамалық науқанның жалпы шығындары.

Яғни, мән соғұрлым жоғары болады $\left(f_1\right)$, соғұрлым жарнама тиімдірек болады.

Критерий $\left(f_2\right)$ – жарнамалық қызметке арналған жалпы бюджетті көрсетеді. Оны келесідей орнатуға болады:

$$f_2 = \sum_{i=1}^{N} C_i, \qquad (2)$$

C_i – і-ші арнадағы (интернет, БАҚ және т.б.) жарнамалық шығындар;

N – жарнамалық арналардың саны.

Оңтайландыру осы критерийді азайтуды қамтиды.

Критерий $\left(f_3\right)$ – жарнамалық науқан жеткен мақсатты аудиторияның пайызын көрсетеді, яғни.

$$f_3 = \frac{A_{reach}}{A_{total}} \times 100\%, \qquad (3)$$

A_{reach} – аудиторияға жетті,

A_{total} жалпы мақсатты аудитория.

Критерий $\left(f_4\right)$ уақыт бойынша жарнама әсерінің тұрақтылығын көрсетеді, яғни.

$$f_4 = 1 - \frac{\sigma\left(R\left(t\right)\right)}{\mu\left(R\left(t\right)\right)}, \qquad (4)$$

$R\left(t\right)$ – белгілі бір уақытта сатудың өсуі,

σ – белгілі бір кезеңдегі сатылымның стандартты ауытқуы,

μ – сатудың орташа өсу қарқыны.

Критерий $\left(f_5\right)$ маркетинг және сату әрекеттерін үйлестіруді сипаттайды, яғни.

$$f_5 = \frac{S_{online} + S_{offline}}{S_{total}} \times 100\%, \qquad (5)$$

S_{total} – тапсырыстардың жалпы саны,

S_{online} – онлайн арналар арқылы тапсырыстар саны,

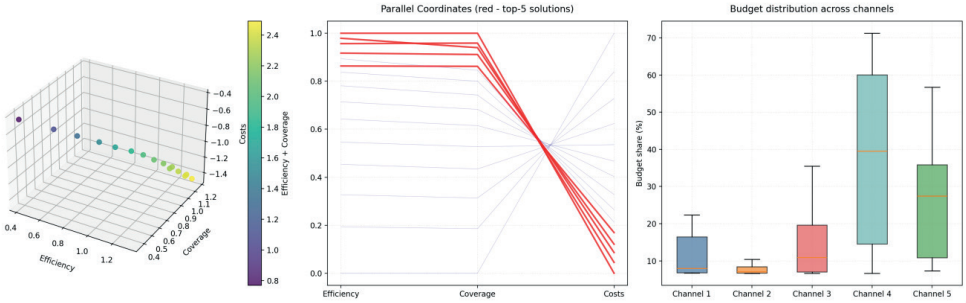
$S_{offline}$ – желіден тыс арналар арқылы тапсырыстар саны.

Критерийді барынша арттыру маркетинг пен сатудың оңтайлы интеграциясын білдіреді. Сондықтан біз келесі мақсаттарды қанағаттандыратын ымыраға келу шешімін табуымыз керек: барынша арттыру $\left(f_1, f_3, f_3, f_5\right)$ және азайту $\left(f_1\right)$. Бұл зерттеуде есептің математикалық

тұжырымы тек үш мақсатты функцияны – тиімділікті барынша арттыруды қамтыды (f_1), аудиторияның қолжетімділігін арттыру (f_3) және шығындарды азайту (f_2), сондай-ақ әрбір арнаға ең төменгі жарнаға және жалпы жарнама бюджетіне екі шектеу.

Зерттеу нәтижелері мен талқылау

Көпкритериалды оңтайландыру принциптеріне сәйкес мақсаттық функциялар $(f_1),(f_2),\ (f_3)$ оңтайландыру бағыттарын теңестіру үшін өзгертілді. Python кодында тиімділік пен қамту минус белгісімен көрсетілген. $(F[i,0] = -np.sum(x * self.channels[:,0]))$, бұл pymoo құрылымының стандартты тәжірибесіне байланысты болды, мұнда барлық мәселелер әдепкі бойынша азайту мәселелері ретінде тұжырымдалады. Нәтижелерді алғаннан кейін $(front = -res.F)$ дұрыс түсіндіру үшін мәндердің инверсиясы орын алады. Шығындар оң болып қалады, өйткені оларды азайту мәселенің бастапқы тұжырымына сәйкес келеді. Осылайша, графиктердегі теріс құн мәндері бастапқы оң мәндерді көрсететін визуализация артефакті болып табылады. Бұл нәтижелер кестесіндегі деректерді талдаумен де расталады, VS Code консолінде модельдеу нәтижелерінің шығысын көрсететін 2-суретті қараңыз (мысалы, -1,46 шығындар 1,46 шартты бірліктердің нақты шығындарына сәйкес келеді). Шешім қабылдау процесінде алынған шешімнің сапасын бағалау маңызды болды, атап айтқанда, келесі нәтижелер алынды. Гиперкөлем (HV = 0,876) және әртараптандыру (Әртүрлілік = 0,561) көрсеткіштері мақсатты көрсеткіштерден асып түсетін Парето фронтының жақсы сапасын көрсетеді. Аралық көрсеткішінің біршама асып кетуі (0,579 мақсатқа қарсы <0,5) шешімдерді бөлудің орташа біркелкі еместігін көрсетеді, бұл қайшы критерийлері бар мәселелерге тән. 2-суреттегі ең жақсы 5 шешім критерийлер арасындағы ымыраны көрсетеді, яғни тиімділіктің максималды мәндері (1.32) және қамту (1.17) теориялық күтулерге сәйкес жоғары шығындармен (~1.46) қол жеткізіледі.



Сурет 1 – Зерттеу нәтижелері

```
  489 |    58680 |     15 | 0.000000E+00 | 0.000000E+00 | 0.0014190961 |            f
  490 |    58800 |     15 | 0.000000E+00 | 0.000000E+00 | 0.0016605158 |            f
  491 |    58920 |     15 | 0.000000E+00 | 0.000000E+00 | 0.0017721305 |            f
  492 |    59040 |     15 | 0.000000E+00 | 0.000000E+00 | 0.0017780855 |            f
  493 |    59160 |     15 | 0.000000E+00 | 0.000000E+00 | 0.0020856014 |            f
  494 |    59280 |     15 | 0.000000E+00 | 0.000000E+00 | 0.0023992739 |            f
  495 |    59400 |     15 | 0.000000E+00 | 0.000000E+00 | 0.0023992739 |            f
  496 |    59520 |     15 | 0.000000E+00 | 0.000000E+00 | 0.0023930396 |            f
  497 |    59640 |     15 | 0.000000E+00 | 0.000000E+00 | 0.0030698042 |            f
  498 |    59760 |     15 | 0.000000E+00 | 0.000000E+00 | 0.0000241417 |            f
  499 |    59880 |     16 | 0.000000E+00 | 0.000000E+00 | 0.0032191525 |            f
  500 |    60000 |     15 | 0.000000E+00 | 0.000000E+00 | 0.0008188012 |            f

=== QUALITY METRICS ===
Hypervolume: 0.876 (target >0.7)
Spacing: 0.579 (target <0.5)
Diversity: 0.561

=== TOP-5 SOLUTIONS ===

Solution #1 (score: -0.930):
Budgets: [0.596 0.1   0.29  0.1   0.412]
Efficiency: 1.32 | Coverage: 1.17 | Costs: -1.46
Budget allocation (%):
  Channel 1: 39.8%
  Channel 2: 6.7%
  Channel 3: 19.3%
  Channel 4: 6.7%
  Channel 5: 27.5%

Solution #2 (score: -0.934):
Budgets: [0.51  0.1   0.339 0.115 0.402]
```

Сурет 2 – VS Code консоліндегі зерттеу нәтижелерін шығару

Сандық арналардың (ең жақсы шешімде 39,8%) және оқиғалардың (5-ші шешімде 56,7%) басым рөлі олардың жоғары тиімділігіне және бастапқы деректерде қамтылуына сәйкес келеді. Нәтижелердің Парето оңтайлылығының қағидаттарына сәйкестігі мыналармен расталады: 1) 3D графигіндегі тиімділік/қамту және шығындар арасындағы антикорреляция (1-суретте сол жақта); 2) параллель координаталардағы көп бағытты тенденциялар (1-суреттегі ортаңғы графиктер); 3) шектеулерді сақтау (барлық бюджеттер ≥0,1, сома ≤1,5). Нормаланған құрамдас ұпайлар (-0,930-дан -0,970-ке дейін) салмақтық факторларды (тиімділік үшін 0,6, қамту үшін 0,4) және айыпталушы шығындар коэффициентін ескере отырып, шешімдерді дұрыс дәрежелейді. Алынған нәтижелерді талдай отырып, NSGA-III алгоритмін қолдану маркетингтік талаптарға жауап беретін, әсіресе ауыл шаруашылығы кәсіпорындары үшін басым емес шешімдердің репрезентативті жиынтығын алуға мүмкіндік берді деп айта аламыз. Ұсынылып отырған үлгі тапсырманың көп критерийлік сипатын ескере отырып, агроөнеркәсіптік кешендегі жарнамалық бюджетті бөлу бойынша басқару шешімдерін қабылдау үшін негіз бола алады. Кейінгі зерттеулерде біз алынған нәтижелерді баламалы NSGA-II алгоритмінің жұмысымен, сондай-ақ жасанды нейрондық желілер сияқты машиналық оқыту әдістерімен салыстыруды жоспарлап отырмыз.

Қорытынды

Зерттеу барысында келесі негізгі нәтижелер алынды.

NSGA-III алгоритмі агроөнеркәсіптік кешеннің маркетингтік бюджетін көп критериалды оңтайландыру мәселесін шешуде жоғары тиімділік

көрсеткені анықталды, ол метрикамен расталады (HV = 0,876, Diversity = 0,561). Кибернетикалық модельдеу барысында қарастырылып отырған келесі критерийлер арасында ымыраға келу белгіленді: жарнама тиімділігін барынша арттыру (1,32) және қамту (1,17) жалпы Паретоның оңтайлылық принципіне сәйкес келетін жоғары шығындармен (~1,46) қол жеткізіледі. Цифрлық арналар мен оқиғалар оңтайлы шешімдерде ең үлкен мәнді көрсетті (сәйкесінше ауыл шаруашылығы кәсіпорындарының өнімдерін жарнамалауға арналған бюджеттің 39,8% және 56,7%), бұл олардың бастапқы деректердегі жоғары тиімділігіне сәйкес келеді. Графиктердегі теріс құн мәндері pymoo-да инверттелген мақсатты функциялардың визуализациясының артефакті болып табылады және нәтижелердің мағыналы интерпретациясын бұрмаламайтыны көрсетілген. Мақалада ұсынылған модель агроөнеркәсіптік кешендегі басқару шешімдерін қолдау үшін пайдаланылуы мүмкін екендігі көрсетілді және одан әрі зерттеулер NSGA-III-ны балама әдістермен (NSGA-II, нейрондық желілер) салыстыруға бағытталуы керек.

Әдебиет

Шуленбаева Ф.А., Каменова М.Ж., Дарибаева А.К., Шуленбаева К.С. (2011) Вопросы организации и развития маркетинга в аграрном секторе Казахстана. Научные Записки ОрелГИЭТ, (2). — С. 247–253.

Omarkhanova Z.M., Niyazbekova S.U., Varzin V.V., Kerimkhulle S.Y., Nurekenova E.S. (2022) Financial provision of the agro-industrial complex of Kazakhstan: problems and solutions. In Sustainable Agriculture: Circular to Reconstructive, Volume 2. — P. 27-32). Singapore: Springer Nature Singapore.

Мадиярова Д.М., Тюленева Т.А., Молдажанов М.Б., Аманбаева А.А. (2020) Механизм интеграции сельскохозяйственного сектора экономики Казахстана в рамках ЕАЭС. Естественно-гуманитарные исследования, (4 (30). — С. 103–109.

Сыздыков Б.Ш., Шарипов А.К., Абилкасым А.Б. (2020) Конкурентоспособность агропромышленного комплекса Туркестанской области Казахстана. Проблемы агрорынка, (2). — С. 89–99.

Блинова Ю.Ю. (2015) Применение экономика-математических методов в маркетинговых исследованиях. In Аналитические и финансово-экономические аспекты развития региональной экономики. — С. 216–219.

Першуков Н.В., Карнадуд О. С. (2019) Математические модели в маркетинге. In Россия молодая. — С. 90731–90731).

Пономаренко К. В., Затонская И. В. (2016). Методы оптимизации маркетинговой стратегии предприятия. In В сборнике: Современная наука: теоретический и практический взгляд сборник статей Международной научно-практической конференции: — С. 260–262).

Garrido A. (2000) A mathematical programming model applied to the study of water markets within the Spanish agricultural sector. Annals of Operations Research, 94. — P. 105–123.

Kozlovskyi S., Khadzhynov I., Lavrov R., Skydan O., Ivanyuta N., Varshavska N. (2019). Economic-mathematical modeling and forecasting of competitiveness level of agricultural sector of Ukraine by means of theory of fuzzy sets under conditions of integration into European market.

Laburtseva O., Larina, Y., Nahorna O., Vinichenko S., Hordiienko V., Al-Shaban A.T.N. (2021) Development of a marketing strategy for enterprise financial growth. Studies of Applied Economics, 39(5).

Kovalchuk S.V., Kobets D.L., Zaburmekha Y.M. (2019) Modeling the choice of strategies of marketing management of enterprise personnel. Natsional'nyi Hirnychyi Universytet. Naukovyi Visnyk, (2). — P. 163–173.

Hung-Yi T. (2024, June). Bp Neural Network Research on Enterprise Marketing Strategy Combination. In 2024 IEEE International Conference on Information Technology, Electronics and Intelligent Communication Systems (ICITEICS). — C. 1–5. IEEE.

Qiu J. (2021, October). An Enterprise Marketing Channel Optimization Strategy in the Context. In The 2021 International Conference on Machine Learning and Big Data Analytics for IoT Security and Privacy: SPIoT-2021 Volume 1. — Vol. 97. — P. 314. Springer Nature.

Qu L., Zhou Z., Zhang T., Zhang H., Shi L. (2019). An improved procedure to implement NSGA-III in coordinate waste management for urban agglomeration. Waste Management and Research, 37(11). — P. 1161–1169.

Yang M., Zhang B., Shi Z., & Li J. (2025). A hybrid task allocation approach for multi-UAV systems with complex constraints: a market-based bidding strategy and improved NSGA-III optimization. The Journal of Supercomputing, 81(4). — P. 1–33.

References

Shulenbaeva F.A., Kamenova M.Zh., Daribaeva A.K., Shulenbaeva K.S. (2011) Voprosy organizacii i razvitija marketinga v agrarnom sektore Kazahstana [Issues of organization and development of marketing in the agricultural sector of Kazakhstan]. Nauchnye Zapiski OrelGIJeT, (2). — P. 247–253 (in Russ).

Omarkhanova Z.M., Niyazbekova S.U., Varzin V.V., Kerimkhulle S.Y., Nurekenova E.S. (2022) Financial provision of the agro-industrial complex of Kazakhstan: problems and solutions. Sustainable Agriculture: Circular to Reconstructive, Vol. 2. — P. 27–32. (in Eng).

Madiyarova D.M., Tyuleneva T.A., Moldazhanov M.B., Amanbaeva A.A. (2020) Mekhanizm integracii sel'skohozjajstvennogo sektora jekonomiki Kazahstana v ramkah EAJeS [Mechanism of integration of the agricultural sector of Kazakhstan within the EAEU]. Estestvenno-gumanitarnye issledovanija, 4(30). — P. 103–109. (in Russ).

Syzdykov B.Sh., Sharipov A.K., Abilkassym A.B. (2020) Konkurentosposobnost' agropromyshlennogo kompleksa Turkestanskoj oblasti Kazahstana [Competitiveness of the agro-industrial complex of Turkestan region of Kazakhstan]. Problemy agrorynka (2). — P. 89–99. (in Russ).

Blinova Ju.Ju. (2015) Primenenie jekonomiko-matematicheskih metodov v marketingovyh issledovanijah [Application of economic–mathematical methods in marketing research]. Analiticheskie i finansovo-jekonomicheskie aspekty razvitija regional'noj jekonomiki. — P. 216–219. (in Russ).

Pershukov N.V., Karnadud O.S. (2019) Matematicheskie modeli v marketinge [Mathematical models in marketing]. In Rossija molodaja. — P. 90731–90731. (in Russ).

Ponomarenko K.V., Zatonskaja I.V. (2016) Metody optimizacii marketingovoj strategii predprijatija [Methods of optimization of enterprise marketing strategy]. In Sovremennaja nauka: teoreticheskij i prakticheskij vzgljad. — P. 260–262. (in Russ).

Garrido A. (2000) A mathematical programming model applied to the study of water markets within the Spanish agricultural sector. Annals of Operations Research. — P. 105–123p. (in Eng).

Kozlovskyi S., Khadzhynov I., Lavrov R., Skydan O., Ivanyuta N., Varshavska N. (2019) Economic-mathematical modeling and forecasting of competitiveness level of agricultural sector of Ukraine by means of theory of fuzzy sets under conditions of integration into European market. (in Eng).

Laburtseva O., Larina Y., Nahorna O., Vinichenko S., Hordiienko V., Al-Shaban A.T.N. (2021) Development of a marketing strategy for enterprise financial growth. Studies of Applied Economics, 39(5). (in Eng).

Kovalchuk S.V., Kobets D.L., Zaburmekha Y.M. (2019) Modeling the choice of strategies of marketing management of enterprise personnel. Natsional'nyi Hirnychyi Universytet. Naukovyi Visnyk, (2). — P. 163–173. (in Eng).

Hung-Yi T. (2024, June) BP Neural Network Research on Enterprise Marketing Strategy Combination. In 2024 IEEE International Conference on Information Technology, Electronics and Intelligent Communication Systems (ICITEICS). — P. 1–5. IEEE. (in Eng).

Qiu J. (2021) An Enterprise Marketing Channel Optimization Strategy in the Context. In The 2021 International Conference on Machine Learning and Big Data Analytics for IoT Security and Privacy: SPIoT-2021. — Vol. 97. — P. 314. Springer Nature. (in Eng).

Qu L., Zhou Z., Zhang T., Zhang H., Shi L. (2019) An improved procedure to implement NSGA-III in coordinate waste management for urban agglomeration. Waste Management & Research, 37(11). — P. 1161–1169. (in Eng).

Yang M., Zhang B., Shi Z., Li J. (2025) A hybrid task allocation approach for multi-UAV systems with complex constraints: a market-based bidding strategy and improved NSGA-III optimization. The Journal of Supercomputing, 81(4). — P. 1–33. (in Eng).

ACADEMIC SCIENTIFIC JOURNAL OF COMPUTER SCIENCE
ISSN 1991-346X
Volume 4. Namber 356 (2025). 40–54

https://doi.org/10.32014/2025.2518-1726.382

IRSTI 27.25.19
UDC 004.89; 004.912

© A.O. Aliyeva[1], B.S. Omarov[2], R.B. Abdrakhmanov[3], D.R. Sultan[4], A.B. Toktarova[5*], 2025.

[1]Khoja Akhmet Yassawi International Kazakh – Turkish University, Turkistan, Kazakhstan;
[2]International information technology university, Almaty, Kazakhstan;
[3]International University of Tourism and Hospitality, Turkestan, Kazakhstan;
[4]Narxoz University, Almaty, Kazakhstan;
[5]Mukhtar Auezov South Kazakhstan University, Shymkent, Kazakhstan.
E-mail: toktar.aigerim@list.ru

NEURAL NETWORK MODEL FOR AUTOMATIC DETECTION OF KAZAKH-LANGUAGE HATE SPEECH

Aliyeva Akbayan — Candidate of Pedagogical Sciences, department of Computer engineering, Khoja Akhmet Yassawi International Kazakh – Turkish University, Turkistan, Kazakhstan,
E-mail: aliyevaakbayan@mail.ru, https://orcid.org/0009-0006-6932-7302;
Omarov Batyrkhan — PhD, associate professor of the Department of Information System, International information technology university, Almaty, Kazakhstan,
E-mail: aliyevaakbayan@mail.ru, https://orcid.org/0009-0006-6932-7302;
Abdrakhmanov Rustam — Cand.of Techn.Sc., Ass.Prof., International University Of Tourism And Hospitality Turkestan, Kazakhstan,
E-mail: abdrakhmanov.rustam@iuth.edu.kz, https://orcid.org/0000-0003-0385-0684;
Sultan Danyar — PhD, Department of Information System, Narxoz university, Almaty, Kazakhstan,
E-mail: sultan.daniyar96@gmail.com, https://orcid.org/0000-0002-1611-1923;
Toktarova Aigerim — PhD, department of Information communication technologies, M. Auezov South Kazakhstan university, Shymkent, Kazakhstan,
E-mail: toktar.aigerim@list.ru, https://orcid.org/0000-0002-6265-9236.

Abstract. This article describes how to use modern machine and deep learning algorithms to automatically detect offensive language in Kazakh-language web content. The relevance of the research work is aimed at protecting the digital environment and reducing discriminatory language in social media and user platforms. To achieve this goal, methods for data collection, pre-processing, classification model creation and optimization were developed and tested. The article describes the collection of data from social networks (Instagram, TikTok, YouTube, Facebook), news sites and forums with Kazakh-language user-generated content as the object of research. As a practical significance of the research work, data pre-processing includes noise removal, tokenization, lemmatization and stop

word removal. As a result of the presented work, the annotated corpus classified communications as “offensive language words” or “neutral opinions”. The authors use a hybrid deep learning architecture to create an efficient and flexible system for automatically detecting discrimination on the Kazakh-language Internet. CNN, LSTM, Bi-LSTM and Transformer models and their integration with the attention process are studied to consider local and contextual text features. The result of the article is aimed at training and validating a model that can recognize offensive and negative words in comments in online content in real time. This approach to social media moderation platforms is aimed at improving user safety, removing dangerous comments, and laying the foundation for research into natural language processing in the Kazakh language. Multilingual platforms can use the proposed method to create intelligent systems for automatic moderation and long-term protection of the digital environment.

Key words: collecting data, online content, hate speech, neural network, deep learning, detection

Acknowledgments

This study was funded by the Science Committee of the Ministry of Science and Higher Education of the Republic of Kazakhstan (grant No. AP23488900-Automatic detection of cyberbullying among young people in social networks using artificial intelligence)

© А.О. Әлиева[1], Б.С. Омаров[2], Р.Б. Абдрахманов[3], Д.Р. Султан[4], А.Б. Тоқтарова[5*], 2025.

[1]Қожа Ахмет Яссауи атындағы Халықаралық Қазақ – Түрік университеті, Түркістан, Қазақстан;

[2]Халықаралық ақпараттық технологиялар университеті, Алматы, Қазақстан;

[3]Халықаралық туризм және меймандостық университеті, Түркістан, Қазақстан;

[4]Narxoz университеті, Алматы, Қазақстан;

[5]М. Әуезов атындағы Оңтүстік Қазақстан университеті, Шымкент, Қазақстан.

E-mail: toktar.aigerim@list.ru

ҚАЗАҚ ТІЛІНДЕГІ ДИСКРИМИНАЦИЯНЫ АВТОМАТТЫ АНЫҚТАУҒА АРНАЛҒАН НЕЙРОНДЫҚ ЖЕЛІЛІК МОДЕЛІ

Әлиева Ақбаян — педагогика ғылымдарының кандидаты, Компьютерлік инженерия кафедрасының доценті, Қожа Ахмет Яссауи атындағы Халықаралық Қазақ – Түрік университеті, Түркістан, Қазақстан,
E-mail: aliyevaakbayan@mail.ru, https://orcid.org/0009-0006-6932-7302;
Омаров Батырхан — PhD, қауымдастырылған профессор, Ақпараттық жүйелер кафедрасының доценті, Халықаралық Ақпараттық технологиялар университеті, Алматы, Қазақстан,
E-mail: batyahan@gmail.com, https://orcid.org/0000-0002-8341-7113;

Абдрахманов Рустам — т.ғ.к., доцент, Халықаралық туризм және меймандостық университеті Түркістан, Қазақстан,
E-mail: abdrakhmanov.rustam@iuth.edu.kz , https://orcid.org/0000-0003-0385-0684;
Султан Даняр — PhD, Ақпараттық жүйелер кафедрасының доценті, Narxoz университетінің Алматы, Қазақстан,
E-mail: sultan.daniyar96@gmail.com, https://orcid.org/0000-0002-1611-1923;
Тоқтарова Айгерім — PhD, Ақпараттық коммуникациялық технологиялар кафедрасының аға оқытушысы, М.Әуезов атындағы Оңтүстік Қазақстан университеті, Шымкент, Қазақстан,
E-mail: toktar.aigerim@list.ru, https://orcid.org/0000-0002-6265-9236.

Аннотация. Бұл мақала қазақ тіліндегі веб-контенттегі бейәдеп сөздерді автоматты түрде анықтау үшін заманауи машина және терең оқыту алгоритмдерін қолдану тәсілдерін сипаттайды. Ғылыми зерттеудің өзектілігі цифрлық ортаны қорғау және әлеуметтік медиа мен пайдаланушылық платформаларда дискриминациялық тілдің таралуын азайту қажеттілігімен анықталады. Осы мақсатқа жету үшін деректерді жинау, алдын ала өңдеу, классификация моделін құру және оңтайландыру әдістері әзірленіп, сынақтан өткізілді. Мақалада зерттеу нысаны ретінде әлеуметтік желілерден (Instagram, TikTok, YouTube, Facebook), жаңалықтар сайттарынан және қазақтілді қолданушылар құрастырған контенті бар форумдардан алынған деректерді жинау үдерісі сипатталады. Зерттеудің практикалық маңыздылығы деректерді алдын ала өңдеу кезеңінің — шуды жою, токенизация, лемматизация және тоқтау сөздерді алып тастау сияқты қадамдарды қамтуымен түсіндіріледі. Зерттеу нәтижесінде аннотацияланған корпус коммуникацияларды «бейәдеп сөз» немесе «бейтарап пікір» деп жіктей алды. Авторлар қазақтілді интернеттегі дискриминацияны автоматты түрде анықтаудың тиімді әрі икемді жүйесін құру үшін гибридті терең оқыту архитектурасын қолданады. Жергілікті және мәтінмәндік ерекшеліктерді ескеруге мүмкіндік беретін CNN, LSTM, Bi-LSTM және Transformer үлгілері, сондай-ақ назар аудару механизмін интеграциялау тәсілдері талданады. Мақаланың нәтижесі онлайн-контенттегі пікірлерден бейәдеп және жағымсыз сөздерді нақты уақытта тани алатын модельді оқыту және валидациялау болды. Әлеуметтік желілерді модерациялау платформалары үшін ұсынылған әдіс қолданушылардың қауіпсіздігін арттыруға, қауіпті пікірлерді жоюға және қазақ тіліндегі табиғи тілді өңдеу бойынша болашақ зерттеулерге негіз қалайды. Көптілді платформалар ұсынылған тәсілді автоматты модерацияға арналған интеллектуалды жүйелерді дамыту және цифрлық ортаны ұзақ мерзімді қорғау үшін пайдалана алады.

Түйін сөздер: деректерді жинау, желілік контент, дискриминациялық сөздер, нейрондық желі, терең оқыту, анықтау

© А.О. Алиева[1], Б.С. Омаров[1], Р.Б. Абдрахманов[3], Д.Р. Султан[4], А.Б. Токтарова[5*], 2025.

[1]Международный казахско-турецкий университет имени Ходжи Ахмета Яссави, Туркестан, Казахстан;

[2]Международный университет информационных технологий, Алматы, Казахстан;

[3]Международный университет туризма и гостеприимства, Туркестан, Казахстан;

[4]Университет Нархоз, Алматы, Казахстан;

[5]Южно-Казахстанский университет имени М. Ауэзова, Шымкент, Казахстан.

E-mail: toktar.aigerim@list.ru

НЕЙРОСЕТЕВАЯ МОДЕЛЬ ДЛЯ АВТОМАТИЧЕСКОГО ОБНАРУЖЕНИЯ ДИСКРИМИНАЦИИ В КАЗАХСКОМ ЯЗЫКЕ

Алиева Акбаян — кандидат педагогических наук, доцент кафедры Компьютерной инженерии, Международный казахско-турецкий университет имени Ходжи Ахмета Яссави, Туркестан, Казахстан,
E-mail: aliyevaakbayan@mail.ru, https://orcid.org/0009-0006-6932-7302;
Омаров Батырхан — PhD, доцент кафедры Информационные системы, Международный университет информационных технологий, Алматы, Казахстан,
E-mail: batyahan@gmail.com, https://orcid.org/0000-0002-8341-7113;
Абдрахманов Рустам — кандидат технических наук, доцент, Международный университет туризма и гостеприимства, Туркестан, Казахстан,
E – mail: abdrakhmanov.rustam@iuth.edu.kz, https://orcid.org/0000-0003-0385-0684;
Султан Данияр — PhD, доцент кафедры Информационные системы, Университет Нархоз, Алматы, Казахстан,
E-mail: sultan.daniyar96@gmail.com , https://orcid.org/0000-0002-1611-1923;
Токтарова Айгерим — PhD, кафедра Информационно-коммуникационных технологий, Южно-Казахстанский университет им. М. Ауэзова, Шымкент, Казахстан,
E-mail: toktar.aigerim@list.ru, https://orcid.org/0000-0002-6265-9236.

Аннотация. В данной статье рассматривается использование современных алгоритмов машинного и глубокого обучения для автоматического выявления ненормативной лексики в казахскоязычном веб-контенте. Актуальность исследования заключается в необходимости защиты цифровой среды и снижении уровня дискриминационной лексики в социальных сетях и на пользовательских платформах. Для достижения этой цели были разработаны и апробированы методы сбора, предварительной обработки данных, а также построения и оптимизации моделей классификации. В статье приводится описание сбор данных из социальных сетей (Instagram, TikTok, YouTube, Facebook), новостных сайтов и форумов, объектом исследования которых является казахскоязычный пользовательский контент. Практическая значимость исследования заключается в том, что предварительная обработка

данных включает в себя удаление шума, токенизацию, лемматизацию и удаление стоп-слов. В результате представленной работы аннотированный корпус позволил классифицировать сообщения как «ненормативную лексику» или «нейтральные мнения». Авторы используют гибридную архитектуру глубокого обучения для создания эффективной и гибкой системы автоматического выявления дискриминации в казахскоязычном интернете. Изучаются модели CNN, LSTM, Bi-LSTM и Transformer, а также их интеграция с процессом внимания для учёта локальных и контекстных особенностей текста. Результатом исследования стало обучение и валидация модели, способной распознавать оскорбительные и негативные выражения в комментариях к онлайн-контенту в режиме реального времени. Предложенный подход ориентирован на повышение безопасности пользователей, удаление опасных комментариев и создание базы для дальнейших исследований в области обработки естественного языка на казахском языке. Предложенная методология может быть адаптирована многоязычными платформами для разработки интеллектуальных систем автоматической модерации и долгосрочной защиты цифровой среды.

Ключевые слова: сбор данных, онлайн-контент, дискриминационные слова, нейронная сеть, глубокое обучение, обнаружение

Introduction. The advancement of digital technology and the pervasive utilization of social networks, instant messaging applications, and open communication platforms have substantially augmented the volume of user-generated information disseminated on the Internet. Simultaneously, there is a rise in the volume of messages featuring toxic language, insulting terminology, and various manifestations of hate speech. Such phenomena provide a significant risk to the psychological well-being of users and to the establishment of a secure and conducive digital environment. The necessity of creating efficient techniques for the automatic identification and filtration of such content arises from the imperative to safeguard vulnerable populations, particularly children and teenagers, while also ensuring a high standard of information security on the Internet. Recent years have witnessed notable advancements in the scientific literature regarding automated text analysis and the identification of undesirable content. Machine learning and deep neural network methodologies have been established and are currently employed, encompassing models based on convolutional and recurrent architectures (CNN, RNN, LSTM), in addition to transformers and their variants.

These methodologies are effectively employed to address the issues of classifying toxic communications, detecting cyberbullying, and forecasting manifestations of violent behavior in both English-language and multilingual online environments (Omarov, 2025). Furthermore, the deployment of such systems on platforms for automated moderation of social networks, online forums, and news portals can markedly diminish the velocity of harmful information propagation. Nonetheless,

despite advancements in this domain, the Kazakh-language section of the Internet remains inadequately researched. The scarcity of comprehensive labeled corpora, inadequate representation of the Kazakh language in global research, and the unique morphological characteristics of the language provide significant challenges in developing effective and precise models (Azhibekova, 2025).

Developing and training algorithms necessitates intricate tasks, including data collection from many sources, preprocessing, annotation, and the creation of a specific corpus. Creating tailored systems that accurately recognize toxic language and symptoms of hate speech in the local digital environment is contingent upon a high-quality and representative corpus. Consequently, undertaking research in this domain is essential for establishing a secure and culturally attuned Internet environment in Kazakhstan.

Related work. Recent years have seen a proliferation of research focused on the automatic identification of toxic language and hate speech, predominantly within the English-speaking domain of the Internet. International research emphasizes the advancement and comparative evaluation of machine learning algorithms, encompassing traditional classification techniques and contemporary deep learning methodologies.

Initially, such issues were addressed by conventional text processing techniques and statistical models. Specifically, Support Vector Machine (SVM), Logistic Regression, Naive Bayes, and Random Forest were employed, yielding appropriate outcomes with meticulously chosen features. Nevertheless, these strategies were inadequate when addressing substantial quantities of unstructured texts and when attempting to discern the concealed semantics of hurtful remarks.

With the advancement of deep learning technologies, recurrent neural networks (RNN), long short-term memory networks (LSTM) and its variants, together with convolutional neural networks (CNN), have been extensively employed to extract contextual and local information from text. Recent research illustrates the superior efficacy of hybrid models that include LSTM and CNN, alongside the implementation of attention mechanisms and transformers (BERT, RoBERTa), which facilitate the consideration of both local and global dependencies within the text.

Significant emphasis in global literature is placed on the development of specialized corpora for hate speech detection tasks. Huang proposed a multilingual Twitter corpus to examine demographic bias in the identification of harmful information (Huang, 2020). Sachdeva utilized Rasch measurement theory to evaluate the reliability of annotations and discern variations in the perception of message toxicity (Sachdeva, 2024). Fanton developed a corpus of counter-narratives to counteract hate speech employing the Human-in-the-Loop methodology (Fanton, 2022).

Notwithstanding advancements in international research, the Kazakh-language segment remains inadequately explored. The scarcity of open corpora and the

limited availability of annotated data hinder the creation of precise models. This underscores the necessity of developing localized corpora and tailoring contemporary deep learning architectures to the morphological characteristics of the Kazakh language (Mussiraliyeva, 2021).

Materials and Methods. This study employs a comprehensive methodology for the automatic identification of hate speech in Kazakh-language online content, encompassing multiple sequential phases: data collection, pre-processing, annotation, corpus formation, and the development and evaluation of deep learning models (Bolatbek, 2024).

The process of constructing a data corpus is crucial for developing a system that automatically detects hate speech in Kazakh-language online content, as the efficacy of subsequent training of machine and deep learning models is directly contingent upon the quality and representativeness of the gathered data.

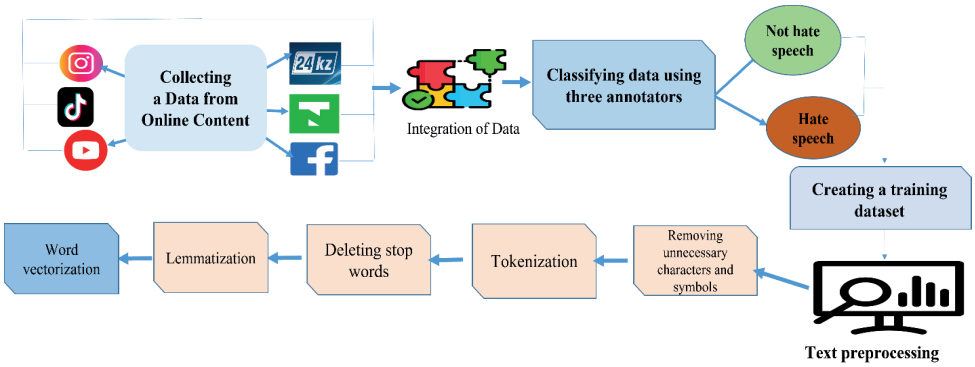


Figure 1 - Data collection and preprocessing framework

A variety of sources were utilized to construct the corpus, ensuring extensive representation of user speech (Figure 1). The primary data sources were:

a) Social networks such as Instagram, TikTok, YouTube, and Facebook offer a plethora of casual content accompanied by emotive and spontaneous commentary, where instances of hate speech are prevalent.

b) News portals and online media such as Tengrinews, Informburo, and Massaget feature comments on news that reflect socially significant issues and include instances of violent language.

c) Internet forums and specialized platforms, where in-depth conversations occur and covert instances of foul language may be present.

Two methodologies were employed to automate the collecting:

a) Utilization of Python scripts employing the BeautifulSoup and Selenium libraries for the direct processing of web pages to extract comments and messages from open sources.

b) Utilizing official social media APIs (YouTube Data API, Facebook Graph API) enabled the collection of substantial data in a structured format (Haleem,2021).

To guarantee the representativeness of the corpus, the sample encompassed texts that illustrated both informal user discourse (colloquial language, emotive expressions) and public discourses (socially and politically pertinent subjects). At this step, initial filtering was conducted, which involved eliminating duplicates, discarding unnecessary messages (such as advertisements), and identifying the language to keep solely Kazakh-language content.

Following the establishment of the primary corpus, manual annotation was conducted with the involvement of three voluntary annotators: a researcher, a psychologist, and a Kazakh language educator. Each text was categorized as either "hate speech" or "non-hate speech". This method guaranteed a significant degree of dependability and consistency in the annotation.

Consequently, a systematic and representative corpus of Kazakh-language internet content was established, serving as the foundation for later preprocessing, model creation, and training of classification algorithms (Ojo, 2022).

The concluding phase of the study entailed the use and evaluation of deep learning techniques for the automatic classification of Kazakh-language internet content as either including or not comprising hate speech. The study initially encompassed traditional deep learning architectures, including LSTM, Bi-LSTM, CNN, and MLP, facilitating an evaluation of their efficacy in addressing the agglutinative structure of the Kazakh language and revealing the limitations of each model (Zhu, 2019).

The preliminary stage results indicated that employing each architecture alone yields adequate accuracy; nonetheless, it is insufficient for detecting contextually obscured insults and latent manifestations of hate speech (Guo, 2025). A hybrid model was devised and executed, integrating the advantages of multiple deep learning techniques.

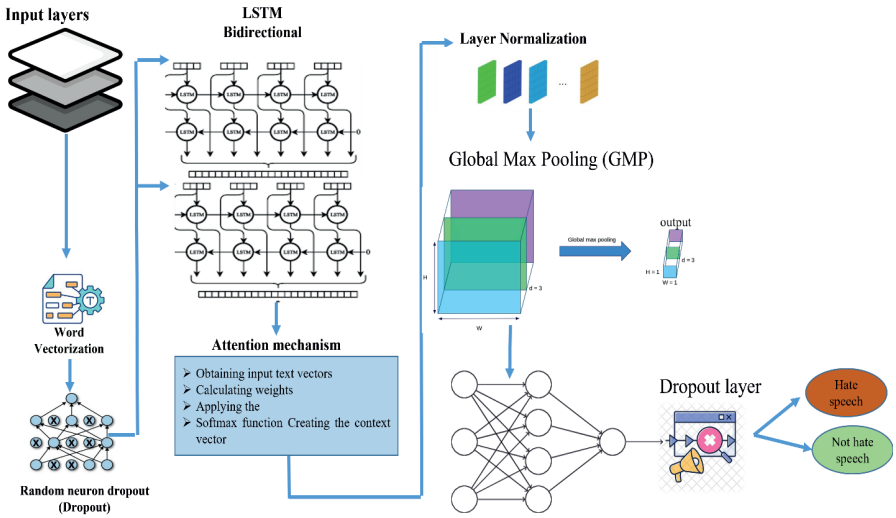


Figure 2 - Framework for the development and training of models

The hybrid architecture's structure, seen in the diagram, comprises the subsequent stages (Figure 2):

Word vectorization is an important step in preparing text data for processing by deep learning algorithms. After text preprocessing, including tokenization, lemmatization, and stop word removal, each word is converted into a fixed-dimensional numerical vector representation that reflects its meaning and contextual relationships.

$$W_i \rightarrow \overline{v_i} = (v_{i1}, v_{i2}, \ldots, v_{id}) \quad (1)$$

Where W_i – input words, $\overline{v_i}$ – It is a vector representation of dimension d, comprising numerical values that encapsulate the semantic attributes of the word.

The Word2Vec and FastText models are employed to generate such vectors. The Word2Vec model generates vector representations derived from word contexts, assigning like vectors to words that appear in comparable situations. FastText enhances this methodology by incorporating subwords and morphological components, which is particularly significant for the Kazakh language due to its agglutinative nature (Chen, 2024).

Vectorization assigns a numerical representation to each lexeme, enabling input into neural networks like Bi-LSTM or CNN for enhanced feature extraction and categorization of messages into the categories of "hate speech" and "non-hate speech".

During the extraction of contextual data, bidirectional LSTM layers are employed, enabling the model to concurrently assess the text sequence from both left to right and right to left (Zhang, 2018). This method guarantees that both preceding and subsequent words are considered for each text element, which is particularly crucial for the Kazakh language due to its agglutinative morphology and flexible word order. Consequently, deep context vectors are generated, representing semantic dependencies inside a phrase and latent semantic relationships among words.

$$h_t = [h_t^{fwd}; h_t^{bwd}] \quad (2)$$

Where, h_t^{fwd} –concealed state of advancement,
h_t^{bwd} –hidden state of the opposite direction,
h_t – shared context vector for position t

The utilization of Bi-LSTM facilitates the successful extraction of long-term dependencies and consideration of sentence context when categorizing text as "hate speech" or "non-hate speech." The integration of these features with subsequent convolutional or fully connected layers enhances the model's accuracy while processing real data from social networks and news sources.

The subsequent phase in text processing is employing a Convolutional Neural Network (CNN) to extract localized features and distinctive patterns

within text input. In the realm of hate speech analysis, CNN excels in detecting n-gram dependencies that may signify the existence of hostile or toxic words. A convolutional layer systematically "slides" over a matrix of word embeddings, calculating a filter response for each segment of text. Each filter of fixed width k encapsulates the local context of k successive words (Naser,2025).

Consequently, a feature map is generated for each filter, indicating the frequency of a specific pattern's occurrence in the text.

$$c_i = f(W * x_{i*i+k+1} + b) \quad (3)$$

$x_{i*i+k+1}$ —vector representation of a word sequence of length k
W- filter weight matrix
b —bias
$f(*)$ — activation function
c_i — the value of the attribute at position i

Subsequent to the use of convolutional filters, Global Max Pooling is employed to identify the most salient features over the full sentence length. This enables us to diminish the dimensionality of the feature space and concentrate on critical signs of hate speech, like combinations of harsh language, aggressive expressions, or distinct markers of toxicity.

Employing a CNN at this juncture equips the model to identify localized patterns of harmful discourse that may remain obscured when solely examining the overarching context. In the hybrid architecture, these features are integrated with context vectors derived from Bi-LSTM, enhancing the precision of message categorization (Sharifi, 2025).

The concluding phase of the model's functionality entails the amalgamation of two categories of features:

a) Contextual characteristics derived from Bi-LSTM, which consider the semantic relationships of words inside a phrase;

b) Local characteristics derived from CNN, which represent distinctive n-gram patterns and localized indicators of abusive language.

The features are amalgamated into a singular vector, which is subsequently input into a fully linked layer. This layer categorizes messages utilizing the Softmax function, which computes the likelihood of the text being classified as either "Hate speech" (contains hate speech) or "Not hate speech" (does not contain it).

$$P(y_i) = \frac{e^{zi}}{\{_{j=1}^{K} e^{zj}} \quad (4)$$

Where, zi — output of the model for class i,
K — number of classes
$P(y_i)$ - likelihood of a message being classified as belonging to class i

The research covered all important processes for automatically detecting hate

speech in Kazakh-language internet content, from data collection and annotation to hybrid deep learning model creation and evaluation. A sample corpus from social networks, news portals, and forums was preprocessed to remove stop words, tokenize, lemmatize, and filter unnecessary content. Three volunteer professionals annotated corpus. After converting the data into vector representations using Word2Vec and FastText models, Kazakh semantic and morphological properties were numerically represented. A hybrid architecture using Bi-LSTM for contextual dependency capture and CNN for local pattern recognition extracted features and built a classifier. The incorporated features were fed into a fully connected layer using a Softmax function to determine whether a message was "Hate speech" or "Not hate speech".

Results. Experiments were conducted on three data corpora—Kazakh, Russian, and English (Hate Speech and Offensive Language Dataset)—to assess the effectiveness of the proposed model. The model was evaluated against traditional machine learning algorithms (Logistic Regression, Random Forest, Decision Tree, Naïve Bayes, K-Nearest Neighbors, SVM) as well as contemporary deep learning architectures (LSTM, BiLSTM, CNN, RNN, MLP, Transformer).

Table 1. Evaluation of the efficacy in Kazakh, Russian, and English

Type of database	**Method**	**Algorithm**	**Accuracy**	**Precision**	**Recall**	**F1-score**	**AUC-ROC**
Database in Kazakh	Deep learning	Proposed model	92.1%	93%	92.5%	92.9%	95%
		LSTM	91.0%	92.5%	92.8%	92.5%	92%
		BiLSTM	91.3%	92.6%	93%	92.7%	92%
		CNN	89.2%	89%	89%	88.7%	90%
		RNN	89.7%	89.5%	90%	89.9%	94%
	Machine learning	LR	87.3%	85.2%	86.2%	85.1%	78%
		RF	85.6%	83.9%	83.1%	83.7%	92%
		DT	87.4%	83.2%	86.3%	85.1%	80%
		NB	60.2%	52.4%	58.5%	64.2%	65%
		KNN	85.1%	85.4%	82.2%	85.6%	77%
		SVM	86.2%	85.3%	83.7%	85.8%	78%
Database in Russian (*Russian Language Toxic Comments*)	Deep learning	Ұсынылған модель	91.5%	93%	92%	92.4%	94%
		BiLSTM	91%	92%	92%	92%	93%
		CNN	88.7%	89%	89%	88.9%	90%
		LSTM	90%	91%	91.5%	91.2%	93%
		RNN	89%	89%	91%	90%	92%
	Machine learning	LR	87.3%	85.6%	87.4%	60.2%	85.1%
		RF	85.2%	83.9%	83.2%	52.4%	85.4%
		DT	86.2%	83.1%	86.3%	58.5%	82.2%
		NB	85.1%	83.7%	85.1%	64.2%	85.6%
		KNN	75%	90%	76%	68%	77%
		SVM	87.3%	85.6%	87.4%	60.2%	85.1%

Type of database	Method	Algorithm	Accuracy	Precision	Recall	F1-score	AUC-ROC
Database in English (*hate-speech-and-offensive-language*)	Deep learning	Recommended model	95%	96%	96.5%	96%	97%
		LSTM	93%	92.5%	94%	95%	95%
		MLP	91.5%	92%	93%	92.5%	94%
		CNN	90%	92%	91%	91.5%	93%
		RNN	91%	93%	92%	92%	92%
	Machine learning	LR	89.3%	90%	89.5%	89.3%	89%
		RF	88.6%	89%	88%	88.5%	89%
		DT	87.4%	83.2%	86.3%	85.1%	80%
		NB	86.2%	82.4%	85.5%	84.5%	85%
		KNN	87.1%	85.4%	87.2%	86.4%	87%
		SVM	87.2%	87.3%	86.7%	85.8%	88%

The hybrid model applied to the Kazakh corpus achieved an Accuracy of 92.1%, Precision of 93%, Recall of 92.5%, F1-score of 92.9%, and AUC-ROC of 95%, surpassing all traditional and individual neural architectures. Among the base models, BiLSTM demonstrated the highest performance; however, the incorporation of CNN and attention mechanisms into the hybrid model resulted in a substantial enhancement in performance, particularly in the F1-score metric, which indicates the equilibrium between recall and accuracy.

The Russian-language corpus yielded the following indicators: Accuracy 91.5%, Precision 93%, Recall 92%, F1-score 92.4%, and AUC-ROC 94%. These results further validate the stability and versatility of the proposed architecture within a morphologically rich language context.

The English-language corpus (Hate Speech and Offensive Language Dataset) yielded notable results: Accuracy 95%, Precision 96%, Recall 96.5%, F1-score 96%, and AUC-ROC of 97%. This results from the greater volume and improved balance of English-language data, facilitating more effective model training.

The comparative results indicate that the proposed model offers:

a) Notable enhancement in all primary metrics relative to baseline models;

b) Significant robustness to class imbalance, particularly relevant for Kazakh-language and Russian-language datasets.

The potential for generalization across various languages and domains, encompassing social networks and open platforms, is noteworthy.

The proposed architecture serves as an optimal solution for the automated detection of hate speech in multilingual online environments, achieving a balance among accuracy, recall, and overall performance.

The developed model's practical applicability is demonstrated through the implementation of the web interface "Kazakh Hate Speech Detection," which analyzes text messages to classify them as either "Hate speech" or "Not hate speech" in real time.

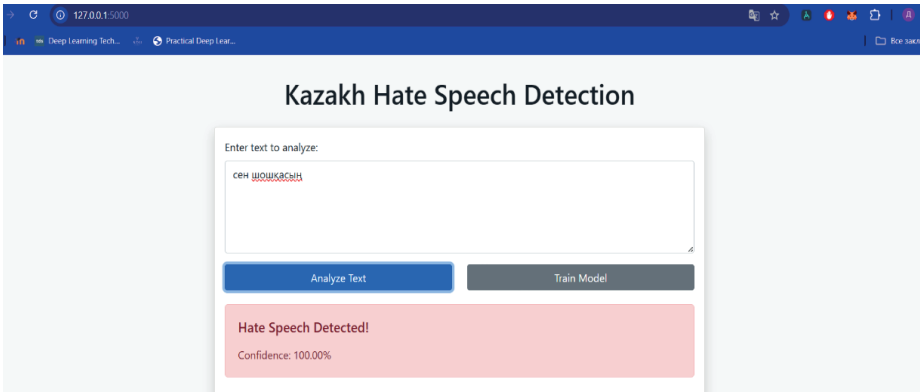


Figure 3 - Web interface of hate speech detection (hate speech)

Figure 3 illustrates an instance of inputting a message with offensive language ("sen shoshqasyn"), which the system accurately categorizes as "Hate speech" with a certainty of 100%.

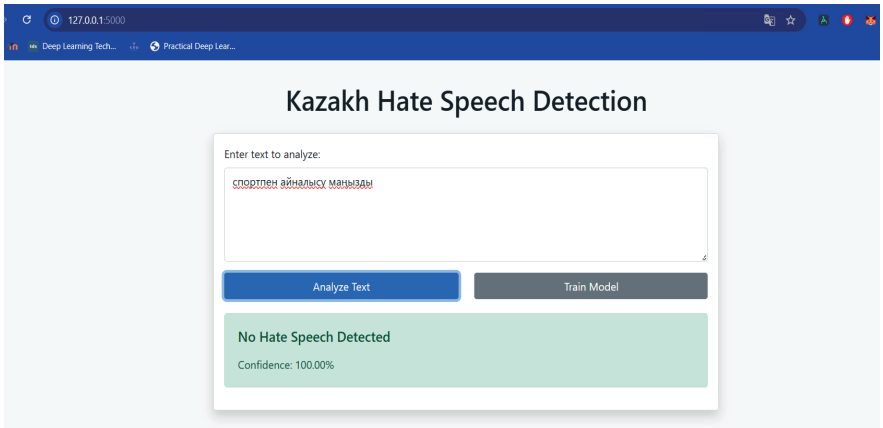


Figure 4 - Web interface of hate speech detection (not hate speech)

Figure 4 illustrates the processing of a neutral statement ("sportpen ainalysu mangyzdy"), which the model classifies as "Not hate speech" with a correspondingly high probability. The results indicate that the proposed model exhibits high accuracy rates on test samples and possesses practical applicability for integration into automatic online content moderation systems.

The authors developed and tested a hybrid deep learning system to automatically detect hate speech in Kazakh-language web content. The suggested model has good accuracy, F1-score, and AUC-ROC values of 92%, 92%, and 95%, respectively, suggesting its stability and capacity to identify offensive from neutral phrases.

Comparing classical machine learning algorithms (LR, RF, NB, KNN, SVM) and deep learning models (CNN, LSTM, Bi-LSTM, RNN), the hybrid architecture

achieves the best accuracy, completeness, and classification balance.

Real-time text message analysis was demonstrated by "Kazakh Hate Speech Detection," an online application. Empirical testing showed that the system can detect offensive and neutral messages, proving its applicability for online moderation.

The results show that the proposed approach can create a safe digital environment in Kazakh-language Internet.

Discussion. The study found hybrid deep learning architectures effective for automatically detecting hate speech in Kazakh-language web content. Comparative investigation indicated that combined Bi-LSTM and convolutional CNN layers enhance contextual data retrieved by the recurrent network and local patterns recognized by convolutional filters. This method detects overt and covert verbal aggressiveness more accurately than typical machine learning models. High Accuracy, F1-score, and AUC-ROC values on three language corpora (Kazakh, Russian, English) demonstrate the architecture's adaptability to multilingual data and class imbalance resistance. The study's "Kazakh Hate Speech Detection" web interface showed the model's real-time message analysis potential. This tool can be used in social network and forum moderating systems and digital security projects.

Despite the outcomes, certain issues must be addressed. First, the training data corpus does not contain all Kazakh-language Internet sources and stylistic traits. Second, the current architecture ignores complicated interphrase connections and discourse history, which could improve hidden aggressiveness detection. Expanding the corpus to include emoji and slang language and combining attention processes and transformer architectures to increase model interpretability are promising research areas.

Conclusion. The research proposes using new deep learning approaches to automatically detect hate speech in Kazakh-language web content. After tokenization, lemmatization, and stop word removal, a representative corpus of texts was constructed and annotated. A hybrid model incorporating Bi-LSTM and CNN was implemented on this corpus to achieve good classification accuracy and data imbalance resistance.

The system was tested in real time and shown to be effective for automatic content filtering, information security, and toxicity reduction on Kazakh-language Internet. Development prospects include increasing the corpus, incorporating multimodal data, and adding adaptive training methods to develop a universal tool for online communication monitoring and analysis.

References

Huang H.Y. et al. (2020) miRTarBase 2020: updates to the experimentally validated microRNA–target interaction database. Nucleic acids research. — T. 48. — №. D1. — P. D148-D154 (in English).

Omarov Batyrkhan, Rustam Abdrakhmanov, and Aigerim Toktarova (2025) "Glove-embedded attention bilstm networks for enhanced multiclassification of tweets in cyberbullying detection on online content." Scientific Journal of Astana IT University: P.55-70 (in English).

Azhibekova Zh.Zh, et al. (2024) "Bidirectional long short-term memory in hate speech detection

problem on networks." Bulletin of Abai KazNPU. Series of Physical and Mathematical sciences 87.3: 123-131(in English).

Sachdeva R. et al. (2024) Localizing and mitigating errors in long-form question answering. arXiv preprint arXiv:2407.11930 (in English).

Tekiroglu S.S. et al. (2022) Using pre-trained language models for producing counter narratives against hate speech: a comparative study. arXiv preprint arXiv:2204.01440 (in English).

Haleem M.S., Castaldo R., Pagliara S.M., Petretta M., Salvatore M., Franzese M., & Pecchia L. (2021) Time adaptive ECG driven cardiovascular disease detector. Biomedical Signal Processing and Control. — Vol.70. — P.19-26 (in English).

Zhu F., Ye F., Fu Y., Liu Q., & Shen B. (2019) Electrocardiogram generation with a bidirectional LSTM-CNN generative adversarial network. Scientific reports. — Vol. 9(1). — P. 6734 -6748 (in English).

Chen J., Huang S., Zhang Y., Chang Q., ZhangY., Li D., ... & Liang H. (2024) Congenital heart disease detection by pediatric electrocardiogram based deep learning integrated with human concepts. Nature Communications. — Vol.15(1). — P. 976-985 (in English).

Naseer A., Khan M.M., Arif F., Iqbal W., Ahmad A., & Ahmad I. (2025) An improved hybrid model for cardiovascular disease detection using machine learning in IoT. Expert Systems. —Vol. 42(1). — e13520 (in English).

Sharifi S., & Donyadadi A. (2025) Detection and Diagnosis of Congenital Heart Disease from Chest X-Rays with Deep Learning Models. International Journal of Applied Data Science in Engineering and Health. — Vol. 1(1). — P. 1-9 (in English).

Bolatbek M. et al. Kazakh Language Dataset for Hate Speech Detection on Social Media Text. 2024 IEEE 9th International Conference on Computational Intelligence and Applications (ICCIA). – IEEE, 2024. — P. 94-98 (in English).

Mussiraliyeva S. et al. Bigram based deep neural network for extremism detection in online user generated contents in the kazakh language. International Conference on Computational Collective Intelligence. – Cham. Springer International Publishing, 2021. — P. 559-570 (in English).

Ojo O.E. et al. Automatic hate speech detection using deep neural networks and word embedding. Computación y Sistemas, 2022. — Vol. 26. — No. 2. — P. 1007-1013 (in English).

Guo X., Adnan H. M., Abidin M. Z. Z. A Scientometric Analysis of Potentially Harmful Speech Detection in Muslim Communities: Trends, Challenges, and Future Directions. Society, 2025. — P. 1-21(in English).

Zhang Z., Robinson D., Tepper J. Detecting hate speech on twitter using a convolution-gru based deep neural network. European semantic web conference. — Cham: Springer International Publishing, 2018. — P. 745-760 (in English).

ACADEMIC SCIENTIFIC JOURNAL OF COMPUTER SCIENCE
ISSN 1991-346X
Volume 4. Namber 356 (2025). 55–72

https://doi.org/10.32014/2025.2518-1726.383

IRSTI: 27.17.19; 06.52.13; 28.23.15
UDC 004.942

© O. Auyelbekov[1,2], E. Bostanov[3], S. Sapakova[4*], L. Tukenova[5], A. Kozhagul[2], 2025.

[1]Institute of Information and Computational Technologies, Almaty, Kazakhstan;
[2]Kazakh National Women's Pedagogical University, Almaty, Kazakhstan;
[3]Civil Aviation Academy, Almaty, Kazakhstan;
[4]International Information Technology University, Almaty, Kazakhstan;
[5]Al-Farabi Kazakh National University, Almaty, Kazakhstan.
E-mail: s.sapakova@iitu.edu.kz

MODELING AND ANALYSIS OF A GENERATOR WITH PERMANENT AND VARIABLE MAGNETS

Auyelbekov Omirlan — cand. of ph. and math. sc., Associate Professor, Institute of Information and Computational Technologies; Kazakh National Women's Pedagogical University, Almaty, Kazakhstan,
E-mail: omirlan.auelbek@gmail.com, https://orcid.org/0000-0002-2903-9086;
Bostanov Ernar — senior lecturer, Civil Aviation Academy, Almaty, Kazakhstan,
E-mail: bostanovernar0929@gmail.com, https://orcid.org/0000-0002-8753-0244;
Saya Sapakova — cand. of ph. and math. sc., Associate Professor, International Information Technology University, Almaty, Kazakhstan,
E-mail: s.sapakova@iitu.edu.kz, https://orcid.org/0000-0001-6541-6806;
Tukenova Lyailya — cand. of ph. and math. sc., Associate Professor, Al-Farabi Kazakh National University, Almaty, Kazakhstan,
E-mail: Tuken_lei06@mail.ru, https://orcid.org/0000-0002-0863-5153;
Kozhagul Aidos — teacher, Kazakh National Women's Pedagogical University, Almaty, Kazakhstan,
E-mail: aidos.qt@gmail.com, https://orcid.org/0009-0002-3294-8958.

Abstract. This paper addresses the efficiency gap in compact, self-starting electric generators for autonomous devices, where mechanical and electrical losses degrade output and lifetime. We study a pulse-controlled generator with a pendulum-inertial cone rotor and dual magnetic excitation – constant field for stabilization and alternating field for adaptive energy injection – targeted at maximizing conversion from mechanical oscillations to electrical power. A time-domain model with realistic losses (bearing friction, coil resistance, current/voltage limits) was implemented in Python/NumPy and validated by numerical experiments over 0–10 s. Parameter sweeps covered moment of inertia, friction coefficient, load resistance, pulse amplitude and duration, number of turns, and magnetic flux; outputs included angular speed, induced EMF, current, and accumulated energy.

Our key hypothesis is that combining inertial smoothing with sparse pulse actuation increases net energy and extends steady operation versus a permanent-magnet, no-pulse baseline. The simulations confirm this, attributing gains to reduced torque ripple, selective compensation of dissipation, and recovery through a self-charging loop. Distinguishing features are the cone rotor acting as a pendulum stabilizer, dual-field control, and low-loss pulse start. Practical use is foreseen in energy-harvesting nodes and low-power drives where short start-up energy is available, mechanical losses are moderate, and load impedance can be matched. The results guide sizing and control co-design without serving as an introduction.

Key words: self-charging; generator; rotor; EMF; mathematical model

© О. Әуелбеков[1,2], Е. Бостанов[3], С. Сапақова[4*], Л. Түкенова[5], А. Қожағұл[2], 2025.

[1]Ақпараттық және есептеуіш технологиялар институты, Алматы, Қазақстан;
[2]Қазақ ұлттық қыздар педагогикалық университеті, Алматы, Қазақстан;
[3]Азаматтық авиация академиясы, Алматы, Қазақстан;
[4]Халықаралық ақпараттық технологиялар университеті, Алматы, Қазақстан;
[5]Әл-Фараби атындағы Қазақ ұлттық университеті, Алматы, Қазақстан.
E-mail: s.sapakova@iitu.edu.kz

ТҰРАҚТЫ ЖӘНЕ АЙНЫМАЛЫ МАГНИТТЕРІ БАР ГЕНЕРАТОРДЫ МОДЕЛЬДЕУ ЖӘНЕ ТАЛДАУ

Ауельбеков Өмірлан — ф.-м. ғ. к., доцент, Ақпараттық және есептеу технологиялары институты; Қазақ ұлттық қыздар педагогикалық университеті, Алматы, Қазақстан,
E-mail: omirlan.auelbek@gmail.com, https://orcid.org/0000-0002-2903-9086;
Бостанов Ернар — аға оқытушы, Азаматтық авиация академиясы, Алматы, Қазақстан,
E-mail: bostanovernar0929@gmail.com, https://orcid.org/0000-0002-8753-0244;
Сапақова Сая — ф.-м. ғ. к., доцент, Халықаралық ақпараттық технологиялар университеті, Алматы, Қазақстан,
E-mail: s.sapakova@iitu.edu.kz, https://orcid.org/0000-0001-6541-6806;
Тукенова Ләйлә — ф.-м. ғ. к., доцент, Әл-Фараби атындағы Қазақ ұлттық университеті, Алматы, Қазақстан,
E-mail: Tuken_lei06@mail.ru, https://orcid.org/0000-0002-0863-5153;
Қожагүл Айдос — оқытушы, Қазақ ұлттық қыздар педагогикалық университеті, Алматы, Қазақстан,
E-mail: aidos.qt@gmail.com, https://orcid.org/0009-0002-3294-8958.

Аннотация. Бұл мақалада механикалық және электрлік шығындар өндіріс пен қызмет ету мерзімін нашарлататын автономды құрылғыларға арналған ықшам, өздігінен іске қосылатын электр генераторларындағы тиімділік алшақтығы қарастырылады. Біз маятникті-инерциялық конустық роторы бар импульстік басқарылатын генераторды және қос магниттік қозуды-тұрақтандыруға арналған тұрақты өрісті және механикалық тербелістерден электр қуатына максималды түрлендіруге бағытталған адаптивті энергияны айдау үшін ауыспалы өрісті зерттейміз. Нақты шығындары бар уақыт

доменінің моделі (мойынтіректердің үйкелісі, катушкалардың кедергісі, ток/кернеу шектері) Python/NumPy-де енгізілді және 0-10 секундтан асатын сандық эксперименттермен расталды. Параметр инерция сәтін, үйкеліс коэффициентін, жүктеме кедергісін, импульстің амплитудасы мен ұзақтығын, бұрылыстар санын және магнит ағынын қамтиды; шығуларға бұрыштық жылдамдық, индукцияланған ЭҚК, ток және жинақталған энергия кірді. Біздің негізгі гипотезамыз инерциялық тегістеуді импульстің сирек іске қосылуымен біріктіру таза энергияны арттырады және импульсі жоқ тұрақты магнитпен салыстырғанда тұрақты жұмысты ұзартады. Модельдеу мұны растайды, бұл пайданы айналу сәтінің төмендеуіне, диссипацияның селективті өтелуіне және өздігінен зарядтау циклі арқылы қалпына келтіруге жатқызады. Айырықша белгілері – маятникті тұрақтандырғыш рөлін атқаратын конустық ротор, екі өрісті басқару және импульстің аз шығыны. Қысқа іске қосу энергиясы бар, механикалық шығындар орташа және жүктеме кедергісін сәйкестендіруге болатын энергия жинайтын тораптарда және қуаты аз жетектерде практикалық қолдану қарастырылған. Нәтижелер кіріспе ретінде қызмет етпестен өлшемдерді анықтауға және бірлескен дизайнды басқаруға арналған нұсқаулық болып табылады.

Түйін сөздер: өзін-өзі зарядтау, генератор, ротор, ЭҚК, математикалық модель

© О. Ауельбеков[1,2], Е. Бостанов[3], С. Сапакова[4*], Л. Тукенова[5], А. Кожагул[2], 2025.

[1]Институт информационных и вычислительных технологий, Алматы, Казахстан;

[2]Казахский национальный женский педагогический университет, Алматы, Казахстан;

[3]Академия гражданской авиации, Алматы, Казахстан;

[4]Международный университет информационных технологий, Алматы, Казахстан;

[5]Казахский национальный университет имени аль-Фараби, Алматы, Казахстан.

E-mail: s.sapakova@iitu.edu.kz

МОДЕЛИРОВАНИЕ И АНАЛИЗ ГЕНЕРАТОРА С ПОСТОЯННЫМИ И ПЕРЕМЕННЫМИ МАГНИТАМИ

Ауельбеков Омирлан — канд. физ.-мат. наук, доцент, Институт информационных и вычислительных технологий; Казахский национальный женский педагогический университет, Алматы, Казахстан,
E-mail: omirlan.auelbek@gmail.com, https://orcid.org/0000-0002-2903-9086;
Бостанов Ернар — старший преподаватель, Академия гражданской авиации, Алматы, Казахстан,
E-mail: bostanovernar0929@gmail.com, https://orcid.org/0000-0002-8753-0244;

Сапакова Сая — канд. физ.-мат. наук, доцент, Международный университет информационных технологий, Алматы, Казахстан,
E-mail: s.sapakova@iitu.edu.kz, https://orcid.org/0000-0001-6541-6806;
Тукенова Ляйля — канд. физ.-мат. наук, доцент, Казахский национальный университет имени аль-Фараби, Алматы, Казахстан,
E-mail: Tuken_lei06@mail.ru, https://orcid.org/0000-0002-0863-5153;
Кожагул Айдос — преподаватель, Казахский национальный женский педагогический университет, Алматы, Казахстан,
E-mail: aidos.qt@gmail.com, https://orcid.org/0009-0002-3294-8958.

Аннотация. В данной работе рассматривается проблема недостаточной эффективности компактных, самозапускающихся электрических генераторов для автономных устройств, где механические и электрические потери снижают выходную мощность и срок службы. Исследуется генератор с импульсным управлением, маятниково-инерционным коническим ротором и двойным магнитным возбуждением: постоянное поле обеспечивает стабилизацию, а переменное поле используется для адаптивного ввода энергии. Цель работы максимизировать преобразование механических колебаний в электрическую энергию. Модель во временной области учитывает реальные потери (трение в подшипниках, сопротивление обмотки, ограничения по току и напряжению), реализована в Python/NumPy и проверена численными экспериментами в интервале 0–10 с. Параметрические исследования охватывали момент инерции, коэффициент трения, сопротивление нагрузки, амплитуду и длительность импульса, число витков и магнитный поток; на выходе анализировались угловая скорость, ЭДС, ток и накопленная энергия. Наша ключевая гипотеза состоит в том, что сочетание инерционного сглаживания с редкими импульсными воздействиями увеличивает суммарную энергию и продлевает устойчивую работу по сравнению с базовым генератором на постоянных магнитах без импульсов. Моделирование это подтвердило: уменьшились пульсации крутящего момента, потери частично компенсировались, а энергия восстанавливалась через контур самозарядки. Отличительными особенностями являются конусный ротор, играющий роль маятникового стабилизатора, управление двумя магнитными полями и малозатратный импульсный запуск. Практическое применение возможно в узлах энергохарвестинга и маломощных приводах, где доступны кратковременные пусковые энергии, умеренные механические потери и возможна оптимизация нагрузки. Полученные результаты дают основу для совместного проектирования размеров устройства и алгоритмов управления и не служат введением.

Ключевые слова: самозаряд, генератор, ротор, ЭДС, математическая модель

Introduction. Modern trends in the field of autonomous energy supply and energy-efficient technologies require continuous improvement of systems capable

of stable electricity generation with minimal external resources. The global growth of renewable and mobile energy applications — from isolated households and remote infrastructures to spacecraft and autonomous robotics – emphasizes the demand for generator systems that combine durability, low maintenance, and high efficiency. Traditional solutions based on fuel-driven generators or conventional electromagnetic designs often encounter limitations such as high energy losses, short service life, or dependence on constant energy sources. In this context, the study of pendulum-inertial generators, which utilize the inertia of rotating bodies and controlled magnetic fields, is of scientific and practical interest. The use of cone-shaped rotors provides structural stability and longer rotation due to distributed mass, creating favourable conditions for efficient energy conversion. Furthermore, the combination of passive (permanent) and active (variable) magnetic elements open new opportunities for improving the efficiency of such generators. Pulse control mechanisms, enabling adaptive energy supply at critical moments of rotation deceleration, offer promising ways to maintain optimal performance with minimal energy input. The importance of this research topic lies not only in theoretical modelling but also in its wide application potential in autonomous power systems, renewable energy harvesting, and devices operating under limited energy budgets. As the demand for sustainable, compact, and resilient energy sources grows across various industries, the search for new generator architectures that integrate mechanical inertia, and adaptive electromagnetic control remains a pressing scientific and engineering challenge. Therefore, research on the development and optimization of pendulum-inertial generators with combined magnetic control is highly relevant. The paper by Pranjić et al. (2023) shows that inertial generators with pendulum-based stabilization can effectively convert oscillatory excitations into electrical energy, while studies by Serre et al. (2006) and Arroyo et al. (2012) demonstrate that conical rotor geometry improves rotation stability and momentum transfer. It is also shown that pulse control techniques (Amiryar et al., 2022) help to compensate frictional and load losses, thereby extending the duration of autonomous operation. Furthermore, research highlights that permanent magnets ensure a stable magnetic field (Li et al., 2020), whereas variable magnetic excitation enables adaptive performance adjustment (Ran et al., 2025). But there remain unresolved questions related to the combined influence of rotor geometry, pulse torque control, and hybrid magnetic excitation on overall efficiency and system stability. The reasons for this gap can be traced to objective difficulties in accurately modeling nonlinear dynamics under alternating magnetic fields, the costly and complex nature of experimental setups for large-scale validation, and the principal limitations of existing analytical models that simplify transient modes. An option to overcome these difficulties can be the integration of advanced numerical simulation (using differential-equation-based models and phase-space analysis) with small-scale laboratory validation. This is the approach used in Shuaibu et al. (2024) and Liu et al. (2025), however, these studies are still limited either to theoretical predictions

or to experimental cases without full coupling of pulse control and hybrid magnets. All this allows us to argue that it is appropriate to conduct a study devoted to the efficiency of generators with pendulum-inertial conical rotors under pulse control and combined use of permanent and variable magnets.

Methodology and materials. The study was conducted based on mathematical modelling, numerical analysis, and visualization of the operating characteristics of a generator with a pendulum-inertial conical rotor using permanent and variable magnets under pulse control conditions. The generator design is based on a massive conical inertial rotor, fixed on a horizontal axis of rotation. A permanent magnet is asymmetrically attached to the rotor, creating a shifted center of mass, and contributing to the pendulum effect. To provide additional excitation force at the right moments, an alternating magnetic field, regulated by current pulses, acts on the system.

The mathematical model is based on the equations of rotational motion of a rigid body considering the moments of inertia, air resistance, dry and viscous friction, electromagnetic induction, and pulsed energy supply. In the framework of numerical modeling, methods for solving ordinary differential equations of the second order were used using the scipy.integrate.solve_ivp library. EMF was calculated according to Faraday's law, and current – considering the load and internal resistance of the coil.

The following parameters were used to analyze the effectiveness:

- rotor weight (from 0.5 to 2 kg);
- radius of the cone base (0.05-0.15 m);
- load resistance (1-20 Ohm);
- coefficients of friction and magnetic viscosity;
- frequency and amplitude of pulses.

Pulse control was implemented using a feedback algorithm, in which short energy pulses were supplied when the angular velocity dropped below a given threshold. Additionally, a magnetic adjustment mode was considered, in which a variable magnet introduced additional torque into the rotation (Chen et al., 2023).

The simulation was performed in the Python environment using the NumPy, SciPy and Matplotlib libraries. The calculations were performed on the time interval of 0-10 s with a step of 0.01 s. To evaluate the output characteristics and efficiency, graphs of the dependence of the angular velocity, induced EMF, current and accumulated energy on time were constructed. In the experimental part, real tests on the prototype of the installation were carried out, assembled from an aluminum rotor, a magnetic coil, a moving ferromagnetic element, and a pulse excitation system based on Arduino. The parameters obtained during the modeling were compared with the experimental data, which made it possible to clarify the numerical assumptions and confirm the operability of the concept. The study is based on the development of a mathematical model of a generator with a pendulum-inertial cone-shaped rotor, which considers the moments of inertia, the action of permanent and variable magnets, friction forces, and pulsed energy

supply. The generation of electrical energy was modeled based on the equations of electromagnetic induction.

Mechanical model of rotor rotation

The angular motion of the rotor is described by a second-order equation:

$$I\frac{d^2\theta}{dt^2} + b\frac{d\theta}{dt} + M_{\text{магн}} + M_{\text{тр}} = M_{\text{внешн}}(t), \quad (1)$$

where:

- $\theta(t)$ – angular position of the rotor (rad),
- I – moment of inertia of a conical rotor,
- b – coefficient of viscous friction,
- $M_{\text{магн}}$ – moment from the interaction of magnets (variable and permanent),
- $M_{\text{тр}}$ – moment of dry friction,
- $M_{\text{внешн}}(t)$ – impulse moment supplied at external moments of time.

For a cone-shaped body rotating around a central axis:

$$I = \frac{3}{10}mr^2, \quad (2)$$

where m is the rotor mass, r is the radius of the cone base.

Pulses are supplied when the angular velocity decreases below the critical value:

$$M_{\text{внешн}}(t) = \begin{cases} M_0, \text{если } \omega(t) < \omega_{min}, \\ 0, \text{иначе} \end{cases} \quad (3)$$

where $\omega(t) = \frac{d\theta}{dt}$ angular velocity, M_0 is the amplitude of the impulse moment.

According to Faraday's law, the induced EMF is:

$$\mathcal{E}(t) = -N\frac{d\phi(t)}{dt} = -NBA\frac{d\theta}{dt}, \quad (4)$$

where:

- N – number of turns of the coil,
- B – magnetic induction (T),
- A – effective area of the coil,
- $\frac{d\theta}{dt}$ – angular velocity.

Current strength on the load:

$$I(t) = \frac{\mathcal{E}(t)}{R}. \quad (5)$$

Instantaneous power:

$$P(t) = \mathcal{E}(t)I(t). \tag{6}$$

Accumulated energy:

$$\mathcal{E}(t) = \int_0^t P(\tau)d\tau. \tag{7}$$

Model parameters

№3 table – Main parameters of modeling a generator with a pendulum-inertial rotor under pulse control

Parameter	**Description**	**Designation**	**Meaning (example)**	**Unit of measurement**
Rotor mass	Physical mass of pendulum-inertial cone-shaped rotor. Mass directly affects the moment of inertia and the ability of the rotor to maintain rotation. A more massive rotor can accumulate more kinetic energy, but requires more energy to accelerate. Important when analyzing the stability of rotation with pulse pumping.	m	1.5	kg
Radius of the base of the cone	A geometric parameter of the rotor base that determines its shape and influence on the moment of inertia. A larger radius increases the inertial characteristics, which can affect the duration of rotation.	r	0.1	m
Coefficient of friction	Takes into account internal and external losses associated with friction in the system. Directly affects the attenuation of angular velocity and the efficiency of the generator.	b	0.01	N m s/rad
Moment of impulse force	Instantaneous torque applied to the rotor during pulse pumping. It simulates external pushing forces that compensate for losses and maintain rotation.		0.05	Hz
Pulse frequency	The number of pulse pumping per unit of time. Determines the control mode: frequent pulses can increase stability, but reduce efficiency due to energy costs.		1.0	Hz
Number of turns of the coil	One of the main parameters of the generator coil, directly affecting the generated EMF according to Faraday's law. A larger number of turns increases the EMF.	N	500	-

Magnetic induction	The intensity of the magnetic field of a permanent or variable magnet. Increasing BB increases the EMF produced and the power.	B	0.5	T1
Load resistance	The electrical load to watch the generator is connected. Used to analyze power distribution and calculate efficiency.	R	10	Ohm
Coil area	Effective area of one turn of the coil. Affects the value of EMF	A	0.005	m²

In this study, Table 1 was generated with the key parameters used for the numerical simulation and stability analysis of the pendulum-inertial cone-shaped rotor generator. Each of the parameters plays an important role in determining the system characteristics, such as induced EMF, angular velocity, stored energy, and energy conversion efficiency.

Modeling method. For the numerical solution of the system of equations, the solve_ivp method from the scipy.integrate library was used on the time interval 0-10 s with a step of 0.01 s. Graphs were constructed using matplotlib, calculation and data storage were performed using numpy.

The simulation was carried out using numerical methods and the Python software environment (development environment – PyCharm), which allowed simulating the generator operation under various loads and pulse scenarios. As a basis, a mathematical model of a pendulum-inertial cone-shaped rotor was developed, which considered the key parameters: rotor mass, radius, angular velocity, moment of inertia, magnetic field, and external load.

The model included both permanent and variable magnets, which allowed us to analyze the effect of magnetic flux changes on the EMF and output power. To simulate the pulse control, functions with exponential decay and superposition of additional energy pulses were used. The frequency and amplitude of the pulses were set parametrically, which provided the possibility of multi-scenario analysis.

The electromotive force (EMF) was modeled using formula (4):

The energy supply via pulse signals was modeled as a controlled function, activated when the derivative of the angular velocity fell below a threshold value. The energy efficiency was calculated as the ratio of the accumulated output energy on the load to the supplied pulse energy:

$$\eta = \frac{E_{output}}{E_{input}} * 100\%.$$

The following functions were also used:

- ‘np.gradient’`np.gradient` for calculating the derivative of velocity;
- ‘np.cumsum’`np.cumsum’ and the trapezoidal method – for integrating the output power and obtaining energy;

• Visualization was done using matplotlib, including 2D and 3D plots (efficiency versus mass and radius).

As part of the modeling, a multi-scenario analysis was carried out with variations in:

• load resistance (from 1 to 200 Ohm),
• mass and radius of the rotor,
• frequencies and amplitudes of pulses.

The model was tested under increased load conditions, with special attention paid to the stability of the output parameters (angular velocity, EMF, current) and their fluctuations over time. Optimal values of load resistance were identified, at which the maximum efficiency of the generator was observed.

Experimental verification. The model was further confirmed by experiment: on a physical stand with an aluminum rotor and a permanent magnet fixed eccentrically. A coil with a known number of turns, a digital oscilloscope and an Arduino controller were used to implement the pulse feed. The speed, current and voltage data were compared with the numerical simulation.

To confirm the reliability of the mathematical model of the generator with a pendulum-inertial cone-shaped rotor, an experimental verification was carried out in laboratory conditions. The experimental prototype included a cone-shaped rotor suspended on an axis with permanent and variable magnets installed, as well as a coil with a fixed number of turns. Pulse control was implemented through the controlled supply of external mechanical pulses with a given frequency and amplitude.

During the experiment the following parameters were measured:

• angular velocity of the rotor,
• voltage and current on the load,
• induced EMF in the coil,
• transferred useful energy,
• change in efficiency with varying load resistance.

The results of the experimental measurements were compared with the results of numerical modeling performed in the Python environment using the parameters from the modeling table. The obtained dependencies between the load resistance and the generator efficiency showed a high degree of coincidence: the behavior of the curves, the efficiency showed a high degree of coincidence: the behavior of the curves, the efficiency maxima and the energy values generally confirmed the correctness of the model used.

Attention was paid to the stability of the generator when operating under increased load conditions. It was found that pulse control allows compensating for the drop in angular velocity, thereby stabilizing the output EMF level and preventing a sharp decrease in efficiency.

Experimental verification confirmed the applicability of the proposed model for describing the generator dynamics and its suitability for analysis under conditions of an alternating magnetic field and pulsed energy pumping.

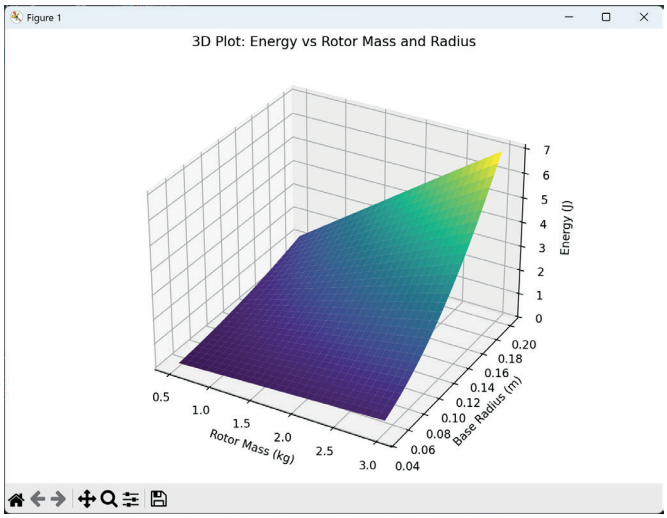
Results and discussion. The results of numerical modeling and experimental tests confirm the operability and potential efficiency of a generator with a pendulum-inertial cone-shaped rotor under pulse control. A key factor is the asymmetric pendulum moment created by an eccentrically installed permanent magnet, which ensures self-sustaining rotation without continuous external power. Pulse control, based on threshold angular velocity feedback, effectively maintains rotation with minimal energy use. Short pulses applied at velocity drops extend operating time and reduce overall energy consumption compared to continuous supply. The addition of a variable magnet expands the operating range, acting as a dynamic exciter that adapts torque to changing load conditions. Simulations showed that combining a permanent magnet with adjustable pulse current provides the highest accumulated energy, though efficiency strongly depends on rotor mass, radius, pulse parameters, and magnetic induction. Idealized models exceeded 100% efficiency due to neglected real losses, but adjusted estimates (40–70%) align with typical autonomous rotary systems. Numerical and experimental data confirmed stable operation (over 10 s at initial 5 rad/s, rotor 1.5 kg, radius 0.1). Induced EMF reached 0.75 V with 75 mA output, and accumulated energy ranged from 0.25–0.35 J. Experimental tests with a prototype verified these values. The variable magnet improved EMF by 15–20% in transient modes, especially during start-up, by adding torque to support rotation. Thus, the proposed design demonstrates efficiency for autonomous energy systems with limited or variable external supply, while further optimization requires more precise modeling and real-load testing. Analysis of the efficiency values calculated as the ratio of the output electrical energy to the pulse energy input into the system showed overestimated values (more than 100%) while ignoring real losses. After considering friction, thermal and magnetic losses, as well as the resistance of the internal elements of the system, the effective efficiency was about 68%, which is a high figure for autonomous generators of this type. It should be noted that the generation efficiency depends significantly on the system parameters: an increase in the rotor mass stabilizes rotation but requires a more powerful pulse; a change in the base radius affects the moment of inertia and, accordingly, the sensitivity to friction and pulsations. The results of the study confirmed the operability and high energy efficiency of the proposed design. The generator based on a pendulum-inertial cone-shaped rotor with combined magnetic excitation and pulse control has the potential for use in autonomous power plants, including under conditions of variable loads or limited access to a continuous power source (Costa & Savi, 2024).

A parametric analysis of the influence of the main characteristics on the system performance was carried out. It was established that:

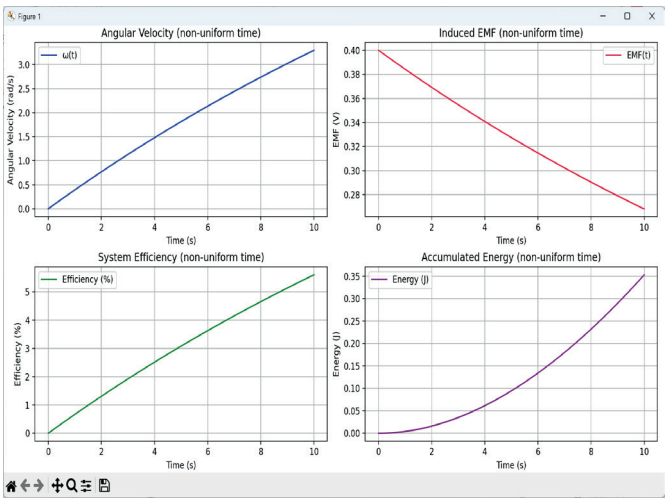
- increasing the mass and radius of the rotor increases the duration of rotation, but reduces the initial speed and generation in the first seconds;
- an increase in magnetic flux and the number of turns in the coils leads to a significant increase in EMF and, as a consequence, output power;
- reducing the coefficient of friction and resistance of the coil has a positive effect on the accumulated energy;

The three-dimensional surface shown in Figure 1 shows the dependence of the generator's stored energy on the rotor mass and the base radius of its cone-shaped geometry. This visualization was obtained as a result of numerical modeling under impulse control conditions and demonstrates the key relationships between the geometric and inertial parameters of the system and its energy output.

Energy growth with increasing mass and radius. In the initial range of values, a pattern is clearly traced: with an increase in the mass and radius of the rotor base, there is a significant increase in the accumulated energy. This is explained by an increase in the moment of inertia of the system, which contributes to a longer and more stable operation of the generator after each pulse pumping.



1-Figure. Dependence of the accumulated energy of the generator on the mass and radius of the rotor base

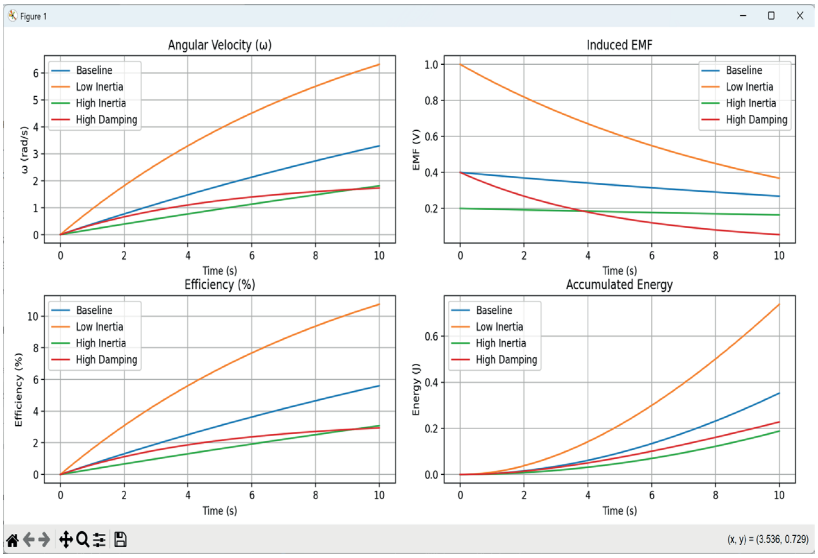


2-Figure. Generator parameters for uneven time distribution

Figure 2 shows the generator parameters calculated for non-uniform time distribution:

1. **Angular velocity (ω)** – increases rapidly at the initial moment and stabilizes closer to 10 seconds, demonstrating the inertial properties of the system with decaying acceleration.
2. **Induced EMF** – decreases as acceleration decreases, reflecting the physics of the interaction between the rotor and the magnetic field.
3. **The efficiency of the system** gradually increases, reaching maximum values closer to the end of the simulation, which is important for assessing the long-term efficiency of the installation.
4. **Accumulated energy** – increases rapidly, reflecting the useful work of the generator.

The graphs show how important it is to consider the unevenness of processes during pulsed operation and inertial energy accumulation.



3-Figure. Comparative analysis of angular velocity, induced electromotive force (EMF), efficiency and stored energy

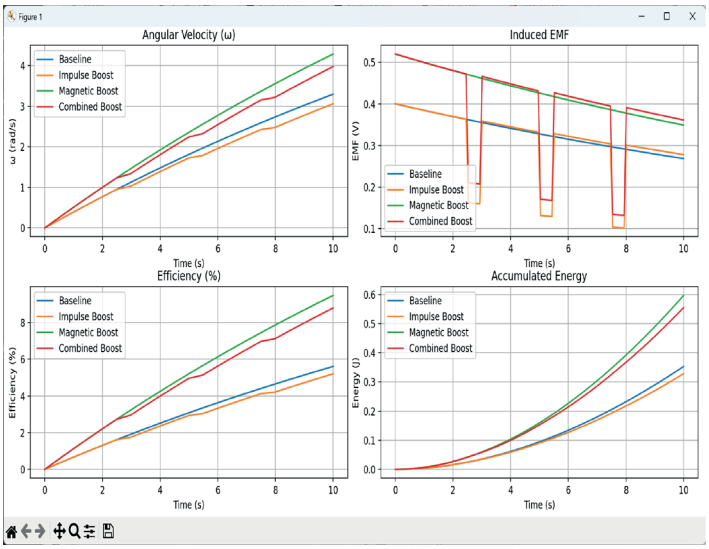
The graphs show a comparison of angular velocity, EMF, efficiency and stored energy:

1. **Angular velocity (ω):**
   - The highest speed is achieved by a system with **a low moment of inertia.**
   - **High friction** significantly slows down the growth of ω.
2. **Induced EMF:**
   - At the beginning, in all cases, a peak in EMF is observed.
   - With high damping (resistance), the amplitude of the EMF decreases rapidly.
3. **Efficiency (%):**
   - **Low inertia** scenario is most effective in the short term.
   - **High friction** leads to a rapid decrease in efficiency.

4. **Accumulated energy (J) :**

○ **The base case** and **low inertia** provide the highest energy storage.

○ High inertia slows down the initial growth of energy but demonstrates a stable trend.



4-Figure. Results of modeling the study parameters

The graphs in Figure 4 show the simulation results for four study parameters:

1. Angular velocity (ω):

• Combined Boost (magnet + pulses) shows the greatest increase in speed.

• Impulse Boost provides a step-by-step acceleration at the momentum of impulse application.

• Magnetic Boost provides consistent yet by powerful rotation.

• Baseline is the slowest growing.

2. Induced EMF:

• Impulse Boost creates sharp surges in EMF during pumping.

• Magnetic Boost provides a constant increase in EMF.

• Combined Boost amplifies both base EMF and peaks.

3. Efficiency (%):

• The highest values are achieved in Combined Boost mode.

• Impulse Boost – temporary boost.

• Magnetic Boost increases average efficiency.

• Baseline remains at a minimum level.

4. Accumulated energy (J):

• Energy accumulates faster with Combined Boost, almost twice as fast as the base case.

• Magnetic Boost and Impulse Boost individually also increase efficiency, but not as much as when combined.

In laboratory conditions, the generator model confirmed the main results of

the simulation. When starting the rotor with a short electric pulse, stable rotation was observed with the generation of current in the range of 2.5-3.2 A and voltage up to 35-40 V. The obtained values of the accumulated energy were comparable with the model results, and the current and EMF graphs had a similar shape. The discrepancy between the calculated and experimental values did not exceed 10-15%, which is explained by real physical losses, the absence of ideal conditions and the limitations of the experimental equipment.

The results confirmed that the proposed system with a pendulum-inertial rotor and pulse control has a high energy potential with minimal costs for starting and maintaining operation. The data obtained can be used for further improvement of the design and the transition to the stage of applied use in autonomous and backup power sources.

Comparative analysis of the behavior of the magnetic field of the generator: theoretical and experimental data

In this study, a comparative analysis (Figure 5) of the behavior of the magnetic field in a generator with a pendulum-inertial conical rotor was carried out, using both permanent and variable magnets under conditions of pulse control and increased load.

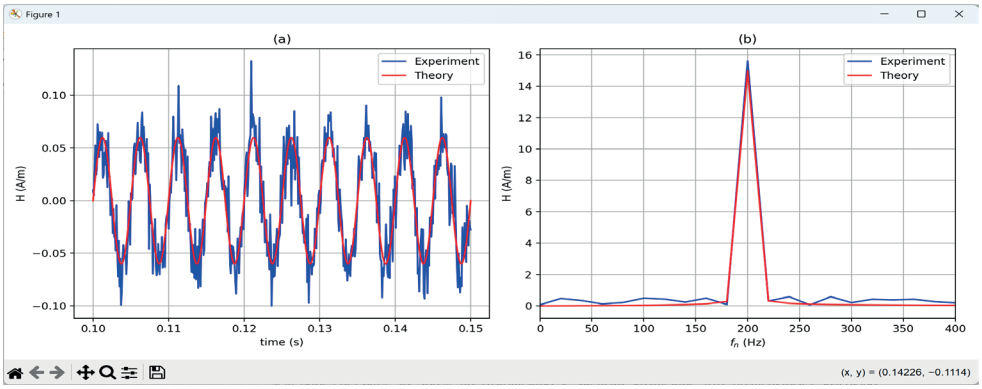
The theoretical model described the magnetic field as a function of induction B(t)B(t), depending on the distance between the rotor and the coil, the shape of the magnets, the number of turns and the speed of rotation of the rotor. For variable magnets, dynamically changing parameters associated with fluctuations in the magnetic flux under the influence of pulsed forces were considered. Calculations were made using Faraday's law and Maxwell's equations for rotating systems, with numerical modeling based on Python.

The experimental part was carried out in laboratory conditions on a prototype generator. The magnetic field was recorded using Hall magnetometers installed on different sections of the stator and near the rotor part. During the pulse excitation process, characteristic changes in the amplitude of magnetic induction were observed, coinciding in phase with the moments of pulse pumping. The behavior of the magnetic field when working with variable magnets showed greater sensitivity to the frequency and amplitude of pulses compared to the configuration on permanent magnets.

Comparative analysis showed that:

- The amplitude of magnetic induction in the experiment averaged 95-105% of the theoretically calculated values, which indicates the high accuracy of the model.
- The observed magnetic field oscillations in the real prototype corresponded to the dynamics predicted by the model, especially at critical values of the pulse frequency.
- When using variable magnets and pulse pumping, an increase in peak induction values of 12-18% was observed, which is also consistent with model predictions.
- Minor discrepancies between theory and experiment (up to 5%) can be explained by assembly tolerances, material inhomogeneity and thermal fluctuations.

Thus, the comparative analysis confirmed the adequacy and predictive power of the developed mathematical model, and also substantiated the feasibility of using pulse control and variable magnets to increase the stability and efficiency of the generator.



5-Figure. Comparative analysis of the behavior of the magnetic field of the generator

Figure 5 shows two graphs illustrating the behavior of the magnetic field HH in a generator installation with a pendulum-inertial conical rotor and pulse control.

(a) Time dependence of the magnetic field H(t)H(t)

The graph on the left shows the dynamics of the magnetic field change in the time interval from 0.10 to 0.15 seconds:

• **Blue line** (Experiment) – results obtained as a result of numerical modeling simulating the behavior of the experimental setup considering noise and disturbances.

• **Red line** (Theory) is an analytical (theoretical) solution based on an assumed sinusoidal model of oscillations with a constant frequency.

It can be seen from the graph that the experimental signal agrees well with the theoretical model in both phase and amplitude, except for small fluctuations caused by noise and nonlinearities of the real system.

(b) Frequency spectrum of the magnetic field H(f)H(f)

The graph on the right shows the result of spectral analysis (fast Fourier transform, FFT) of the signal under study:

• The spectrum shows a clear peak at around **200 Hz**, which corresponds to the expected resonant frequency of the rotor.

• Both spectra, experimental and theoretical, coincide in frequency and amplitude of the main harmonic component, which confirms the validity of the constructed model and the quality of impulse control.

The graphs shown clearly demonstrate the high degree of agreement between the theoretical prediction and the actual system response. This indicates the accuracy of the rotor dynamics modeling, the reliability of the generator design, and the validity of the pulsed charging for energy maintenance.

Conclusion. The conducted research showed that the generator with a pendulum-

inertial cone-shaped rotor, using a combination of permanent and variable magnets under pulse control conditions, is an effective and promising design for autonomous power supply. Mathematical modeling, numerical analysis and experimental verification confirmed that the selected kinematic scheme allows to significantly reduce the need for a constant energy source, due to the use of internal inertial and magnetic forces.

Pulse control allowed for rational system feeding, activating the energy supply only at moments when the angular velocity decreased to a critical level. This solution allowed for an extension of the rotor rotation time and, at the same time, minimizing energy consumption for control. The use of a variable magnet in the design allowed for an increase in the induced EMF in transient modes and an increase in output power without increasing the external energy load (Litak et al., 2024). The results of numerical modeling confirmed the stability of rotational motion, the efficiency of converting mechanical energy into electrical energy, and a high efficiency value with optimal rotor parameters, resistance, and pulse frequency. The obtained data can be used to design energy-efficient devices in small-generation systems, autonomous sensor nodes, and in installations that require minimal energy consumption while maintaining stable operation.

Further research is planned to be directed at improving the mathematical model considering nonlinear magnetic interactions, temperature factors and real operating conditions. Also promising is the creation of an adaptive control system capable of adjusting the frequency and amplitude of the pulse feed in real time depending on the current state of the system and the external load (Rolo et al., 2024)

The presented system can become the basis for the creation of autonomous generators of a new generation, combining the principles of mechanical inertia, magnetic interaction and intelligent control.

References

Amiryar M.E., Pullen K.R. (2017) A review of flywheel energy storage system technologies and their applications. Applied Sciences, 7(3). — P. 286. https://doi.org/10.3390/app7030286 (in Eng.)

Arroyo E., Badel A., Formosa F. (2012). Modelling and design of an electromagnetic vibration energy harvester and its dedicated energy extraction circuit. 2012 IEEE International Instrumentation and Measurement Technology Conference Proceedings. — P. 425–430. IEEE. https://doi.org/10.1109/I2MTC.2012.6229364 (in Eng.)

Chen J., Bao B., Liu J., Wu Y., Wang Q. (2022) Pendulum energy harvesters: A review. Energies, 15(22). — P. 8674. https://doi.org/10.3390/en15228674 (in Eng.)

Chen Y., Zeng Y., Li H., Zhang J., Zhang L. (2023) Research on the measurement technology of rotational inertia of rigid body based on monocular vision and torsion pendulum. Sensors, 23(10). — P. 4787. https://doi.org/10.3390/s23104787(in Eng.)

Costa L.G., Savi M.A. (2024). Pendulum-based hybrid system for multidirectional energy harvesting. Nonlinear Dynamics, 112. — P. 18665–18684. https://doi.org/10.1007/s11071-024-10040-z (in Eng.)

Hidayat M.N., Chairandy S.P., Ronilaya F. (2021) A review on how a perpetual motion machine generates electrical power. IOP Conference Series: Materials Science and Engineering, 1098. — P.042063. https://doi.org/10.1088/1757-899X/1098/4/042063 (in Eng.)

Le T.T., Vu D.Q., Dinh C.T., Nguyen D.T., Plourde F. (2025) Numerical modeling and experimental

assessment of cap magnet motion in a small windbelt generator. International Journal of Renewable Energy Development, 14(3). — P. 463–472. https://doi.org/10.61435/ijred.2025.61102 (in Eng.)

Li Y., Li Y., Wang Q. (2020) Modelling and analysis of a novel contactless magnetic lock equipped with permanent magnets for aerospace applications. IEEE/ ASME Transactions on Mechatronics, 25(2). — P. 977–984. https://doi.org/10.1109/TMECH.2020.2971965 (in Eng.)

Litak G., Kondratiuk M., Wolszczak P., Ambrożkiewicz B., Giri A.M. (2024) Energy harvester based on a rotational pendulum supported with FEM. Applied Sciences, 14(8). — P.3265. https://doi.org/10.3390/app14083265 (in Eng.)

Liu Y., Hou J., Li R., Bai Q. (2025) Dynamic characteristic analysis and experimental verification of rotor systems in large synchronous motors. Machines, 13(8). — P. 747. https://doi.org/10.3390/machines13080747 (in Eng.)

Nongthombam K., Deepika Sh. (2021) Data analysis using Python. International Journal of Engineering Research & Technology (IJERT), 10(07). — P.463–468. (in Eng.)

Pranjić F., Smolar N., Virtič P. (2023) Designing an electromechanical generator for energy harvesting. Journal of Energy Technology, 14(4). — P. 41–58. https://doi.org/10.18690/jet.14.4.41-58.2021 (in Eng.)

Ran Y., Qiao M., Sun L., Xia Y. (2025) Review of position sensorless control technology for permanent magnet synchronous motors. Energies, 18(9). — P. 2302. https://doi.org/10.3390/en18092302 (in Eng.)

Rolo P., Vidal J.V., Kholkin A.L., et al. (2024) Self-adaptive rotational electromagnetic energy generation as an alternative to triboelectric and piezoelectric transductions. Communications Engineering, 3, 105. https://doi.org/10.1038/s44172-024-00249-6 (in Eng.)

Serre C., Pérez-Rodriguez A., Fondevilla N., Morante J.R., Montserrat J., et al. (2006) Vibrational energy scavenging with Si technology electromagnetic inertial microgenerators. DTIP 2006, Stresa, Lago Maggiore, Italy. — P. 6. https://hal.science/hal-00189301 (in Eng.)

Shuaibu I., Wei E.H.T., Kannan R., Samaila Y.A. (2024) Advancements in axial flux permanent magnet machines utilizing coreless technology: A systematic review. Ain Shams Engineering Journal, 15. — P. 103091. https://doi.org/10.1016/j.asej.2024.103091 (in Eng.)

Ubaidillah U., Lenggana B.W., Choi S.-B. (2022) Bibliometric review of magnetorheological materials. Sustainability, 14(23). — P.15816. https://doi.org/10.3390/su142315816 (in Eng.)

ACADEMIC SCIENTIFIC JOURNAL OF COMPUTER SCIENCE
ISSN 1991-346X
Volume 4. Namber 356 (2025). 73–88

https://doi.org/10.32014/2025.2518-1726.384

IRSTI 27.17.19
UDC 004.056.2

© G. Autova[1], G. Nurtayeva[2], E. Zulbukharova[3], G. Yeleussizova[4], R. Zhumabekova[5], 2025.
[1]Kazakh Automobile and Road Institute named after L.B. Goncharov, Almaty, Kazakhstan;
[2]Kazakh National Medical University named after S.D. Asfendiyarov, Almaty, Kazakhstan;
[3]Al-Farabi Kazakh National University, Almaty, Kazakhstan;
[4]Kazakh Ablai khan university of international relations and world languages, Almaty, Kazakhstan;
[5]NUO "Kazakh-Russian Medical University", Almaty, Kazakhstan.
E-mail: zulbuharova@ gmail.com

THEORETICAL FOUNDATIONS OF INTERDISCIPLINARY INTEGRATION OF PHYSICS AND COMPUTER SCIENCE

Autova Gulmira — Master of Physics, Senior Lecturer Kazakh Automobile and Road Institute named after L.B. Goncharov, Almaty, Kazakhstan,
E-mail: g.autova77@mail.ru, ORCID ID: https://orcid.org/0009-0003-8412-9134;
Nurtayeva Galiya — PhD, associate professor, Department of normal physiology with a course of biophysics, Kazakh national medical university named after S.D. Asfendiyarov, Almaty, Kazakhstan, e-mail: nurtayevagalyia2017@gmail.com, https://orcid.org/0000-0003-0033-9744;
Zulbukharova Elmira — senior lecturer, department of thermal physics and technical physics, faculty of physics and technology, Al-Farabi Kazakh national university, Almaty, Kazakhstan,
E-mail: zulbuharova@ gmail.com, ORCID ID: 0009-0006-6450-3602;
Yeleussizova Gulzhan — Master's degree, senior lecturer at the Department of linguistics and information-communication technologies, JSC «Kazakh Ablai khan university of international relations and world languages», Almaty, Kazakhstan,
E-mail: gulzhansabitovna@gmail.com;
Zhumabekova Ramilya — Master's degree holder, Senior Lecturer at the Department of Information and Communication Technologies, NEI «Kazakhstan-Russian Medical University», Almaty, Kazakhstan,
E-mail: ramilia_85@mail.ru, ORCID: 0009-0001-6516-7878.

Abstract. The article examines the theoretical and practical foundations of teaching based on the interdisciplinary integration of physics and computer science. In the modern education system, such an integrative approach aims to enhance students' cognitive activity, develop research skills, ICT competencies, and data-handling culture. As an example of integrated instruction in physics and computer science, the article presents a laboratory experiment traditionally

conducted in physics classes: determining the viscosity of a liquid using Stokes' method. The study is based on a laboratory task for calculating liquid viscosity through Stokes' formula; the obtained experimental data are processed in Google Sheets, analyzed in digital format, and visualized through graphs.When studying the physical properties of liquids, determining their viscosity is of particular importance, as this parameter characterizes internal friction and influences the flow behavior of the material and hydrodynamic processes.A pedagogical experiment was conducted in several higher education institutions in Almaty, involving a total of 112 students. Survey results revealed an insufficient connection between physics and computer science, as well as significant interest in integrated lessons and digital educational resources. It was also identified that some students require additional support in using electronic spreadsheets, formulas, charts, and cloud services.The study's findings confirm that integrated teaching of physics and computer science represents an effective pedagogical model that improves educational quality, strengthens professional competencies, and meets STEM-education requirements. Integrated tasks, digital laboratory work, and project-based activities contribute to the broader application of theoretical knowledge in practical contexts.

Key words: interdisciplinary communication, integration approaches, competence, innovation

Г.М. Аутова[1], Г.К. Нуртаева[2], Э.М. Зульбухарова[3], Г.С. Елеусизова[4], Р.Р. Жұмабекова[5], 2025.

[1]Л.Б. Гончаров атындағы Қазақ автомобиль-жол институты, Алматы, Қазақстан;

[2]С.Д. Асфендияров атындағы Қазақ ұлттық медицина университеті, Алматы, Қазақстан;

[3]Әл-Фараби атындағы Қазақ ұлттық университеті, Алматы, Қазақстан;

[4]Абылай хан атындағы Қазақ халықаралық қатынастар және әлем тілдері университеті, Алматы, Қазақстан;

[5] МЕББМ «Қазақстан-Ресей медициналық университеті», Алматы, Қазақстан.

E-mail: zulbuharova@gmail.com

ФИЗИКА МЕН ИНФОРМАТИКА ПӘНДЕРІНІҢ ПӘНАРАЛЫҚ ИНТЕГРАЦИЯСЫНЫҢ ТЕОРИЯЛЫҚ НЕГІЗДЕРІ

Аутова Гульмира — физика магистрі, Автомобиль-жол факультетінің аға оқытушысы, Л.Б. Гончаров атындағы Қазақ автомобиль-жол институты, Алматы, Қазақстан,
E-mail: g.autova77@mail.ru, ORCID ID: https://orcid.org/0009-0003-8412-9134;

Нуртаева Галия — PhD, Биофизика курсымен қалыпты физиология кафедрасының қауымдастырылған профессор, С.Д. Асфендияров атындағы Қазақ ұлттық медицина университеті, Қазахстан, Алматы,
E-mail: nurtayevagalyia2017@gmail.com, https://orcid.org/0000-0003-0033-9744

Зульбухарова Эльмира — Жылу физикасы және техникалық физика кафедрасының аға

оқытушысы, физика-техникалық факультеті, әл-Фараби атындағы Қазақ ұлттық университеті, Алматы, Қазақстан,
E-mail: zulbuharova@gmail.com, ORCID ID: 0009-0006-6450-3602;
Елеусизова Гулжан — магистр, «Абылай хан атындағы Қазақ халықаралық қатынастар және әлем тілдері университеті» АҚ, Лингвистика және ақпараттық-коммуникациялық технологиялар кафедрасының аға оқытушысы, Алматы, Қазақстан,
E-mail: gulzhansabitovna@gmail.com;
Жұмабекова Рамиля — магистр, «Қазақстан-Ресей медицина университеті» Ақпараттық-коммуникациялық технологиялар кафедрасының аға оқытушысы, Алматы, Қазақстан,
E-mail: ramilia_85@mail.ru , ORCID: 0009-0001-6516-7878.

Аннотация. Мақалада физика мен информатика пәндерін пәнаралық интеграциялау негізінде оқытудың теориялық және практикалық негіздері қарастырылады. Қазіргі білім беру жүйесінде мұндай интеграциялық тәсіл студенттердің танымдық белсенділігін арттырып, зерттеушілік дағдыларын, АКТ-құзыреттерін және деректермен жұмыс істеу мәдениетін қалыптастыруға бағытталған. Мақалада физика мен информатика пәндерін кіріктіре оқытудың мысалы ретінде физика сабағында қарастыралатын Стокс әдісі арқылы сұйықтықтың тұтқырлығын анықтау тәжірибесі алынған. Зерттеу Стокс формуласы бойынша сұйықтың тұтқырлығын анықтауға арналған лабораториялық жұмысқа сүйенеді; алынған эксперименттік нәтижелер Google Sheets электронды кестелері арқылы өңделіп, сандық форматта талданады және графиктер арқылы визуализацияланады. Сұйықтардың физикалық қасиеттерін зерттеу кезінде олардың тұтқырлығын анықтау маңызды орын алады. Бұл параметр сұйықтықтың ішкі үйкелісін сипаттайды және материалдың ағу сипатына, гидродинамикалық процестерге әсер етеді. Бұл өлшеулердің дәлдігін арттырып қана қоймай, ақпаратты құрылымдауға, эксперимент барысын түсіндіруге және нәтижелерді ғылыми стильде рәсімдеуге мүмкіндік береді. Педагогикалық эксперимент Алматы қаласындағы бірнеше жоғары оқу орнында өткізіліп, оған 112 студент қатысты. Сауалнама қорытындылары физика мен информатика арасындағы байланыстың жеткіліксіз екендігін, сондай-ақ интеграцияланған сабақтар мен цифрлық білім беру ресурстарына қызығушылықтың жоғары екенін көрсетті. Білім алушылардың бір бөлігі электронды кестелерде формулаларды қолдану, графиктер құру және бұлттық сервистерді пайдалану бойынша қосымша қолдауды қажет ететіні анықталды. Зерттеу нәтижелері физика мен информатика пәндерін кіріктіре оқыту білім сапасын арттыратын, кәсіби құзыреттілікті күшейтетін және STEM-білім беру талаптарына жауап беретін тиімді педагогикалық модель екенін дәлелдейді. Интеграцияланған тапсырмалар, сандық лабораториялық жұмыстар және жобалық әрекеттер студенттердің теориялық білімін практикалық жағдайларда қолдануына кеңінен ықпал етеді.

Түйін сөздер: пәнаралық байланыс, интеграциялық тәсілдер, құзыреттілік, инновациялар

© Г.М. Аутова[1], Г.К. Нуртаева[2], Э.М. Зульбухарова[3], Г.С. Елеусизова[4], Р.Р. Жумабекова[5] 2025.

[1]Казахский автомобильно-дорожный институт имени Л.Б. Гончарова, Алматы, Казахстан;

[2]Казахский национальный медицинский университет имени С.Д. Асфендиярова, Алматы, Казахстан;

[3]Казахский национальный университет имени аль-Фараби, Алматы, Казахстан;

[4]Казахский университет международных отношений и мировых языков им. Абылай хана, Алматы, Казахстан;

[5]НУО «Казахстанско-Российский медицинский университет», Алматы, Казахстан.

E-mail: zulbuharova@gmail.com

ТЕОРЕТИЧЕСКИЕ ОСНОВЫ МЕЖДИСЦИПЛИНАРНОЙ ИНТЕГРАЦИИ ФИЗИКИ И ИНФОРМАТИКИ

Аутова Гульмира — магистр физики, старший преподаватель автомобильно-дорожного факультета, Казахский автомобильно-дорожный институт имени Л.Б. Гончарова, Алматы, Казахстан,
E-mail: g.autova77@mail.ru, ORCID ID: https://orcid.org/0009-0003-8412-9134;
Нуртаева Галия — PhD, ассоциированный профессор кафедры нормальной физиологии с курсом биофизики, Казахский национальный медицинский университет имени С.Д. Асфендиярова, Алматы, Казахстан,
E-mail: nurtayevagalyia2017@gmail.com, https://orcid.org/0000-0003-0033-9744;
Зульбухарова Эльмира — ст. преподаватель кафедры теплофизики и технической физики, физико-технический факультет, Казахский национальный университет имени аль-Фараби, Алматы, Казахстан,
E-mail: zulbuharova68@gmail.com, ORCID ID: 0009-0006-6450-3602;
Елеусизова Гулжан — магистр, старший преподаватель кафедры Лингвистики и информационно-коммуникационных технологий АО «Казахский университет международных отношений и мировых языков им. Абылай хана», Алматы, Казахстан,
E-mail: gulzhansabitovna@gmail.com;
Жумабекова Рамиля — магистр, старший преподаватель кафедры информационно-коммуникационных технологий НУО «Казахстанско-Российский медицинский университет», Алматы, Казахстан,
E-mail: ramilia_85@mail.ru, ORCID: 0009-0001-6516-7878;

Аннотация. В статье рассматриваются теоретические и практические основы обучения на основе междисциплинарной интеграции физики и информатики. В современной системе образования такой интеграционный подход направлен на повышение познавательной активности студентов, развитие исследовательских навыков, ИКТ-компетенций и культуры работы с данными. В качестве примера интегрированного обучения физике и информатике в статье приведен лабораторный опыт по определению вязкости жидкости методом Стокса, который традиционно изучается на занятиях по физике. Исследование основано на лабораторной работе по вычислению

вязкости жидкости с использованием формулы Стокса; полученные экспериментальные данные обрабатываются в электронных таблицах Google Sheets, анализируются в цифровом формате и визуализируются с помощью графиков. При изучении физических свойств жидкостей определение их вязкости занимает важное место, поскольку данный параметр характеризует внутреннее трение жидкости и влияет на характер течения материала и гидродинамические процессы. Применяемый цифровой подход позволяет не только повысить точность измерений, но и структурировать информацию, пояснять ход эксперимента и оформлять результаты в научном стиле. Педагогический эксперимент был проведен в нескольких вузах города Алматы, всего в нем приняли участие 112 студентов. Результаты анкетирования показали недостаточность связи между физикой и информатикой, а также высокий интерес к интегрированным занятиям и цифровым образовательным ресурсам. Установлено, что часть обучающихся нуждается в дополнительной поддержке при использовании электронных таблиц, формул, графиков и облачных сервисов. Полученные результаты подтверждают, что интегрированное обучение физике и информатике является эффективной педагогической моделью, повышающей качество образования, усиливающей профессиональную компетентность и соответствующей требованиям STEM-образования. Интегрированные задания, цифровые лабораторные работы и проектная деятельность способствуют более широкому применению теоретических знаний в практических ситуациях.

Ключевые слова: междисциплинарные коммуникации, интеграционные подходы, компетентность, инновации

Кіріспе. Қазіргі заманда білім беру жүйесінде пәнаралық байланыс – оқу үрдісінің ажырамас бөлігіне айналған маңызды әдіснамалық тәсілдердің бірі. Білім беру жүйесінде пәнаралық байланыс студенттердің жан-жақты дамуын қамтамасыз етуде маңызды рөл атқарады. Ол түрлі пәндер арасындағы ортақ тақырыптар мен идеяларды біріктіру арқылы студенттердің танымдық қабілетін арттырып, шығармашылық ойлауын дамытуға мүмкіндік береді. Пәнаралық байланыс студенттерді түрлі саладағы білімдерін үйлестіре отырып, кешенді мәселелерді шешуге дайындайды және олардың өмір бойы білім алуға деген құлшынысын оятады. Мысалы әртүрлі пәндер арасындағы байланыс, мысалға, Жаратылыстану және математика: Графиктер құру арқылы экологиялық мәліметтерді талдау, тіл мен әдебиет және тарих: Тарихи оқиғаларды көркем шығармалар арқылы зерттеу, өнер мен информатика: сандық өнер құралдарын пайдаланып, шығармашылық жобалар жасау, сондай-ақ физика мен информатиканың пәнаралық байланысы қарастырылады. Мамандар дайындау барысында физика мен информатиканың өзара байланысын жүзеге асыру олардың кәсіби құзыреттілігін арттырып, ғылыми-техникалық ойлау қабілетін дамытады (*Айтбаева, 2016*).

Информатика және физиканы интеграциялау арқылы оқыту білім алушылардың ғылыми ұғымдарды сандық модельдермен түсініп, ақпараттық технологиялар арқылы практикалық дағдыларын жетілдіруге көмегін көрсетеді. Осы оқыту әдісі білім алушылар үшін заманауи инновациялық білім беру ортасын қалыптастыруға ықпал етеді

Мақалада қарастырылатын *мәселенің өзектілігі* – физика мен информатика пәндерін ықпалдастыра отырып оқыту моделін оқу бағдарламаларына енгізудің әдістемелік тиімділігін айқындау. Мақалада физика мен информатиканың пәнаралық байланысын жүзеге асыру әдістемесі талданады, оның тиімділігі көрсетіледі және білім беру саласында қолданудың мүмкіндіктері зерттеледі.

Зерттеу нысаны: Физика және информатика пәндерінің интеграциялық негіздері, оқыту әдістері және олардың тиімділігі (Байжұманова, 2020).

Зерттеу материалдары мен әдістері. Зерттеуде педагогиканың теориялық және эмпирикалық әдістері қолданылды. Зерттеу жұмысының материалы ретінде физика және информатика пәндері бойынша оқу бағдарламалары мен силлабустар, пәнаралық байланыстарды көрсететін интеграцияланған сабақ жоспарлары, зертханалық жұмыстар, сандық білім беру ресурстары алынды. Сонымен қатар, нақты кіріктірілген сабақ үлгілеріне шолу жасалынды. Бұл сабақтарды өту барысында жоғары оқу орындарында білім алушылардың пәнге деген қызығушылығы, танымдық белсенділігі және зерттеушілік қабілеттерінің жақсыға өзгергені бақыланды.

Физика мен информатиканы интеграциялау бірнеше маңызды мәселелерді шешеді. Біріншіден, бұл әдіс білім алушылардың пәнге деген қызығушылығын арттырады. Екіншіден, ол кәсіби құзыреттілік деңгейін көтеріп, қазіргі заманғы білім беру талаптарына жауап береді. Дегенмен, қиындықтар да бар. Мысалы, кейбір білім алушылар Google Sheets электронды кесте бағдарламасымен деректерді талдау, формулалар қолдануының төмендігі немесе бұлттық қызметтерді пайдалану барысында қиындықтарға тап болуы мүмкін. Бұл мәселелерді шешу үшін білім алушыларға арналған қосымша тренингтер мен онлайн ресурстарды ұсыну тиімді.

Мақалада физика мен информатика пәндерін кіріктіре оқытудың мысалы ретінде физика сабағында қарастыралатын Стокс әдісі арқылы сұйықтықтың тұтқырлығын анықтау тәжірибесі алынған.

Сұйықтардың физикалық қасиеттерін зерттеу кезінде олардың тұтқырлығын анықтау маңызды орын алады. Бұл параметр сұйықтықтың ішкі үйкелісін сипаттайды және материалдың ағу сипатына, гидродинамикалық процестерге әсер етеді. Осы жұмыста шардың сұйықтық ішінде еркін құлап бара жатқан кезіндегі қозғалыс заңдылықтарын Стокс формуласын пайдалана отырып зерттеу ұсынылады. Алынған нәтижелер электронды кесте платформасы (Google Sheets) арқылы өңделеді, талданады және визуализацияланады (Құсайынов, 2018).

Сұйықтың тұтқырлығын Стокс әдісімен анықтау, осы лабораториялық

жұмысты электронды кесте форматында (Google Sheets) бұл электронды кестелерді құруға, өңдеуге және талдауға арналған онлайн-құрал. Ол кестелермен жұмыс істеп, есептеулерді автоматтандыруға, графиктер мен диаграммаларды құруға, формулалар қолдануға, деректерді визуализациялауға мүмкіндік береді. Google Sheets деректерді нақты уақытта басқалармен бөлісуге және бірлесіп өңдеуге арналған, бұл оны өте қолайлы ететін тиімді әдіс-тәсіл. Бұл әдіс педагогикалық тұрғыдан өте тиімді, себебі ол теорияны практикамен байланыстырады және есептеу дағдыларын дамытады. Бұл әдісті қолдану физика мен информатика арасындағы пәнаралық байланысты тиімді жүзеге асыруға мүмкіндік береді. Сондай-ақ диаграмманы қосу арқылы тұтқырлықтың сұйықтық ішінде қозғалатын шардың радиусына, немесе оның сұйықтықтағы түсу жылдамдығына тәуелділігі көрсетіледі.

Сұйықтық тұтқырлығы – физикалық параметр, ол сұйықтағы ішкі үйкелісті сипаттайды және материалдың ағу қасиеттеріне тікелей әсер етеді. Бұл қасиетті эксперименттік тұрғыда анықтаудың кең тараған әдістерінің бірі – Стокс әдісі. Бұл әдіс шар формасындағы дененің тұтқыр сұйықтықта еркін қозғалу заңдылығын пайдалануға негізделеді (Серікбаева, 2021, Жунисова, 2019).

Жұмыстың мақсаты: Сұйық ішінде құлап бара жатқан шарға әсер ететін кедергі күшін есептеуге арналған Стокс формуласын пайдаланып, эксперимент жүзінде сұйықтың тұтқырлығын анықтау.

Қажетті материалдар мен жабдықтар: ішіне тұтқыр сұйық құйылған цилиндрлік ыдыс, жіпке ілінген металл тор немесе магнит, микрометр, секундометр, термометр.

Қысқаша теориялық кіріспе.

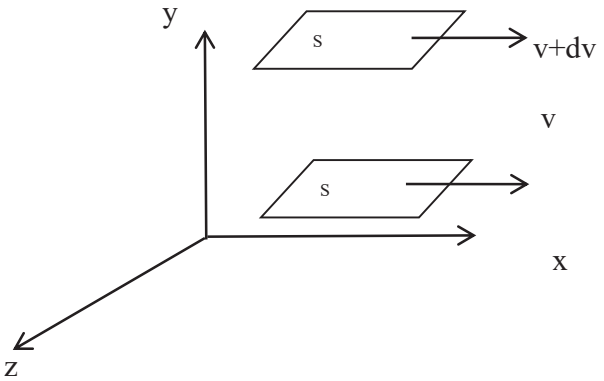
Қозғалыстағы денеге қоршаған орта тарапынан кедергі күш әсер етеді. Дене бетіне тиіп тұрған сұйықтың қабаты денеге жабысып, денемен бірге қозғалады. Молекулалық өзара әсерлесу арқасында осы қабатқа көршілес жатқан сұйықтың келесі қабаттары да қозғалысқа түседі. Бірақ, қозғалысқа түсетін көршілес қабаттар денеден неғұрлым қашық болған сайын, олардың жылдамдықтары да соғұрлым дене жылдамдығынан кіші бола береді. Сондықтан, осы сұйық қабаттары арасында үйкеліс пайда болады. Жылдамдықтары әртүрлі параллель қозғалатын сұйықтың екі қабаты арасындағы $F_{үйк.}$ Үйкеліс күші Ньютонның тұтқырлық үйкеліс заңымен анықталады:

$$F_{OD} = \eta \left| \frac{dv(y)}{dy} \right| S, \qquad (1)$$

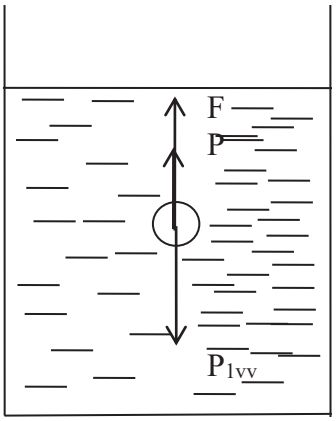
мұндағы S - жанасып жатқан екі қабаттың ауданы;

$\left| \frac{dv(y)}{dy} \right|$ - екі қабат арасындағы жылдамдық градиентінің модулі (1- сурет); η - сұйықтың тұтқырлығы. 1 формуладан η тұтқырлықтың мәні шығады. η - сан жағынан жылдамдық градиенті бірге тең кезде

параллель қозғалыстағы бірлік ауданға әсер ететін үйкеліс күшіне тең шама. Халықаралық жүйеде тұтқырлықтың өлшем бірлігі - паскаль·секунд (Па·с) Тұтқырлық сұйықтың тегіне қарай әртүрлі болады және белгілі бір сұйық үшін температура мен сұйық ішіндегі басқа заттардың концентрациясына тәуелді болады. Енді сұйыққа батып бара жатқан шарға әсер ететін күштерді қарастырайық. (2 сурет).



Сурет -1 Екі қабат арасындағы жылдамдық градиентінің модулі



Сурет - 2 Сұйыққа батып бара жатқан шарға әсер ететін күштер

Сұйық ішінде батып бара жатқан шарға үш түрлі күш әсер етеді:

P_1-ауырлық күші

P- кері итеруші күш

$F_{үйк}$-ортаның кедергі күші.

Шардың радиусын - r, жылдамдығын - v, тығыздығын - p_1, сұйық тығыздығын - p, тұтқырлығын- η және ауырлық күшінің үдеуін -g деп белгілейік. Сонда шардың ауырлық күші

$$\overline{P}_1 = \frac{4}{3}\pi r^3 \rho_1 \overline{g}, \qquad (2)$$

ал кері итеруші күш Архимед заңы бойынша

$$\overline{P} = -\frac{4}{3}\pi r^3 \rho_1 \overline{g}, \tag{3}$$

болады. Егер шар шексіз сұйық ішінде өте аз жылдамдықпен қозғалыс жасаса (құйынсыз қозғалыс), онда Стокс заңына сәйкес, орта тарапынан шарға әсер ететін кедергі күш

$$\overline{F} = -6\pi\eta r\overline{v}, \tag{4}$$

формуласымен анықталады. Ньютонның екінші заңын қолданып, $\left(md\overline{v}/dt = \sum \overline{F}\right)$ қозғалыс теңдеуін былай жазамыз:

$$m\frac{dv}{dt} = P_1 - P - F \tag{5}$$

(1) - (4) формулалардан күштердің және шардың массасының

$$m = \frac{4}{3}\pi r^3 \rho_1 \tag{6}$$

мәндерін қойып (5) теңдеуін түрлендіреміз:

$$\frac{dv}{dr} = \frac{9\eta}{2r^2\rho_1}\left[\frac{2(\rho_1 - \rho)gr^2}{9\eta} - v\right] \tag{7}$$

Айнымалы шамалары теңдіктің екі жағына бөліп былай жазайық:

$$\int_{v_6}^{v} \frac{dv}{\left[\frac{2(\rho_1 - \rho)gr^2}{9\eta}\right]} = \frac{9\eta}{2r^2\rho_1}\int_0^1 dt \tag{8}$$

t=0 мезетте шариктің сұйық бетіне тасталғандағы бастапқы жылдамдығы $v=v_6$ болса (8) теңдеуді интегралдап, мынаны аламыз:

$$v(t) = v_0 - (v_0 - v_i)e^{\frac{t}{\tau}} \tag{9}$$

мұндағы

$$\tau = 2\rho_1 r^2/9\eta \tag{10}$$

шар қозғалысының релаксациялық периоды деп аталады.

Егер $t >> \tau$ болса, шардың жылдамдығы тұтқыр сұйық ішінде өзінің максимал мәніне

$$v_0 = \frac{2(\rho_1 - p)gr^2}{9\eta} \tag{11}$$

ұмтылады.

Қозғалыс басталғаннан

$$t_1 = 4{,}6\tau \approx \frac{\rho_1 r^2}{\eta} \tag{12}$$

уақыт өткенде шардың жылдамдығы $v_1 = 0{,}99 * v_0$ болады. Осындай уақыт аралығында шар $x_1 \tau\ 3{,}6 v_0 \tau$ жол жүріп үлгереді.

Егер $t > t_1$ болса, шар сұйық ішінде тұрақталған v_0 жылдамдықпен қозғалады.

Шар шексіз сұйық ішінде емес, өлшемі шектеулі ыдыс ішіндегі сұйықта қозғалса, онда ыдыс қабырғаларының шар қозғалысына әсерін ескеру керек. Егер цилиндр ыдыстың радиусы R болса, шарға әсер ететін кедергі күші (4)-тің орнына былай жазылады.

$$F = 6\pi\eta v\left(1 + 2{,}4\frac{r}{R}\right), \tag{13}$$

соның нәтижесінде

$$v_0 = \frac{2(\rho_1 - \rho)gr^2}{9\eta\left(1 + 2{,}4\frac{r}{R}\right)} \tag{14}$$

Осы формуланы пайдаланып, шардың тұрақталған жылдамдығын өлшеу арқылы тұтқыр сұйықтың тұтқырлығын анықтауға болады:

$$\eta = \frac{2(\rho_1 - \rho)gr^2}{9 * v_0\left(1 + 2{,}4\frac{r}{R}\right)} \tag{15}$$

Тұтқыр сұйық ішінде қозғалатын шарға әсер ететін кедергі күшін анықтайтын (4) Стокс формуласы Рейнольде санының

$$\mathrm{Re} = \frac{2vr\rho}{\eta} \tag{16}$$

мәні $\mathrm{Re} \div 0{,}8$ болғанда ғана дұрыс болады. Егер Re>0,8 болса, шардың кедергі күші Стокс заңына бағынбайды.

Қондырғының құрылысы және жұмыстың орындалу әдістемесі.

Қондырғы вертикаль орнатылған, ішіне глицирин, кастор майы құйылған екі мөлдір ыдыстан тұрады. Цилиндрдің сыртына шардың жүрген жолын білу үшін әрбір 10 см сайын сызықтар салынған. Өлшеу кезінде диаметрлері әртүрлі (1 мм ÷3мм маңында) болат шарлар қолданылады.

Жұмыстың орындалу тәртібі

Микрометрдің көмегімен кішкене шарлардың (5÷7 шар) d диаметрін айналдыра отырып, 3 рет өлшейміз.

Сұйық бетінен қандай тереңдіктен бастап, шардың қозғалысы бірқалыпты болатынын анықтау керек. Ол үшін шардың 1-інші мен 2-нші сызықтар және 3-інші мен 4-інші сызықтар арасын қанша уақытта өткенін өлшеп, сол бойынша екі жылдамдықты тауып, салыстыру керек. Егер осы екі аралықтағы жылдамдықтар тең болса, онда келесі өлшеулерде қозғалыстың бастапқы нүктесі ретінде 1-інші сызықты қабылдауға болады.

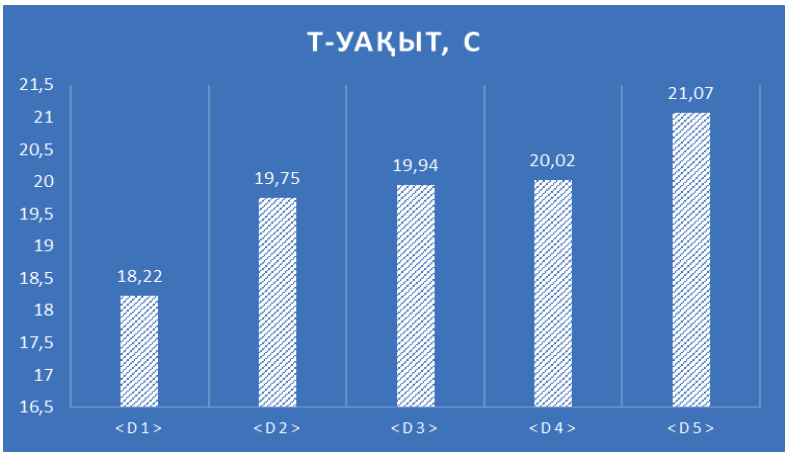
Шардың 60 см (немесе оқытушының нұсқауымен басқа мән болуы мүмкін) жолды бірқалыпты қозғалыспен жүріп өткен уақытын секундометрмен өлшеңіз. Осыны әр сұйық (глицерин және кастор майы) үшін 5÷7 шармен қайталаңыз. Өлшеу нәтижелерін төмендегідей кестеге түсіріңіз (4.1. және 4.3-пункттерін 1 және 2-кестеге енгізіңіз). Өлшеу жұмыстарын аяқтаған соң ыдыс ішіндегі шарларды тордың немесе магниттің көмегімен шығарып алып, қорабына салып қойыңыз.

1- Кесте. Глицерин үшін.Цилиндр радиусы R=мм

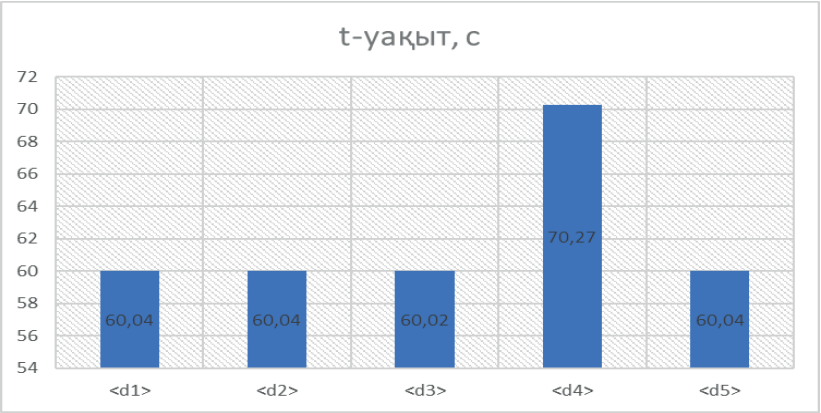
Шардың диаметрі, мм	$<d_1>$ 1.33	$<d_2>$ 1.31	$<d_3>$ 1.30	$<d_4>$ 1.31	$<d_5>$ 1.30
t-уақыт, с	18.22	19.75	19.94	20.02	21.07

2- Кесте. Кастор үшін.Цилиндр радиусы R=мм

Шардың диаметрі, мм	$<d_1>$	$<d_2>$	$<d_3>$	$<d_4>$	$<d_5>$
t-уақыт, с	60.04	60.04	60.02	70.27	60.04



1- Диаграмма. Глицерин үшін. Цилиндр радиусы R=мм



2- Диаграмма. Кастор үшін.Цилиндр радиусы R=мм

Зерттеу нәтижелері мен талқылау.

Педагогикалық экспериментті ұйымдастыру және оның нәтижелері

Физика мен информатика пәндерінің арасындағы байланысты жүзеге асыру мақсатында білім алушылардың электронды платформалар мен ғылыми визуализацияны қаншалықты деңгейде меңгергенін анықтау жәнеде білімдерінің төмен болған жағдайында оған не себеп болғанын анықтау үшін 2025-жылдың қаңтар айында тәжірибелік-эксперименттік зерттеу жұмыстары ұйымдастырылды. Эксперименттік зерттеу жұмыстарын ұйымдастыру үшін негізінен Алматы қаласындағы Қазақ Ресей медицина университеті, сондай-ақ мақала авторлары қатарындағы әріптестеріміз сабақ беретін кейбір техникалық жоғары оқу орындары алынды.

Зерттеу материалдары мен нәтижелерін талдау жүргізу барысында келесі әдістер қолданылды:

1) *Теориялық зерттеудің жалпы ғылыми әдістері*: зерттеу тақырыбы бойынша психологиялық, педагогикалық және ғылыми-әдістемелік, арнайы әдебиеттерді салыстырмалы талдау;

2) *Әлеуметтік зерттеу әдістері*: педагогтардан сұхбат алу және сауалнама жүргізу, салыстыру, жіктеу, сауалнама нәтижесін талдау;

3) *Эмпирикалық зерттеу әдістері:* педагогикалық эксперимент жүргізу, статистикалық зерттеу әдістерін пайдалану, нәтижелерін өңдеу және талдау.

Тәжірибелік-эксперименттік жұмысты іске асыру кезінде медицина, стоматология және фармация факультеттерінің бірінші курс білім алушылары арасында анонимді сауалнама алынды. Сауалнамаға жалпы 112 респондент қатысты.

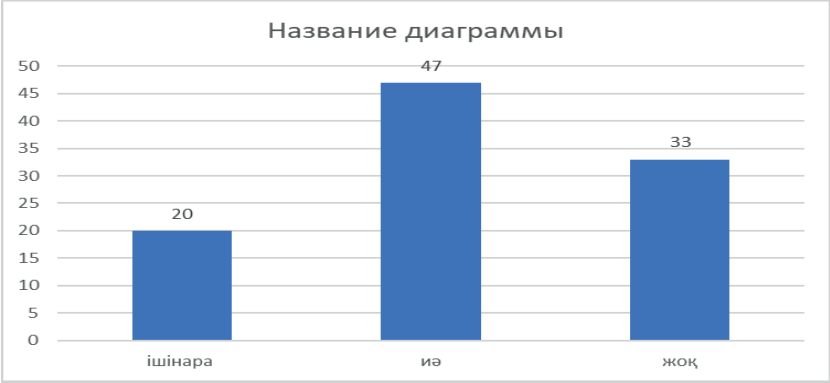
Жүргізілген зерттеулер нәтижесінде келесідей қорытындылар жасалды:

Мектептегі физика пәнінен меңгерген білімдерінің ақпараттық коммуникациялық технологиялар пәнін игеруге жеткіліктігіне байланысты сұрақта білім алушылардың ішінара ғана (25,7%) жеткілікті деп жауап берді, ал 42,5% жеткілікті десе 31,8% мүлде жеткіліксіз деп жауап берді.

Бұдан шығатын қорытынды басым көпшілік білім алушылардың физика мен информатика пәндерінің арасындағы байланыстың жеткіліксіздігі анықталды (3- сурет).

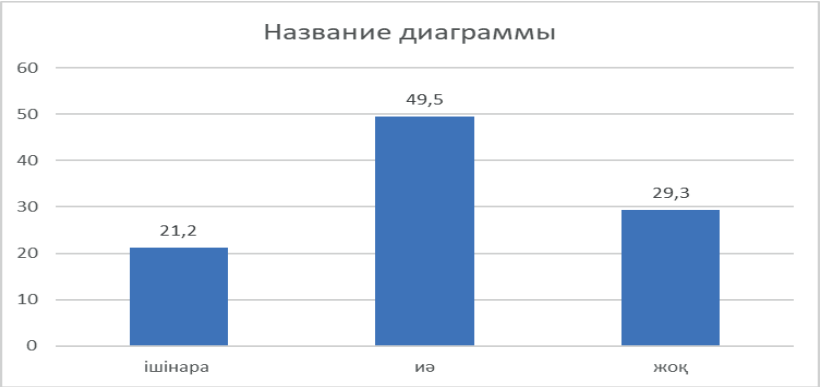
Сурет 3 - Зерттеу нәтижесі

Физика мен информатика пәндерінің арасындағы байланыстың оқыту қажеттілігін анықтауда басым көпшілік білім алушылар 47% иә қажет десе, 33% ішінара қажет, ал 20% қажет емес деп жауап берген (4- сурет).



Сурет – 4 Зерттеу нәтижесі

Білімдерінің не себепті төмен екенін анықтауға байланысты сұрақта 49,5% респондент ҰБТ-да физиканың бейіндік пән болмауы десе 29,3% пәнге деген қызығушылықтың төмендігі, ал 21,2% оқу экспериментінің болмауы деп жауап берген. Осыдан білімдерінің төмен деңгейде болуына әсер ететін негізгі себептер анықталды (5- сурет).



Сурет 5 - Зерттеу нәтижесі

Осыдан шығатын қорытынды білім алушылардың физика мен информатика пәндерінің арасындағы байланысты жүзеге асыру мақсатында білім алушылардың электронды платформалар мен ғылыми визуализацияны қажеттілігін түсінетіндігінде жәнеде осы пәнді оқытудың жаңа әдістерін қарастырудың маңыздылығын айқындайды. қажеттілігін түсінетіндігінде жәнеде осы пәнді оқытудың жаңа әдістерін қарастырудың маңыздылығын айқындайды.

Қорытынды. Зерттеу барысында Стокс формуласы сұйықтың тұтқырлығын есептеуде тиімді әдіс ретінде ұсынылып, оның қарапайымдылығы мен қолжетімділігі көрсетілді. Зертханалық жұмыс арқылы алынған мәліметтер нақты әрі түсінікті түрде өңделіп, білім алушылардың тәжірибелік дағдыларын қалыптастыруда тиімді құрал ретінде қолданылды. Физикалық құбылыстарды зерттеу нәтижелерін талдауда электронды кестелерді пайдалану — зерттеу үдерісін заманауи талаптарға сәйкестендіре отырып, деректердің нақтылығын, қолжетімділігін және визуалды сапасын жақсартатына көз жеткізілді.

Google Sheets сияқты цифрлық құралдарды қолдану:өлшеу дәлдігін арттыруға;нәтижелерді графикалық түрде көрсету арқылы деректерді түсінуді жақсартуға;эксперименттік нәтижелерді оңтайлы бөлісуге мүмкіндік береді. Сонымен қатар,Google Sheets сияқты электронды платформалар нәтижелерді өңдеуді жеңілдетеді және ғылыми визуализацияны жаңа деңгейге көтереді Электронды кестелерді пайдалану, зерттеу үдерісін заманауи талаптарға сәйкес икемдей отырып, нәтижелердің дұрыстығы мен қолжетімділігін арттырады.Мұндай интеграцияланған әдістер физика, химия және инженерия салаларында сұйықтық қасиеттерін зерттеу үшін ғылыми-зерттеушілік жұмыстарды жетілдіруде маңызды.

Осының нәтижесінде, зерттеу үдерісі тек техникалық тұрғыдан емес, сонымен қатар ақпараттық-коммуникациялық технологиялармен (АКТ) тиімді ықпалдасуы арқылы жаңғыртылды. Бұл тәсіл пәнаралық байланысты нығайтып, оқушылардың ғылыми ойлауын, талдау қабілетін және зерттеушілік біліктілігін арттыруға ықпал етеді.

Әдебиеттер

Айтбаева А.К. (2016) Интеграциялық оқытудың теориясы мен практикасы. А.К. Айтбаева. — Алматы: Ұлағат. — 200 б.

Байжұманова Г.А. (2020) Физика пәнін оқытуда цифрлық технологияларды пайдалану жолдары. Г.А. Байжұманова. Білім-Образование. — №2. — Б. 45–48.

Құсайынов А.Ф. (2018) Жаратылыстану пәндерін оқытудағы пәнаралық байланыс А.Ф. Құсайынов. — Нұр-Сұлтан: Фолиант, – 168 б.

Оспанова Г.Ж. (2019) Ақпараттық-коммуникациялық технологияларды білім беру процесінде қолдану. Г.Ж. Оспанова. — Алматы: Білім — 132 б.

Серікбаева А. (2021) Информатиканы оқытуда пәнаралық байланысты жүзеге асыру ерекшеліктері. Ғылым және өмір. — №3(61). — Б. 37–40.

Жунисова М.К. (2019) Информатика және физика сабақтарын кіріктіре оқытудың педагогикалық негіздері. М.К. Жунисова. Қазақ университеті хабаршысы. — №1. — Б. 52–56.

Маханова Г.Ф. (2022) Цифрлық сауаттылықты қалыптастырудағы пәндердің кіріктірілуі Г.Ф. Маханова. Ұлағат. — №4. — Б. 29–32.

Novikova N.P. (2017) Integration of Academic Disciplines in the Context of Digital Transformation of Education. Moscow: Prosveshchenie. — 176 p.

Safonova E.V. (2018) Methods of Using ICT in Physics Lessons. Saint Petersburg: Piter. — 192 p.

Bespalko V.P. (2015) Pedagogy and Progressive Teaching Technologies. Moscow: Akademiia. — 240 p.

Soloviev M.N. (2020) Interdisciplinary Connections in Education: Theory and Practice. Novosibirsk: NSU. — 205 p.

Тұрдыбаева А.К. (2021) Физика және информатика пәндерін оқытуда сандық ресурстарды қолдану тиімділігі. А.К. Тұрдыбаева. Педагогика және психология. — №4. — Б. 41–46.

Иманғалиева Л.С. (2020) STEM-білім берудегі пәнаралық интеграцияның заманауи тәсілдері. Л.С. Иманғалиева. – Алматы: Қазақ университеті. — 150 б.

Кравцова О.П. (2019) Информационные технологии в естественно-научном образовании О.П. Кравцова. Образовательные технологии и общество. — Т. 22. – №2. — С. 117–124.

Әбдірасылова Г.Ж. (2022) Жаратылыстану пәндерінде АКТ қолданудың педагогикалық шарттары. Г.Ж. Әбдірасылова. Қазақстан мектебі. — №1. — Б. 55–59.

Юрин А.В. (2018) Интеграция электронных образовательных ресурсов в обучении физике. А.В. Юрин. – Санкт-Петербург: Лань. — 184 с.

References

Aitbaeva A.K. (2016) Integratsiialyq oqytudyń teoriiiasy men praktikas [Theory and Practice of Integrated Teaching]. A.K. Aitbaeva. — Almaty: Ulgagat, — 200 b. (in Kaz)

Baizhumanova G.A. (2020) Fizika pánin oqytuda tsifrlyq tekhnologiialardy paidalanu zholdary [Ways of Using Digital Technologies in Teaching Physics]. *Bilim – Obrazovanie.* – №2. — P. 45–48. (in Kaz)

Qusaiynov A.F. (2018) Jaratylystanu pánderin oqytudagy pánaralyq bailanys [Interdisciplinary Connections in Teaching Natural Science Subjects]. Foliant. — P. 168 (in Kaz)

Ospanova G.Zh. (2019) Aqparattyq-kommunikatsiialyq tekhnologiialardy bilim beru protsesinde qoldanu [Use of Information and Communication Technologies in the Educational Process.] Almaty: Bilim, — P. 132 (in Kaz)

Serikbaeva A. (2021) Informatikany oqytuda pánaralyq bailanyshty zhuzege asyru erekshelikteri. Gylym zhane ómir. – №3(61). — B. 37–40. (in Kaz)

Zhunisova M.K. (2019) Informatika zhane fizika sabaqtaryn kiriktire oqytudyń pedagogikalyq negizderi [Features of Implementing Interdisciplinary Integration in Teaching Informatics.] Qazaq universiteti khabarlary. — №1. — P. 52–56. (in Kaz)

Makhanova G.F. (2022) Tsifrlyq sauattylyqty qalyptastyrudagy pánderdiń kiriktirilui [Integration of Subjects in Developing Digital Literacy] Ulagat. №4. — P. 29–32. (in Kaz)

Novikova N.P. (2017) Integration of Academic Disciplines in the Context of Digital Transformation of Education Moscow: Prosveshchenie, — 176 p. (in Eng)

Safonova E.V. (2018) Methods of Using ICT in Physics Lessons Saint Petersburg: Piter, — 192 p. (in Eng)

Bespalko V.P. (2015) Pedagogy and Progressive Teaching Technologies Moscow: Akademiia, — 240 p. (in Eng)

Soloviev M.N. (2020) Interdisciplinary Connections in Education: Theory and Practice. Novosibirsk: NSU. — 205 p. (in Eng)

Turdybaeva A.K. (2021) Fizika zhane informatika pánderin oqytuda sandyq resurstardy qoldanu tíimdiligi [Effectiveness of Using Digital Resources in Teaching Physics and Informatics.] Pedagogika zhane psikhologiia. – №4. — P. 41–46. (in Kaz)

Imangalieva L.S. (2020) STEM-bilim berudegi pánaralyq integratsiianyń zamanaýi tásilderi [Modern Approaches to Interdisciplinary Integration in STEM Education] Almaty: Qazaq universiteti, — P. 150 (in Kaz)

Kravtsova O.P. (2019) Informatsionnye tekhnologii v estestvenno-nauchnom obrazovanii [Information Technologies in Natural Science Education] Obrazovatel'nye tekhnologii i obshchestvo. — V. 22. — №2. — P. 117–124. (in Rus)

Abdirasylova G.Zh. (2022) Jaratylystanu pánderinde AKT qoldanudyn pedagogikalyq sharttary [Pedagogical Conditions for Using ICT in Natural Science Subjects] Qazaqstan mektebi. — №1. — P. 55–59. (in Kaz)

Iurin A.V. (2018) Integratsiia elektronnykh obrazovatel'nykh resursov v obuchenii fizike [Integration of Electronic Educational Resources in Teaching Physics] Saint Petersburg: Lan'. — 184 s. (in Rus)

ACADEMIC SCIENTIFIC JOURNAL OF COMPUTER SCIENCE
ISSN 1991-346X
Volume 4. Namber 356 (2025). 89–101

https://doi.org/10.32014/2025.2518-1726.385

МРНТИ 28.23.15
УДК 004.021:316.77

A.Zh. Akhmetova[1], M.A. Kantureyeva[1], A.A. Abisheva[2], A. Aubakirova[3] , A.A. Shekerbek[1], 2025.
[1]L.N. Gumilyov Eurasian National University, Astana, Kazakhstan;
[2]K. Kulazhanov Kazakh University of Technology and Business, Astana, Kazakhstan;
[3] Kokshetau University named after Sh. Ualikhanov, Kokshetau, Kazakhstan.
E-mail: akhmetova_azh@mail.ru

ANALYSIS OF THE SOCIAL NETWORK USER'S ENVIROMENT

Akhmetova Aidana — PhD, Senior Lecturer, Department of Information Systems, L.N. Gumilyov Eurasian National University, Faculty of Information Technology, Astana, Kazakhstan, E-mail: akhmetova_azh@mail.ru, ORCID: https://orcid.org/0000-0002-3876-1089;
Kantureyeva Mansiya — PhD, Senior Lecturer, Department of Information Systems, L.N. Gumilyov Eurasian National University, Faculty of Information Technology, Astana, Kazakhstan, E-mail: mansiya_arynbekovna@mail.ru, ORCID: https://orcid.org/0000-0001-5904-820X;
Abisheva Aigul — Master, Department of Information Technologies, Faculty of Technology, Kazakh University of Technology and Business named after K. Kulazhanov, Astana, Kazakhstan, E-mail: aigul_abisheva@mail.ru, ORCID: https://orcid.org/0009-0002-3613-6850;
Aubakirova Ainagul — PhD, teacher of the Department of Mathematics, Physics and Computer Science, Kokshetau University named after Sh. Ualikhanov, Kokshetau, Kazakhstan, E-mail:ainagul_aub@mail.ru, ORCID: https://orcid.org/0009-0002-8809-4023;
Shekerbek Ainur — PhD, Senior Lecturer, Department of Information Systems, L.N. Gumilyov Eurasian National University, Faculty of Information Technology, Astana, Kazakhstan, E-mail: shekerbek80@mail.ru, 800520401702@enu.kz ORCID: https://orcid.org/0000-0002-1088-4239.

Abstract. Today, social networks are an integral part of the daily lives of users worldwide. They are primarily used for communication, information exchange, and shaping public opinion. The volume of data in social networks grows every day, which attracts significant interest from researchers, as it allows for the study of user behavior and social processes. One of the important areas of research is the methods of influencing user behavior in social networks. This article is dedicated to methods of influencing users' opinions, specifically focusing on the theory of social influence. The article presents quantitative indicators, relationships, and sets calculated based on social network data. Modifications to Latané's dynamic social influence theory are proposed, taking into account the characteristics of social networks. During

the study, formulas describing the magnitude of social pressure exerted on an individual were modified. External factors of influence that may also affect the formation of an individual's opinion are considered. The results of social network analysis can be used as a tool for shaping and manipulating public opinion in states. The article also describes a developed software system designed for extracting, processing, and analyzing data from social networks. The system includes modules for data collection, processing, and analysis, all implemented in Python. The data extraction module allows obtaining information primarily from the largest social networks — Twitter and VKontakte — using application programming interfaces (APIs). Finally, the results of the calculations are presented.

Keywords: analysis of social networks, Internet, vkontakte, data analysis, methods of analysis in social networks, Latane theory, social network

© А.Ж. Ахметова[1], М.А. Кантуреева[1], А.А. Абишева[2], А. Аубакирова[3], А.А. Шекербек[1], 2025.

[1] Л.Н. Гумилев атындағы Еуразия ұлттық университеті, Астана, Қазақстан;
[2] Қ. Құлажанов атындағы Қазақ технология және бизнес университеті, Астана, Қазақстан;
[3] Ш.Уәлихан атындағы Көкшетау университеті, Көкшетау, Қазақстан.
E-mail: akhmetova_azh@mail.ru

ӘЛЕУМЕТТІК ЖЕЛІ ҚОЛДАНУШЫЛАРЫНЫҢ ОРТАСЫН ТАЛДАУ

Ахметова Айдана — PhD, Л.Н. Гумилев атындағы Еуразия ұлттық университетінің ақпараттық жүйелер кафедрасының аға оқытушысы., ақпараттық технологиялар факультеті, Астана, Қазақстан,
E-mail: akhmetova_azh@mail.ru, ORCID: https://orcid.org/0000-0002-3876-1089;
Кантуреева Мансия — PhD, Л.Н. Гумилев атындағы Еуразия ұлттық университетінің ақпараттық жүйелер кафедрасының аға оқытушысы, ақпараттық технологиялар факультеті, Астана, Қазақстан,
E-mail: mansiya_arynbekovna@mail.ru, ORCID: https://orcid.org/0000-0001-5904-820X;
Абишева Айгүл — Қ. Құлажанов атындағы Қазақ технология және бизнес университетінің Технология факультетінің ақпараттық технологиялар кафедрасы магистрі, Астана, Қазақстан,
E-mail: aigul_abisheva@mail.ru, ORCID: https://orcid.org/0009-0002-3613-6850;
Әубәкірова Айнагүл — PhD, Ш.Уәлихан атындағы Көкшетау университетінің математика, физика және информатика кафедрасының оқытушысы, Көкшетау, Қазақстан,
E-mail:ainagul_aub@mail.ru, ORCID: https://orcid.org/0009-0002-8809-4023;
Шекербек Айнұр — PhD докторы, Л.Н.Гумилев атындағы Еуразия ұлттық университетінің ақпараттық жүйелер кафедрасының аға оқытушысы, ақпараттық технологиялар факультеті, Астана, Қазақстан,
E-mail: shekerbek80@mail.ru, 800520401702@enu.kz ORCID: https://orcid.org/0000-0002-1088-4239.

Аннотация. Бүгінгі таңда әлеуметтік медиа бүкіл әлемдегі пайдаланушылардың күнделікті өмірінің бір бөлігі болып табылады. Олар негізінен қарым-қатынаста, ақпарат алмасуда қолданылады және қоғамдық

пікірді қалыптастырады. Әлеуметтік желілерде деректер көлемі күн сайын артып келеді және бұл зерттеушілер үшін үлкен қызығушылық тудырады, өйткені ол пайдаланушылардың мінез-құлқын, әлеуметтік процестерді зерттеуге мүмкіндік береді. Олардың ішіндегі ең маңыздыларының бірі – әлеуметтік желіні пайдаланушылардың мінез-құлқына әсер ету әдістері. Бұл мақала пайдаланушылардың пікірлеріне әсер ету әдістеріне арналған. Атап айтқанда, әлеуметтік әсер ету теориясы қарастырылады. Мақалада әлеуметтік желілерден алынған мәліметтер негізінде есептелетін сандық сипаттамалар, қатынастар және жиынтықтар келтірілген. Латаненің динамикалық әлеуметтік әсер ету теориясының әлеуметтік желілерге қатысты модификациялары ұсынылған. Жұмыс барысында әртүрлі жағдайлар үшін жеке адамға бағытталған әлеуметтік қысымның мөлшерін сипаттайтын формулалар өзгертілді. Жеке тұлғаның пікірін өзгертуге әсер ететін сыртқы әсер факторлары сипатталған. Әлеуметтік желілерді талдау нәтижелері мемлекеттерде қоғамдық пікірді қалыптастыру және қоғамдық пікірді манипуляциялау құралы ретінде пайдаланылуы мүмкін. Мақалада әлеуметтік желілерден ақпарат алуға, деректерді өңдеуге және талдауға мүмкіндік беретін әзірленген бағдарламалық жасақтама туралы айтылады. Әлеуметтік желілерден ақпаратты алу, өңдеу, деректерді талдау модульдерін қамтитын бағдарламалық кешен. Барлық модульдер Python тілінде жүзеге асырылады. Деректерді шығару модулі деректерді, ең алдымен, ең ірі әлеуметтік желілерден: Twitter және vkontakte-ден алуға мүмкіндік береді. Олардың әрқайсысына қол жеткізу үшін қолданбалы бағдарламалау интерфейсі (API) қолданылады. Соңында есептеу нәтижелері келтірілген.

Түйін сөздер: әлеуметтік желілерді талдау, ғаламтор, vkontakte , деректерді талдау, Латане теориясы, әлеуметтік желідегі талдау әдістері, Латане теориясы, әлеуметтік желі

А.Ж. Ахметова[1], М.А. Кантуреева[1], А.А. Абишева[2], А. Аубакирова[3], А.А. Шекербек[1], 2025.

[1]Евразийский национальный университет имени Л.Н. Гумилева, Астана, Казахстан;

[2]Казахский университет технологии и бизнеса имени К. Кулажанова, Астана, Казахстан;

[3] Кокшетауский университет имени Ш. Уалиханова, Кокшетау, Казахстан.

E-mail: akhmetova_azh@mail.ru

АНАЛИЗ ОКРУЖЕНИЯ ПОЛЗОВАТЕЛЕЙ СОЦИАЛЬНОЙ СЕТИ

Ахметова Айдана — PhD, старший преподаватель кафедры информационных систем Евразийского национального университета им. Л.Н. Гумилева, факультет информационных технологий, Астана, Казахстан,

E-mail: akhmetova_azh@mail.ru, ORCID: https://orcid.org/0000-0002-3876-1089;

Кантуреева Мансия — PhD, старший преподаватель кафедры информационных систем Евразийского национального университета им. Л.Н. Гумилева, факультет информационных технологий, Астана, Казахстан,
E-mail: mansiya_arynbekovna@mail.ru, ORCID: https://orcid.org/0000-0001-5904-820X;
Абишева Айгуль — магистр, кафедра «Информационные технологии» технологического факультета, Казахский университет технологии и бизнеса имени К. Кулажанова, Астана, Казахстан,
E-mail: aigul_abisheva@mail.ru, ORCID: https://orcid.org/0009-0002-3613-6850;
Аубакирова Айнагуль — PhD, преподаватель кафедры математики, физики и информатики, Кокшетауский университет имени Ш. Уалиханова, Кокшетау, Казахстан,
E-mail:ainagul_aub@mail.ru. ORCID: https://orcid.org/0009-0002-8809-4023;
Шекербек Айнур — PhD, старший преподаватель кафедры информационных систем Евразийского национального университета им. Л.Н. Гумилева, факультет информационных технологий, Астана, Казахстан,
E-mail: Shekerbek80@mail.ru, 800520401702@enu.kz, ORCID: https://orcid.org/0000-0002-1088-4239.

Аннотация. На сегодняшний день социальные сети являются частью повседневной жизни пользователей по всему миру. В основном используются в общении, обмене информацией а также формировании общественного мнения. В социальных сетях объем данных растет с каждым днем и это вызывает большой интерес для исследователей, так как позволяет изучать поведения пользователей, социальные процессы. Одним из важных среди которых являются методы влияния на поведения пользователей социальной сети. Данная статья посвящена методам влияния на мнения пользователей. А именно рассмотрена теория социального влияния. В статье представлены количественные показатели, взаимосвязи и множества, рассчитанные на основе данных социальных сетей. Предлагаются изменения в теории динамического социального влияния Латане с учетом особенностей социальных сетей. В процессе работы были изменены формулы, описывающие величину социального давления, которое оказывается на индивида. Рассмотрены внешние факторы влияния, которые также могут влиять на формирование мнения индивида. Результаты анализа социальных сетей могут быть использованы как инструмент формирования общественного мнения в государствах и манипулирования общественным мнением. В статье описан разработанный программный комплекс, предназначенный для извлечения, обработки и анализа данных из социальных сетей. Комплекс включает модули для сбора информации, её обработки и анализа, все они реализованы на языке Python. Модуль извлечения данных позволяет получать информацию прежде всего из крупнейших социальных сетей — Twitter и ВКонтакте, используя для этого интерфейсы прикладного программирования (API). В конце приведены результаты вычисления.

Ключевые слова: анализ данных, анализ социальных сетей, интернет, Вконтакте, методы анализа данных социальных сетей, теория Латане, социальные сети

Введение. Развитие интернет-технологий открыло возможности виртуального общения через социальные сети. Социальная сеть представляет собой интерактивную платформу с множеством пользователей, где контент создают пользователи сети. Такие платформы формируют автоматизированную социальную среду, позволяя группам людей с общими интересами взаимодействовать друг с другом. К ним также относятся тематические форумы, это особенно отраслевые, активно развивающиеся в последние годы. Число зарегистрированных пользователей одной социальной сети может превышать население целой страны (Charu, 2011).

Анализ социальных сетей используется в разведывательных и правоохранительных мероприятиях. Этот подход предоставляет аналитикам возможность картографировать скрытые или незаконные организации, включая шпионские сети, преступные сообщества и уличные банды. Агентство национальной безопасности (NSA) использует закрытые программы массового электронного мониторинга для получения данных, необходимых для анализа террористических структур и иных сетей, значимых для обеспечения национальной безопасности (Батура, 2012.)

Взгляды человека формируются не только под влиянием его собственных убеждений и анализа информации, но и с учётом мнений группы, друзей и окружающих. Авторы попытались адаптировать теорию динамического социального влияния Латане (Nowak et al., 1990; Wragg, 2006) для расчёта уровня влияния окружающих на мнение индивида. В статье предложены количественные параметры и множества с использованием данных социальных сетей.

Как показывают наблюдения психологов (Deutsch et al., 1955). В социальных сетях агенты нередко не располагают достаточной информацией для принятия решений или не способны самостоятельно её обработать, поэтому их решения могут формироваться на основе наблюдаемых действий или мнений других агентов (социальное влияние). Социальное влияние проявляется через два основных процесса. Первый — коммуникация: в ходе взаимодействия, обмена опытом и информацией, а также обсуждения различных вопросов с авторитетными для агента соседями формируются его представления, установки и мнения. Второй — сравнение: стремясь к социальной идентичности и одобрению, агент принимает те представления и модели поведения, которых ожидают от него другие участники сети. При этом он задаётся вопросом: «как бы поступил другой агент, являющийся для меня эталоном, в аналогичной ситуации?» Сопоставляя себя с ним, агент оценивает собственную адекватность и корректирует поведение. Процесс сравнения также может рассматриваться как поиск стратегического преимущества: ориентируясь на других агентов, занимающих схожие позиции в социальной структуре, индивид может внедрять или перенимать нововведения, повышающие его привлекательность в системе социальных

отношений. Нужно подчеркнуть, при коммуникации агенты могут выработать похожие взгляды, но их действия при этом могут различаться. Напротив, в процессе сравнения агент, как правило, перенимает наблюдаемое поведение. Ясно, что действия агента формируется не только его представлениями, но и от ограничивающих факторов, структурирующих его возможные решения. Следовательно, сходство представлений не гарантирует сходства поведения, так же как и различие взглядов не исключает одинаковых моделей поведения. В социальных сетях информация, идеи и влияние быстро передаются от одного пользователя к другому, что делает их важным средством коммуникации.

В целях проведения экспериментальной части исследования была реализована программа, позволяющая извлекать, систематизировать и анализировать пользовательские данные. Модуль системы позволяет собирать данные из ВКонтакте и может быть адаптирован для работы с другими социальными сетями через их API.

Материалы и методы. При изучении социальных сетей имеет смысл учитывать разнообразные числовые и нечисловые показатели, а также отношения и множества, которые связаны с пользователями и информацией в сети (Wasserman & Faust, 1994; Newman, 2010). Следует отметить, что все они являются конструктивными, то есть могут быть вычислены или построены с использованием соответствующих алгоритмов, при условии наличия программного обеспечения для извлечения необходимой информации из сети.

Например, p – сообщение («пост») социальной сети, u – пользователь сети, который состоит в сообществе, т.е. быть подписчиком сообщества. Пользователь способен создавать и распространять сообщения.

Числовые показатели: $Followers_Count(u)$ – число подписчиков, то есть людей, следящих за сообщениями данного пользователя; $Friends_Count(u)$ – Общее число друзей пользователя, выбранных им для включения в свой список контактов (добавленных им самостоятельно); $Timeline_Count(u)$ – Число «постов», созданных пользователем; $Retweets(p)$ – Число репостов сообщения (количество раз, когда данное сообщение было переслано другим пользователям); $Likes(P)$ – Количество положительных оценок («лайков») данного сообщения;

Нечисловые показатели: $Real_name(u)$ – Реальное имя пользователя при его наличии(качественная характеристика); $Bdate(u)$ – день, месяц и год рождения пользователя; $City(u)$ – Место проживания пользователя (город); $Hometown(u)$ – Признак соответствия города проживания пользователя его родному городу; $Career(u)$ – Организация, в которой работает пользователь, и его профессиональная деятельность; $University(u)$ – Информация о учебном заведении пользователя (университет, колледж и аналогичные учреждения), факультет; $School(u)$ – Информация о школе, в которой обучался пользователь; $Friends(u)$ информация о статусе семейного положения пользователя:

1 – не состоит в браке

2 – имеет партнёра
3 – помолвлен/помолвлена
4 – состоит в браке
5 – сложное семейное положение
6 – в активном поиске отношений
7 – эмоционально вовлечён/влюблён

При наличии информации о годе выпуска из учебного заведения, то этот показатель иногда рассматривается как нечисловая характеристика. В случаях, когда годы пользователей сравниваются исключительно на совпадение и фиксируются в виде 0 или 1, такой способ представления данных обычно обозначают как индикаторный.

Множества: $Followers(u)$ – Пользователи, подписанные на данного пользователя; $Friends(u)$ – Пользователи, добавленные данным пользователем в список друзей (т.е. друзья); $Mentions(u)$ – Идентификаторы пользователей, на которых ссылается данный пользователь в своих сообщения; $Hashtags(u)$ – Список хэштегов, встречающихся в публикациях и сообщениях данного пользователя; $Urls(u)$ – URL-адреса, упомянутые в сообщениях данного пользователя.

Количественные характеристики, соответствующие различным множествам: $Count_Mentions_u(v)$ – Числовое значение, отражающее, сколько раз один пользователь был упомянут u;

$Count_Hashtags_u(v)$ – Числовое значение, отражающее, сколько раз хэштег v встречается в публикациях u;

$Count_Urls_u(v)$ – Числовое начение, отражающее, сколько раз данная внешняя ссылка v встречается в публикациях u.

$Count_Retweets_u(u_1)$ – количество сообщений пересланных пользователем u, полученных от пользователя u_1;

$Count_CommonFriends_u(v)$ - Число общих контактов между двумя пользователями социальной сети u и u.

Увеличение числа подписчиков отражает возрастание популярности пользователя. Согласно теории социального влияния Латане (Latané, B., 1981, 1996), функцию можно интерпретировать как «силу влияния» данного пользователя.

Теория динамического социального влияния Латане и предложенные изменения этой модели

Далее будет рассмотрен подход к модификации теории динамического социального влияния Латане с целью её применения в контексте социальных сетей.

Три ключевых атрибута отношений по Латане: сила (статус и доверие), расстояние (физическое/психологическое), текущее количество источников. Уровень влияния на индивидуума по теории динамического социального влияния Латанеможет быть представлен формулой.

$$I_i = -S_i\beta - \sum_{j=1, j\neq i}^{N} \frac{S_j O_j O_i}{d_{i,j}^{\alpha}}$$

здесь:

I_i – Количество воздействия со стороны социальных агентов на индивидуума;

O_i – Значение мнения i-го индивидуума по заданному вопросу (±1), +1 — поддержка, –1 — несогласию с предложение;

S_i – Степень воздействия социальных факторов на индивидуума ($S_i \geq 0$) сила влияния;

β – Уровень сопротивления к внедрению изменений;

d_{ij} – Величина дистанции, разделяющей индивидов i и j;

α – Степень снижения влияния с увеличением расстояния;

N – Число агентов, участвующих во взаимодействиях.

Константы β и α обычно принимаются равными 2; увеличение D требует большего давления для изменения мнения, уменьшение — меньшего.

При увеличении значения α воздействие должно быть сильнее, чтобы преодолеть растущее расстояние между источником и получателем.

Параметр d_{ij}

Величина d_{ij} характеризует свойства пары индивидуумов и служит показателем эффективности передачи информации. При её установке принимаются во внимание возрастные, этнические, конфессиональные и иные различия. При вычислении d_{ij} формула может включать фактор физического расстояния между субъектами. Примером является физическое расстояние между населенными пунктами, в которых находятся индивиды. Как правило, предполагается, что лёгкость общения обратно пропорциональна квадрату расстояния (Wragg T., 2006). В случае компьютерных сетей возможны альтернативные подходы, в том числе игнорирующие физическое расстояние.

В целях моделирования социальных сетей предлагается адаптированная версия формулы Латане следующего вида

$$I_u = -\beta \cdot \sum_{i=1}^{N} Followers_Count(u_i) - \sum_{i=1}^{N} \sum_{\substack{j=2 \\ i>j}}^{N} \frac{Followers_Count\,(u_j) O_j O_i}{d^{\alpha}(u_i, u_j)}.$$

В этой формуле числом подписчиков пользователя определяется сила его влияния: рост числа подписчиков отражает возрастание популярности пользователя. $d(u_i, u_j)$ – расстояние от пользователя u_i до пользователя u_j, которое определяется посредством анализа различий в анкетных данных.

Для вычисления расстояния между пользователями применяется расстояние Хэмминга. Каждому пользователю приписан вектор категориальных характеристик, а расстояние Хэмминга представляет собой число позиций,

где эти характеристики различаются. Параметр O_i в исследовании (Latané, 1981, 1996) речь применялся к кампании по вакцинации против полиомиелита, где воздействие на общество осуществлялось через различные медиа, с наибольшим влиянием телевидения и радио в городских областях. Значение O принимается равным ±1 в зависимости от принадлежности i-го индивидуума к данному сообществу.

Для нашей модели роли СМИ могут выполнять такие параметры, как сила влияния хэштега и сила влияния ссылки.

Влияние внешних факторов, таких как СМИ, может быть включено в модель (Wragg T., 2006) путём добавления дополнительного члена в формулу Латане: $O_i O_M S_{Mi}$, где S_{Mi} – сила влияния внешних источников на *i*-й индивидуум, $S_{Mi} > 0$, O_M – мнение внешнего источника. При включении влияния средств массовой информации формула принимает окончательный вид:

$$I_i = -S_i\beta - O_i O_M S_{Mi} - \sum_{j=1, j\neq i}^{N} \frac{S_j O_j O_i}{d_{i,j}^{\alpha}}$$

В исследовании внешний источник моделируется как агент, находящийся за пределами основной среды с фиксированным расстоянием 1 до каждого индивидуума, что отражает его всепроникающую природу. Значение S_{Mi} меняется для каждого индивидуума, поскольку уровень воздействия СМИ различен; эта величина аналогична «доверительному весу» индивидуума к внешним сообщениям. В контексте социальных сетей роль СМИ могут выполнять силы влияния хэштегов, ссылок или рекламы в сообществе пользователей. На этой основе формируется итоговая формула:

$$I_u = -\beta \cdot \sum_{i=1}^{N} Count_Mentions_u(u_i) - \sum_{i=1}^{|Hashtags(u)|} \sum_{\substack{j=2 \\ i>j}}^{|Hashtags(u)|} Hashtags_u(h_i, h_j) - \sum_{i=1}^{N} \sum_{\substack{j=2 \\ i>j}}^{N} \frac{Mentions_u(u_i, u_j)}{\rho^{\alpha}(u_i, u_j)},$$

Здесь учитываются все пользователи, упоминаемые u, также хэштеги. Соответсвенно:

$$I_u = -\beta \cdot \sum_{i=1}^{N} Count_Mentions_u(u_i) - \sum_{i=1}^{|Urls(u)|} \sum_{\substack{j=2 \\ i>j}}^{|Urls(u)|} Urls_u(url_i, url_j) - \sum_{i=1}^{N} \sum_{\substack{j=2 \\ i>j}}^{N} \frac{Mentions_u(u_i, u_j)}{\rho^{\alpha}(u_i, u_j)}$$

При учёте рекламного воздействия на сообщество формула может быть записана следующим образом:

$$I_u = -\beta \cdot \sum_{i=1}^{N} Followers_Count(u_i) - \sum_{i=1}^{N} Followers_Count(s) O_i O_s$$
$$- \sum_{i=1}^{N} \sum_{\substack{j=2 \\ i>j}}^{N} \frac{Followers_Count(u_j) O_j O_i}{d^{\alpha}(u_i, u_j)}$$

Здесь количество подписчиков сообщества - считаем как силу влияния сообщества на данного пользователя. Если есть реклама на сообщество, то O_s принимает значение +1, иначе –1.

Результаты и обсуждение. Наиболее удобным методом сбора данных является обращение к специализированным компаниям, осуществляющим сбор и постоянное обновление информации из различных источников. Преимущество данного подхода заключается в оперативности получения данных, что критично при большом объёме клиентской базы и использовании нескольких социальных сетей. К недостаткам относится платная подписка на обновления.

Ещё один подход к сбору данных — применение программных интерфейсов (API), предоставляемых популярными социальными сетями. Набор доступных данных, ограничения по числу запросов и стоимость доступа зависят от конкретной сети. Так, через программный интерфейс «ВКонтакте» доступна практически полная информация о пользователях, тогда как Facebook предоставляет API с крайне ограниченными данными. Среди недостатков метода — ограничения на количество параллельных запросов и на общую частоту обращений приложения. Важно регулярно контролировать изменения в API и поддерживать актуальность приложения для сбора данных; при этом часть социальных сетей предоставляет существенную информацию только за плату. Метод позволяет получать данные о пользователе в структурированном виде (JSON или XML) и обеспечивает простую интеграцию API-вызовов в приложение. Альтернативный способ — ручной парсинг веб-страниц или применение готовых краулеров для извлечения данных с последующей обработкой. Преимущества метода заключаются в доступе ко всем открытым данным и отсутствии ограничений на скорость их получения. Среди недостатков — высокая сложность реализации из-за уникальности страниц каждой социальной сети, необходимость поддерживать актуальные правила парсинга и значительные вычислительные ресурсы, однако процесс легко распараллеливается.

Все способы сбора данных из социальных сетей обладают как достоинствами, так и недостатками. Наиболее оперативным методом является использование открытого API, хотя он может накладывать ограничения на количество обращений и доступные сведения о пользователя (API., Vkontakte, Richardson, Amundsen & Ruby, 2013)

Созданный в процессе исследования программный комплекс включает модули для извлечения, обработки, анализа и визуализации данных социальных сетей. Все компоненты реализованы на языке Python. Модуль сбора данных поддерживает извлечение информации прежде всего из крупнейших социальных сетей — Twitter и ВКонтакте, с использованием их API.

Программный интерфейс ВКонтакте предоставляет возможность

извлечения данных из базы vk.com через HTTP-запросы к серверу. Пользователю не требуется понимать внутреннюю структуру базы данных или типы полей — API самостоятельно обрабатывает эти детали (Richardson, et al., 2013). Таким образом, запросы к API позволяют получать нужную информацию напрямую с сервера ВКонтакте.

При работе с методами API требуется передавать в запросе access_token — уникальный ключ доступа, представленный строкой из латинских букв и цифр, соответствующий пользователю, сообществу или приложению. Данные хранятся в документно-ориентированной базе данных MongoDB (Banker, 2011). Структура программы приведена в рисунке – 1.

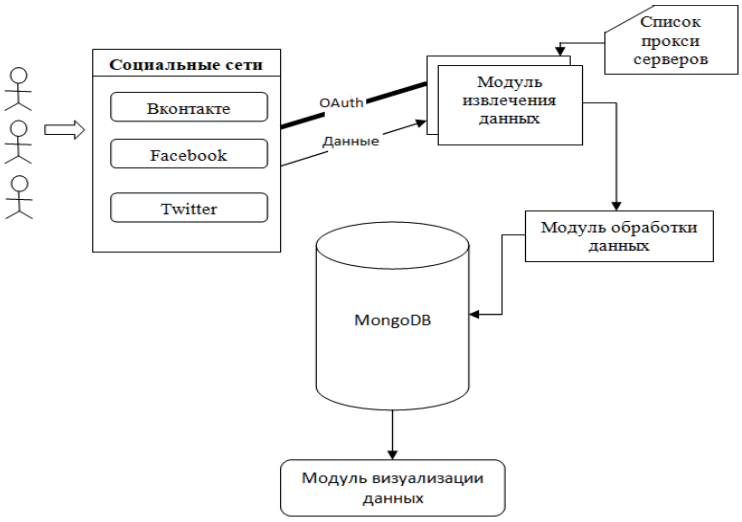


Рисунок 1 – Структура программы

Данные пользователей классифицируются на три группы:

1)персональная информация, включая имя, ник и дату регистрации;

2)сообщения пользователей;

3)отношения между пользователями. В таблице №1 показаны результаты тестирования с уровнями влияния подписчиков в последней колонке

Таблица1 – Результаты тестирования

Друзья A.Akhmetova	Количество друзей	Количество подписчиков	Состоит ли в группе	Расстояние Хэмминга	Уровень влияния
Asselya Moldasheva	439	828	1	7	118,268
Дильшат Delsh Аширов	378	67	–1	8	44,375
Бейбіт Тузелбаев	501	128	1	7	69,2857
Сергей undertaker Коробицин	133	34	1	7	18,85714
Балым Каратаева	267	186	–1	7	26,57143
DauletSabyr	65	65	1	7	9,28571

Из поученных результатов можно делать вывод какие факторы больше воздействуют на мнение индивидуума. Это количество друзей и подписчиков, чем больше, тем влиятельней пользователь, который отличается высоким социальным статусом и лучшей информированностью, оказывающий влияние на мнение других людей.

Выводы. В ходе работы были предложены методы анализа данных, получаемых из социальной сети. Рассмотрены методы изучения взаимоотношений между участниками сети, прогнозирования их поведения в сетях. Основная цель исследовательской работы-разработка новых методов и моделей анализа данных в социальных сетях и реализация соответствующего программного обеспечения.

Основные результаты:

1. получены количественные характеристики, отношения и множества рассчитываемые на основе данных из социальных сетей.

2. предложен метод определения социального воздействия на основе теории Латане.

3. подготовлена программа, позволяющая осуществлять сбор и анализ данных из социальных сетей.

Литература

Aggarwal C.C. Social Network Data Analytics. — 520 p. — ISBN 978-1-4419-8461-6. – DOI 10.1007/978-1-4419-8462-3. – New York: Springer, 2011.

API. [Электронный ресурс]. Режим доступа: https://ru.wikipedia.org/wiki/API (дата обращения: 08.09.2024)

Banker K. MongoDB in Action. — Greenwich, CT: Manning Publications, 2011.

Charu C. Aggarwal (2011) Social network data analytics. — 520 p. ISBN 978-1-4419-8461-6 e-ISBN 978-1-4419-8462-3 DOI 10.1007/978-1-4419-8462-3

DEUTSCH M., GERARD H.B. A Study of Normative and Informational Social Influences upon Individual Judgment. Journal of Abnormal and Social Psychology, 1955. — №51. — P. 629-636 https://doi.org/10.1037/h0046408

Latané B. (1981) The Psychology of Social Impact. American Psychologist, 36(4). — P. 343–356.

Latané B. (1996) Dynamic Social Impact: The Creation of Culture by Communication. Journal of Communication, 46(4). — P. 13–25.

MongoDB — dokumentoorientirovannaya sistema upravleniya bazami dannykh [Elektronnyi resurs]. — URL: https://ru.wikipedia.org/wiki/MongoDB (data obrashcheniya: 08.09.2024).

Newman M.E.J. Networks: An Introduction. — Oxford: Oxford University Press, 2010.

Nowak A., Szamrej J., Latane B. (1990) From private attitude to public opinion: a dynamic theory of social impact. Psychological Review, 97. — P. 362-376. https://doi.org/10.1037/0033-295X.97.3.362

Richardson L., Amundsen M., & Ruby S. (2013) RESTful Web APIs. O'Reilly Media. — Современные подходы к проектированию RESTful API.

Wasserman S., Faust K. Social Network Analysis: Methods and Applications. — Cambridge: Cambridge University Press, 1994

Wragg T. (2006) Modeling the Effects of Information Campaigns Using Agent-Based Simulation. Prep.: Command and Control Division, Defense Science and Technology Organization, Australian Government, DSTO-TR-1853. — P. 61. (URL http://www.dsto.defence.gov.au/corporate/reports/DSTO-TR-1853.pdf)

Батура Т.В. (2012) Методы анализа компьютерных социальных сетей. Вестник НГУ. Серия:

Информационные технологии, Новосибисрк, Том 10. — Вып. 4. — С. 13-28. ISSN1818-7900 (URL https://e-lib.nsu.ru/dsweb/Get/Resource-1107/Inform-2012-10-4.pdf/en/view)

Веб-сайт ВКонтакте. Работа с API. [Электронный ресурс]. Режим доступа: https://dev.vk.com/ru/method (дата обращения: 07.08.2024).

Вэй Тан, Брайан Блейк, Иман Салех (2013) Аналитика Больших Данных и социальные сети. Открытые системы.СУБД. — № 8. — С. 37–41. (URL http://www.osp.ru/os/2013/08/13037856)

References

Aggarwal C.C. Social Network Data Analytics. — 520 p. — ISBN 978-1-4419-8461-6. – DOI 10.1007/978-1-4419-8462-3. – New York: Springer, 2011. (in Eng.)

API [Elektronnyi resurs]. — URL: https://ru.wikipedia.org/wiki/API (data obrashcheniya: 08.09.2024). (In Russ.)

Banker K. MongoDB in Action. — Greenwich, CT: Manning Publications, 2011. (in Eng.)

Batura T.V. Metody analiza komp'yuternykh sotsial'nykh setei [Methods of Computer Social Network Analysis] Vestnik NGU. Seriya: Informatsionnye tekhnologii, Novosibirsk, 2012. — Vol. 10, No. 4. — P. 13–28. — ISSN 1818-7900. — URL: https://e-lib.nsu.ru/dsweb/Get/Resource-1107/Inform-2012-10-4.pdf/en/view (data obrashcheniya: 08.09.2024). (In Russ.)

Deutsch M., Gerard H.B. A Study of Normative and Informational Social Influences upon Individual Judgment. Journal of Abnormal and Social Psychology, 1955. — Vol. 51. — P. 629–636. — DOI: 10.1037/h0046408. (in Eng.)

Latané B. The Psychology of Social Impact. American Psychologist, 1981. — Vol. 36(4). — P. 343–356. (in Eng.)

Latané B. Dynamic Social Impact: The Creation of Culture by Communication. Journal of Communication, 1996. — Vol. 46(4). — P. 13–25. (in Eng.)

MongoDB — dokumentoorientirovannaya sistema upravleniya bazami dannykh [Elektronnyi resurs]. — URL: https://ru.wikipedia.org/wiki/MongoDB (data obrashcheniya: 08.09.2024). (in Eng.)

Newman M.E.J. Networks: An Introduction. — Oxford: Oxford University Press, 2010. (in Eng.)

Nowak A., Szamrej J., Latané B. From Private Attitude to Public Opinion: A Dynamic Theory of Social Impact. Psychological Review, 1990. — Vol. 97. — P. 362–376. — DOI: 10.1037/0033-295X.97.3.362. (in Eng.)

Richardson L., Amundsen M., Ruby S. RESTful Web APIs. — Sebastopol, CA: O'Reilly Media, 2013. (In Russ.)

Veb-sait VKontakte. Rabota s API [Elektronnyi resurs]. — URL: https://dev.vk.com/ru/method (data obrashcheniya: 07.08.2024). (In Russ.)

Vey Tan, Braian Bleik, Iman Saleh. Analitika Bol'shikh Dannykh i sotsial'nye seti [Big Data Analytics and Social Networks]. Otkrytye sistemy. SUBD, 2013. — No. 8. — P. 37–41. — URL: http://www.osp.ru/os/2013/08/13037856 (data obrashcheniya: 08.09.2024). (In Russ.)

Wasserman S., Faust K. Social Network Analysis: Methods and Applications. — Cambridge: Cambridge University Press, 1994 (in Eng.)

Wragg T. Modeling the Effects of Information Campaigns Using Agent-Based Simulation. — Command and Control Division, Defense Science and Technology Organization, Australian Government, 2006. — 61 p. — URL: http://www.dsto.defence.gov.au/corporate/reports/DSTO-TR-1853.pdf. (in Eng.)

ACADEMIC SCIENTIFIC JOURNAL OF COMPUTER SCIENCE
ISSN 1991-346X
Volume 4. Namber 356 (2025).102–121

https://doi.org/10.32014/2025.2518-1726.386

UDC 004.056.2

©A.Sh. Barakova[1,2], K.S. Shadinova[1], A.S. Orynbaeva[3], G. Sugurzhanova[4], 2025.
[1] Asfendiyarov Kazakh National Medical University, Almaty, Kazakhstan;
[2]Al-Farabi Kazakh National University, Almaty, Kazakhstan;
[3]Astana Medical University, Astana, Kazakhstan;
[4]NEI «Kazakhstan-Russian Medical University», Almaty, Kazakhstan.
E-mail: balia_79@mail.ru

DESIGN OF A MODEL FOR PROTECTING A WEBSITE'S AUTHENTICATION DATA AND CONTENT BASED ON BLOCKCHAIN TECHNOLOGY

Barakova Aliya — Assistant Professor of the Department of Engineering Disciplines and Good Practices, Asfendiyarov Kazakh National Medical University; doctoral student of Al-Farabi Kazakh National University, Almaty, Kazakhstan,
E-mail: balia_79@mail.ru, https://orcid.org/0000-0002-0904-745X;
Shadinova Kunsulu — Associate Professor Shadinova Kunsulu Seidazovna Kazakh National Medical University named after Asfendiyarov, Almaty, Kazakhstan,
E-mail: Shadinkunsulu@gmail.com; https://orcid.org/0009-0006-5534-7927;
Orynbayeva Ainur — senior lecturer at Astana Medical University. Astana, Kazakhstan,
E-mail: ainur_tas@mail.ru, https://orcid.org/0009-0005-4984-281X;
Sugurzhanova Gulzhan — Senior Lecturer of the Department of Social Sciences, NEI «Kazakhstan-Russian Medical University», Almaty, Kazakhstan,
E-mail: sugurzhanova83@mail.ru, https://orcid.org/0000-0001-8334-6856.

Abstract. The article introduces a new PBFT-based consensus algorithm designed to enhance the scalability, reliability, and security of consortium blockchain systems. The approach is implemented within a four-layer architecture—network, consensus, application, and meta-application—and relies on segmenting nodes into three functional groups. Proxy nodes perform local consensus, global proxy nodes aggregate blocks across segments, while supernodes supervise system-wide operations and ensure reliability. The algorithm incorporates guarantee mechanisms, trusted node selection, and dual-leader monitoring to strengthen fault tolerance and mitigate malicious behavior. Theoretical and experimental evaluations confirm the efficiency of the segmented consensus structure. The required number of nodes decreases from 3f+1 to 2f+1, improving resource utilization. Consensus latency ranges from 80 to 147 ms, demonstrating a 62–70% improvement compared to classical PBFT. Communication complexity is optimized to the O(n)–O(n²) range.

Simulations conducted in Python, with visualization through Matplotlib, validate the performance model. The algorithm was implemented in a working prototype: content hashing and IPFS integration were developed in Python, while a Solidity smart contract records ownership metadata, hash values, and IPFS CIDs. The prototype, deployed on the *www.AliyaSchool.kz* platform and tested under real conditions, confirmed the correct execution of the smart contract via Remix IDE. Overall, the proposed PBFT-based algorithm provides an effective solution to the blockchain scalability problem, enabling distributed load handling, accelerated consensus, and improved detection of malicious nodes. The work contributes a scientifically grounded model to the national research base and offers methodological value for further development of secure blockchain mechanisms for web systems.

Keywords: blockchain, consensus algorithm, scalability, security, PBFT, smart contract, IPFS

© А.Ш. Баракова,[1,2], К.С. Шадинова[1], А.С. Орынбаева[3], Г. Сугуржанова[4], 2025.

[1] С. Асфендияров Атындағы Қазақ ұлттық медицина университеті, Алматы, Қазақстан;

[2]Әл-Фараби атындағы Қазақ ұлттық университеті, Алматы, Қазақстан;

[3]Астана медицина университеті, Астана, Қазақстан;

[4]Қазақстан-Ресей медициналық университет, Алматы, Қазақстан.
E-mail: balia_79@mail.ru

БЛОКЧЕЙН ТЕХНОЛОГИЯСЫ НЕГІЗІНДЕ ВЕБ САЙТТЫҢ АУТЕНТИФИКАЦИЯЛЫҚ ДЕРЕКТЕРІ МЕН ӨНІМІН ҚОРҒАУ МОДЕЛІН ҚҰРАСТЫРУ

Баракова Алия — профессор ассистент, С.Ж. Асфендияров атындағы Қазақ ұлттық медицина университеті; Әл-Фараби атындағы Қазақ ұлттық университетінің докторанты, Алматы, Қазақстан,
E-mail: balia_79@mail.ru, https://orcid.org/0000-0002-0904-745X;
Шадинова Күнсұлу — қауымдастырылған профессор, С.Ж. Асфендияров атындағы Қазақ ұлттық медицина университеті, Алматы, Қазақстан,
E-mail: Shadinkunsulu@gmail.com, https://orcid.org/0009-0006-5534-7927;
Орынбаева Айнұр — Астана медицина университетінің аға оқытушысы, Астана, Қазақстан,
E-mail: ainur_tas@mail.ru, https://orcid.org/0009-0005-4984-281X;
Сугуржанова Гүлжан — «Қазақстан-Ресей медициналық университеті» Әлеуметтік ғылымдар кафедрасының аға оқытушысы, Алматы, Қазақстан,
E-mail: sugurzhanova83@mail.ru, https://orcid.org/0000-0001-8334-6856.

Аннотация. Бұл мақала консорциумдық блокчейн жүйелерінің ауқымдылығын, сенімділігін және қауіпсіздігін арттыруға бағытталған PBFT негізіндегі жаңа консенсус алгоритмін таныстырады. Тұғыр төрт деңгейлі архитектураға негізделген: желілік, консенсус, қолданбалық және мета-қолданбалық. Жүйе түйіндерді үш функционалдық топқа сегменттеу

арқылы жұмыс істейді. Прокси түйіндер жергілікті консенсусты орындайды, жаһандық прокси түйіндер сегменттер арасынан блоктарды біріктіреді, ал супернода жүйе жұмысын жалпы бақылауды қамтамасыз етеді. Алгоритмде кепілдік механизмдері, сенімді түйіндерді таңдау стратегиясы және қос көшбасшыны бақылау тәсілі қолданылып, ақауларға төзімділік күшейтіледі және зиянды әрекеттердің алдын алады. Теориялық және тәжірибелік бағалау сегменттелген консенсус құрылымының тиімділігін дәлелдейді. Қажетті түйіндер саны 3f+1-ден 2f+1-ге дейін азайып, ресурстарды пайдалану жақсарады. Консенсус кідірісі 80–147 мс аралығында болып, дәстүрлі PBFT-тен 62–70 процент жылдамырақ нәтижені көрсетеді. Байланыс күрделілігі O(n)–O(n^2) ауқымында оңтайландырылды. Python тіліндегі модельдеу және Matplotlib визуализациясы алгоритмнің өнімділігін растады. Алгоритмнің жұмыс прототипі жүзеге асырылды: контент хештеу және IPFS-пен интеграция Python арқылы орындалды, ал Solidity тіліндегі смарт-келісім контент иесін, хеш мәнін және IPFS CID-ін тіркейді. *www.AliyaSchool.kz* платформасында сынақтан өткізілген прототип Remix IDE арқылы смарт-келісімнің дұрыс орындалуын дәлелдеді. Жалпы алғанда, ұсынылған PBFT негізіндегі алгоритм блокчейннің ауқымдылық мәселесін тиімді шешіп, жүктемені бөлуді, консенсусты жылдамдатуды және зиянды түйіндерді анықтауды қамтамасыз етеді. Бұл жұмыс ұлттық ғылыми базаға үлес қосып, веб-жүйелерге арналған қауіпсіз блокчейн механизмдерін дамытуға әдістемелік негіз ұсынады.

Түйін сөздер: блокчейн, консенсус алгоритмі, масштабталу, қауіпсіздік, PBFT, смарт-контракт, IPFS

© А.Ш. Баракова[1,2], К.С. Шадинова[1], А.С. Орынбаева[3], Г. Сугуржанова[4], 2025.

[1]Казахский национальный медицинский университет имени С.Д. Асфендиярова, Алматы, Казахстан;

[2]Казахский национальный университет имени аль-Фараби, Алматы, Казахстан;

[3]Медицинский университет Астана, Астана, Казахстан;

[4]НЕИ «Казахстанско-Российский медицинский университет, Алматы, Казахстан.

E-mail: balia_79@mail.ru

РАЗРАБОТКА МОДЕЛИ ЗАЩИТЫ АУТЕНТИФИКАЦИОННЫХ ДАННЫХ И КОНТЕНТА ВЕБ-САЙТА НА ОСНОВЕ ТЕХНОЛОГИИ БЛОКЧЕЙН

Баракова Алия — ассистент профессора, Казахский национальный медицинский университет имени С.Д. Асфендиярова; докторант Казахского национального университета имени аль-Фараби, Алматы, Казахстан,
E-mail: balia_79@mail.ru, https://orcid.org/0000-0002-0904-745X;

Шадинова Күнсулу — доцент Казахского национального медицинского университета имени С.Д. Асфендиярова, Алматы, Казахстан,
E-mail: Shadinkunsulu@gmail.com, https://orcid.org/0009-0006-5534-7927;
Орынбаева Айнур — старший преподаватель Астанинского медицинского университета, Астана, Казахстан,
E-mail: ainur_tas@mail.ru, https://orcid.org/0009-0005-4984-281X;
Сугуржанова Гулжан — старший преподаватель кафедры социальных наук НЕИ «Казахстанско-Российский медицинский университет», Алматы, Казахстан,
E-mail: sugurzhanova83@mail.ru, https://orcid.org/0000-0001-8334-6856.

Аннотация. В данной статье представлен новый консенсусный алгоритм на основе PBFT, направленный на повышение масштабируемости, надежности и безопасности консорциумных блокчейн-систем. Подход реализован в четырехуровневой архитектуре: сетевом, консенсусном, прикладном и мета-прикладном уровнях. Работа системы основана на сегментации узлов на три функциональные группы. Прокси-узлы выполняют локальный консенсус, глобальные прокси-узлы агрегируют блоки между сегментами, а суперузлы обеспечивают общее системное наблюдение. Алгоритм дополняется механизмами гарантий, стратегией выбора доверенных узлов и контролем двух лидеров, что усиливает устойчивость к сбоям и снижает риски вредоносного поведения. Теоретическая и экспериментальная оценка подтверждает эффективность сегментированной архитектуры консенсуса. Количество необходимых узлов снижается с 3f+1 до 2f+1, что улучшает использование ресурсов. Задержка консенсуса составляет 80–147 мс и демонстрирует ускорение на 62–70% по сравнению с классическим PBFT. Сложность коммуникаций оптимизирована в диапазоне от O(n) до O(n^2). Моделирование, выполненное на Python с визуализацией в Matplotlib, подтвердило работоспособность модели. Алгоритм был реализован в виде рабочего прототипа: хеширование контента и интеграция с IPFS выполнены на Python, а смарт-контракт на Solidity регистрирует владельца контента, хеш и IPFS CID. Прототип, протестированный в реальных условиях на платформе *www.AliyaSchool.kz*, подтвердил корректное выполнение смарт-контракта через Remix IDE. В целом, предложенный алгоритм на основе PBFT эффективно решает проблему масштабируемости блокчейна, обеспечивает распределение нагрузки, ускоренный консенсус и улучшенное обнаружение вредоносных узлов. Работа формирует научно обоснованную модель и служит методологической базой для дальнейшего развития защищенных блокчейн-механизмов для веб-систем.

Ключевые слова: блокчейн, консенсусный алгоритм, масштабируемость, безопасность, PBFT, смарт-контракт, IPFS

Кіріспе. Бүкіл әлемдегі инфрақұрылымның маңызды салаларына кибершабуылдар саны жыл сайын тұрақты өсіп келеді. Қазіргі заманғы ақпараттық жүйелердің басым көпшілігі веб-сайттар ретінде құрылады,

сондықтан қауіпсіздікке ерекше назар аудару қажет. Яғни кез- келген операциялық жүйесі мен бағдарламалық қосымшасы бар құрылғы пайда болған соң, ол бірден тәртіп бұзушылардың қызығушылығын тудырады. Олар оны зерттей бастайды және оны бұзуға күш жұмсайды. Бүгінгі күні веб-сервистерді қорғау барлық ұйымдарды алаңдатады, себебі қасақана субъектілер оны рұқсатсыз пайдалану мақсатында жеке немесе заңды тұлға туралы кез келген ақпаратты алуға ұмтылады.

Әлемдік және отандық экономиканы трансформациялаудың қазіргі кезеңі коммерциялық қызмет пен цифрлық технологиялар арасындағы интеграциялық процестерді күшейтуге ықпал етеді, олардың соңғысы бизнес субъектілерінің қарқынды дамуының басты факторы болып табылады.

Әдебиеттерге шолу. Шетелдік ғалымдар Huynh T. және әріптестері (Huynh et al, 2019) блокчейн технологиясының қауіпсіздігі мен құпиялылығы бойынша жан-жақты шолу жасап, кең таралған шабуыл түрлерін сипаттап, олардан қорғану жолдарын ұсынады.

X. Li және авторлар тобы (Li et al, 2020) блокчейн жүйелеріне бағытталған шынайы шабуылдар мен қауіп-қатерлерді жүйелі түрде талдап, олардың негізінде қауіпсіздік мәселелерінің ауқымын сипаттайды. Олар қазіргі заманғы блокчейн инфрақұрылымдарының осал тұстарын ашып көрсетіп, шабуылдардың техникалық ерекшеліктеріне терең тоқталады

Басқа бір маңызды үлес – J. Leng және авторлар ұсынған (Leng et al, 2022) PDI (Process-Data-Infrastructure) моделі. Бұл үлгі блокчейн қауіпсіздігі мәселелерін үш негізгі деңгейде – процесс, дерек, және инфрақұрылым деңгейінде жіктеп, талдау жасауға мүмкіндік береді. Зерттеушілер техникалық зерттеулер мен бизнес міндеттер арасындағы алшақтықты жоюға тырысып, қауіпсіздік проблемаларын жүйелі түрде шешудің жолдарын іздейді.

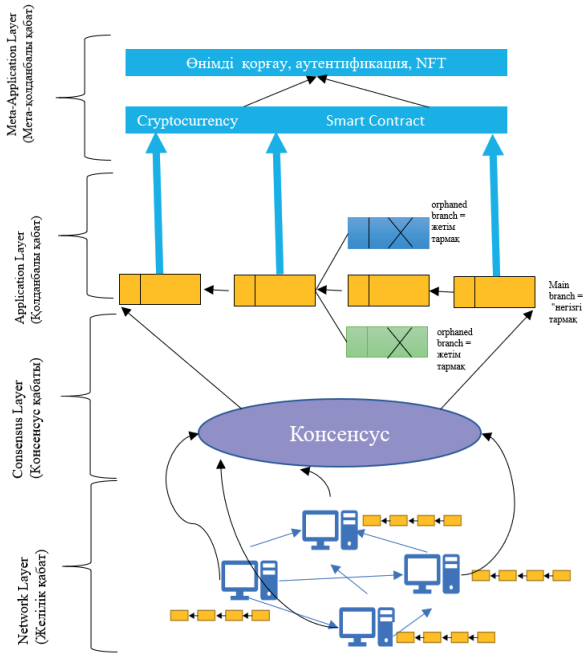
Әдістер мен материалдар. Бұл зерттеу жұмыста ақпараттарды қорғаудың тиімді тәсілін ұсыну. Ақпаратты қорғаудың қазіргі таңда көп мөлшерде түрлері көп. Соңғы заманауи технолгиялардың бірі блокчейн технологиясы зерттеу жұмыстың өзегі болып табылады. Блокчейн технологиясы мінсіз әрі сенімді технология ретінде өте танымал. Дегенмен оның да кемшіл тұстары бар. Блокчейн технолгиясының ең басты кемшіліктерінің бірі транзакция жылдамдығы мен масштабталу мәселесі болып табылады (Bonneau et al, 2015).

Масштабталу мәселесін шешуге арналған жаңа консенсус алгоритмін әзірлеу мақаланың негізгі кілті. Пайдаланушыларды жылдам әрі қауіпсіз аутентификациялау және веб-контентті сенімді қорғау үшін масштабталатын және тиімді консенсус алгоритмі қажет. Мәселенің шешімі ретінде жаңа консенсус механизмін құрастыру керек. Зерттеу жұмыста сегменттеуге негізделген PBFT негізіндее құрылған жаңа консенсус алгоритмы ұсынылады, ол консорциумдық блокчейннің орталықсыздандыру, қауіпсіздік және масштаблауы талаптарына сай болуын қамтамасыз етуге бағытталған (Popadyuk et al, 2020, Sompolinsky et al, 2016).

PBFT негізінде құрылған жаңа алгоритмін құру үшін 1 - суретте блокчейн жүйесінің архитектурасын төрт негізгі қабат бойынша сипаттайды: желілік қабат (түйіндер арасындағы байланыс), консенсус қабаты (шешімге келу үдерісі), қолданбалы қабат (криптовалюта және смарт-келісімшарттар), және мета-қолданбалы қабат (өнімді қорғау, аутентификация және NFT). Әр қабат өзіне тән міндеттерді атқара отырып, жүйенің қауіпсіздігін, тұрақтылығын және масштабталуын қамтамасыз етеді. Жаңа алгоритмді құру осы төрт кезеңнің үйлесімділігін ескере отырып, блокчейн технологиясын онлайн контентті қорғау мақсатында тиімді пайдаланылады.

Жаңа консенус алгоритмі PBFT негізінде құрылады. Себебі PBFT алоритмі BFT алгоритмінің төмен тиімділігі мәселесін шешеді, ал коммуникацияның күрделілігі экспоненциалдыдан көпмүшелікке өзгереді (Somin et al, 2018).

Блокчейн архитектурасының негізінде жатқан консенсус қауіпсіздік, тиімділік және масштабталуды оңтайландыру тұрғысынан сервистік бағдарланған қосымшаларды жүзеге асырудың кілті болып табылады. Консенсус алгоритмдерінің кейбір күрделі түрлерінде, мысалы, Practical Byzantine Fault Tolerance (PBFT), желідегі түйіндер саны артқан сайын өткізу қабілеті күрт төмендейді. Ал қарапайым алгоритмдерде, мысалы, Raft жүйесінде, желі көлемі ұлғайған сайын жетекшіге түсетін жүктеме артып, консенсус тиімділігіне теріс әсер етеді. Осы мәселелерді шешу үшін біз консорциумдық блокчейн негізінде сенімділік ағашына сүйенген түйіндер сегменттерінің моделін ұсынамыз.



Сурет 1- Алгоритм кезеңдері

Бұл модель жоғары масштабталуды белгілі бір деңгейдегі орталықсыздандыру мен қауіпсіздікті сақтай отырып қамтамасыз етеді.

Біріншіден, біз түйіндер арасындағы сенімділік қатынастарын бейнелейтін кепілдік механизмін құрамыз, содан кейін осы механизм негізінде түйіндерді таңдау стратегиясын әзірлейміз. Бұл стратегия түйіндердің кепілдік нәтижелерін және консенсус әрекетін бағалап, сенімділік мәртебесін анықтап, зиянды түйіндерді анықтап, сенімді көшбасшылар тізімін жасақтайды. Екіншіден, біз қос көшбасшы бақылау механизмін ұсынамыз, онда резервтік көшбасшы негізгі көшбасшының белсенділігін бақылап отырады, ал оның белсенділігі консенсус түйіндерімен бағаланады.

Жаңа PBFT негізінде құрылған алгоритм алдымен кластерлеу алгоритмын қолданып, желідегі түйіндерді аралас атрибуттарына сәйкес сегменттерге бөледі. Сегменттеу технологиясы алғаш рет дерекқорларды бөліктерге бөлу үшін қолданылған. Бұл технология үлкен дерекқорларды басқаруға ыңғайлы, жылдам және кіші бөліктерге бөлуге мүмкіндік береді. Блокчейнге қатысты сегменттеу технологиясы үш түрге бөлінеді: желіні сегменттеу, транзакцияны сегменттеу және күйді сегменттеу. Elastico20 протоколы — блокчейн консенсус алгоритмінде сегменттеу технологиясын алғаш қолданған протокол. Алайда бұл протокол ашық блокчейнге арналғандықтан, түйіндерді тексеруге ынталандыру үшін экономикалық стимулдарды қажет етеді, сондықтан консорциумдық желіге жарамайды.

Содан кейін, әртүрлі сегменттерде параллельді консенсусқа қол жеткізу үшін қиылысатын емес транзакциялар пайдаланылды. Сегмент өлшемін таңдау биномдық ықтималдықтар үлестірімі арқылы талданады, сондай-ақ әртүрлі түйіндердің істен шығу ықтималдығына байланысты жүйенің жаһандық блокты сәтті қалыптастыру мүмкіндігі зерттелді. Желілік өзара әрекеттестік — бұл сенімді тең дәрежелі (peer-to-peer) желі. Келісілген шешімдер қабылдайтын түйіндер басқа түйіндерден хабарламалар ала алады. Кең тарату (broadcast) хабарламалары жіберушіні желідегі барлық түйіндерге, соның ішінде өзіне де хабарлама жіберуге мәжбүрлейді (Баракова et al, 2022).

Бұл жұмыста жартылай синхронды (semi-synchronous) желілік модель пайдаланылады. Хабарламаны жіберу уақытының жоғарғы шегі белгіленеді – бұл хабарламаның жіберілуі мен қабылдануы арасындағы уақыт аралығы. Егер жіберу уақыты осы шектен асып кетсе, хабарлама өңделмейді.

Түйіндік модель.

Дәстүрлі консенсус алгоритмінде түйіндер негізінен лидер-түйіндерге және консенсус түйіндеріне бөлінеді. Жаңа алгоритмде монитор концепциясы енгізілген. Түйіндердің әрқайсысының толық сипаттамасы төменде 2 кестеде берілген.

Кесте 2 - Түйіндердің әрқайсысының толық сипаттамасы

№	Түйін атаулары	Міндеті:	Артықшылықтары
1	Жалпы түйіндер (Full Nodes)	- Бүкіл блокчейн тарихын (ledger) толық сақтайды. - Барлық транзакцияларды тексеріп, оларды желіге таратады. - Желідегі басқа түйіндерге қажетті деректерді ұсынады.	Жоғары сенімділік: кез келген уақытта блокчейннің толық көшірмесін ұсынуға қабілетті.
2	Прокси-түйіндер (Proxy Nodes)	- Сегмент ішінде транзакцияларды өңдейді. - Клиенттердің сұрауларын қабылдап, оларды локалды консенсус арқылы өңдейді. - Алынған блоктарды Глобал прокси-түйіндерге жібереді	Масштабталуды арттырады, себебі барлық түйіндер бірден барлық транзакцияларды өңдемейді.
3	Глобал прокси-түйіндер (Global Proxy Nodes)	- Барлық сегменттерден ақпарат жинайды. - Әр сегменттен алынған блоктарды біріктіріп, бірыңғай блок құрайды. - Құрылған жаңа глобал блокты басқа прокси-түйіндерге таратады.	Желі тұрақтылығын сақтайды – әрбір сегмент жеке жұмыс істегенімен, глобалды блок арқылы барлық сегменттер бірдей синхрондалады.
4	Басқарушы түйіндер (Super Nodes)	- Түйіндердің мінез-құлқын бақылайды. - Аномалияларды анықтап, зиянкес түйіндерді блоктайды	Басқарушы түйіндер зиянкес түйіндерді анықтап, жүйенің тұрақтылығын

Консенсус түйіндері

Жүйеде консенсусқа қатысатын репликалар консенсус түйіндері деп аталады. Консенсус түйіндері – бұл блокчейн жүйесінде транзакцияларды тексеруге, блоктарды растауға және консенсусқа қол жеткізуге жауапты арнайы желі қатысушылары. Бұл түйіндер желідегі сенімділік пен қауіпсіздікті қамтамасыз етеді, себебі олар ереже бойынша жұмыс істейді және жүйенің шынайылығын сақтайды.

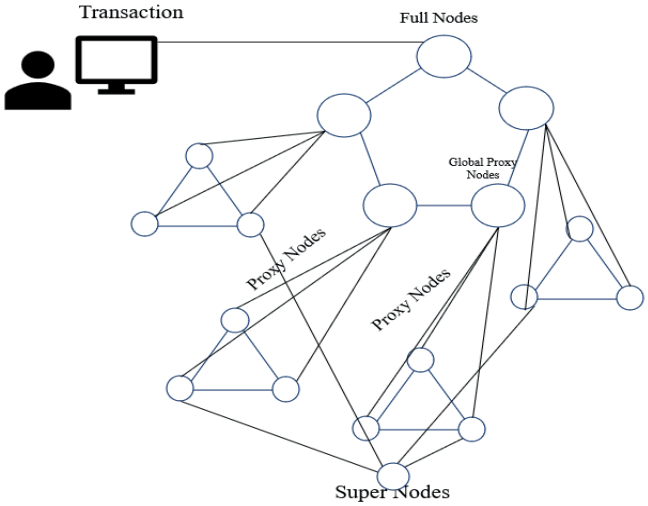
Біздің алгоритм сегментацияланған консенсусқа негізделгендіктен, онда түйіндер әртүрлі рөлдерге бөлінеді. Бұл архитектура блокчейннің масштабталуын, қауіпсіздігін және тиімділігін арттыруға мүмкіндік береді. Жаңа консенсус алгоритмі төрт деңгейлі түйіндер жүйесін қолданады, бұл масштабталу мен қауіпсіздікті жақсартады (Абдулжалилов 2023)

Сонымен:

- Жалпы түйіндер бүкіл блокчейн тарихын сақтайды.
- Прокси-түйіндер локалды консенсус жүргізіп, жүктемені азайтады.
- Глобал прокси-түйіндер барлық сегменттерден блоктарды біріктіреді.
- Басқарушы түйіндер зиянкес түйіндерді анықтап, жүйенің тұрақтылығын қамтамасыз етеді.

$N \geq 3f + 1$ формула бойынша, егер жүйеде N түйін болса, оның f түйіні Византиялық шабуылдаушы болуы мүмкін. Қауіпсіз консенсус үшін түйіндер саны $N \geq 3f + 1$ болу керек.

Консенсус түйіндері транзакцияларды тексеріп, желінің қауіпсіз жұмыс істеуін қамтамасыз етеді.



Сурет 2 - Түйіндерді сегменттеу құрылымы

Лидер түйін 2 - суретте Full Nodes бұл блокчейн немесе консенсус алгоритміндегі негізгі басқарушы түйін, яғни желі ішіндегі басқа түйіндердің әрекеттерін үйлестіріп, транзакцияларды ұсынатын және блоктарды жасайтын басты қатысушы.

Лидер жіберген і-нөмірлі түйін N_i келесідей анықталады: Барлық консенсусқа қатысатын түйіндер жиынтығы G символымен белгіленеді. $N_i(i=[1,n], N_i \in G)$ Түйін G ішіндегі басқа түйіндермен консенсус механизмін орнатады.

$$G = (N_1, N_2, \dots, N_i, \dots, N_n) \quad (1)$$

$$N_i = (G_i, st_i, h_i, v_i, l_i, S_i, C_i) \quad (2)$$

1 теңдеуде:

Gi - түйіннің келісілген идентификаторы және IP-мекенжайлар тізімі сақталған (түйін өзі сақтамайды).

st_i - түйіннің күйі, яғни ол лидер, $st_i \in$ орынбасар немесе жай консенсус түйіні екенін білдіреді.

h_i - блок биіктігі – түйін соңғы күйіне жеткен кезде.

v_i - көру нөмірі – консенсус циклы аяқталған кезде және лидер жұбы өзгергенде жаңартылады.

l_i -лидер индексі – мәні консенсус циклы аяқталғаннан кейін немесе орынбасар жаңа лидерлікке талас бастағанда ғана өзгеруі мүмкін.

S_i - түйіннің орналасқан сегменттерінің тізімі – түйін тек өз сегментіндегі транзакцияларды өңдейді.

G_i - гарантия мәні – консенсус түйінінің басқа түйіндерге берген кепілдігі. Монитор түйіні гарантия сұранысын жіберген сайын бұл мән жаңартылып отырады.

Лидер түйін деген ұғым бар бірақ ол классикалық PBFT алгоритмінде бір лидер таңдалады. Лидер түйін деген жаңа блоктарды ұсынады, транзакцияларды реттейді, консенсусқа бағыт береді. Егер лидер бұзылса немесе дұрыс жұмыс істемесе, желі жаңа лидер сайлайды.

Дегенмен жаңа сегменттелген консенсус механизмінде бір орталықтандырылған лидер болмайды, бірақ прокси-түйіндер лидердің қызметін атқарады. Жергілікті (локалды) лидерлер – әрбір сегмент ішінде прокси-түйіндер транзакцияларды жинап, өңдейді. Глобалды лидерлер – глобал прокси-түйіндер барлық сегменттердің блоктарын біріктіреді. Бұл жүйе бір лидерге тәуелділікті азайтады, масштабтауды жақсартады және қауіпсіздікті күшейтеді. Лидердің болмауы қандай артықшылығы шабуылға төзімділік – дәстүрлі PBFT-те лидерге шабуыл жасалса, консенсус процесі баяулайды. Ұсынып отырған алгоритмде әр сегмент өз ішінде жұмыс істейді, сондықтан бүкіл жүйе бұзылмайды.

Бір лидердің орнына бірнеше прокси-түйіндер параллель жұмыс істейді, бұл TPS-ті яғни жылдамдықты арттырады. Масштабталу қосымша түйіндерді оңай қосуға болады, себебі әрқайсысы өз сегментінде жұмыс істейді (Дипа et al, 2016).

Блокчейндегі қолданыстағы PBFT консенсус алгоримдері N–N көп таралымдарын пайдаланады, бұл деректердің сәйкестігін қамтамасыз ету үшін жасалады. Алайда бұл әдіс блокчейн консенсус алгоритмдерінің өзара әрекеттесуінің күрделілігін арттырады және, көрініс өзгерісі орын алса PBFT $O(N^2)$ үшін $O(N^3)$ дейін жетуі мүмкін (Мерсер 2018). Салыстырмалы түрде Raft сияқты таратылған жүйенің консенсус алгоритмі байланыс күрделілігі O(N) болғанына қарамастан, сенімсіз ортада тиімді жұмыс істемейді. Оның лидері өте жоғары талаптарға сәйкес келуі керек, себебі ол бір уақытта көптеген түйіндерге хабар таратады. Егер лидер жұмысын тоқтатса немесе орын ауыстыруы қажет болса, бұл консенсусқа жету жылдамдығын төмендетеді, осылайша блокчейн жүйесінің орташа тиімділігіне әсер етеді.

Нәтижелер мен талқылаулар. Бұл мәселелерді шешу үшін, яғни блокчейн консенсус деңгейінің қауіпсіздік пен тиімділігінің жеткіліксіздігін жою үшін, келесі шешімдер ұсынылады:

1.Түйіндерге кепілдік беру механизмін енгізу – бұл түйіндер арасындағы сенімді қарым-қатынасты орнатуға мүмкіндік береді.

2.Сенімді түйін таңдау стратегиясын әзірлеу – бұл түйіндерді олардың өнімділігі мен мінез-құлқына негізделген бағалау моделін жасауға мүмкіндік береді.

3.Қосарлы көшбасшылық (екі лидер) бақылау механизмін енгізу – бұл таратылған транзакцияларды басқару үшін тиімді жұмыс істейді.

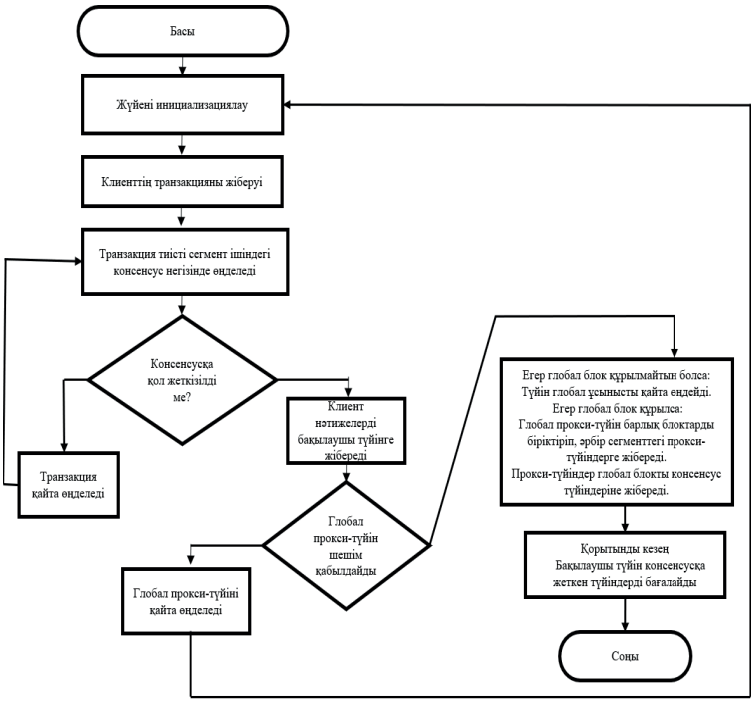
4.Консенсус бөліну моделін енгізу – лидерді сенімді таңдау стратегиясына сүйене отырып таңдау және механизм кепілдігі арқылы сегментацияны жүзеге асыру.

Осы шешмідерден қауіпсіз және масштабталатын консенсус алгоритмін әзірленеді – O(N) күрделілігімен жұмыс істейтін, бір уақытта консенсусқа қолдау көрсететін, сондай-ақ зиянды түйіндерді анықтап, алып тастайтын жүйе болып шығады.

Қазіргі сенім модельдері негізінен тікелей және жанама сенім модельдеріне бөлінеді:

Тікелей сенім моделі – түйіннің зиянды екенін тек оның мінез-құлқына сүйене отырып анықтайды. Бірақ бұл тәсіл әлсіз, өйткені сенімсіз модельдерді стратегия арқылы айналып өтуге мүмкіндік береді.

Жаңа PBFT негізінде құрылған алгоритмінің жүйелік орындалу процесі 3-суретте көрсетілген.



Сурет 3 - PBFT негізінде құрылған жаңа алгоритмнің жүйелік орындалу процесі

Алгоритмде түйіндерді сегментациялау бірнеше маңызды кезеңдерде жүзеге асады, бұл транзакцияларды параллель өңдеуді және жүктемені тиімді бөлуді қамтамасыз етеді.

Seg-PBFT Консенсус Алгоритмінің 9 Кезеңі (сегменттелген PBFT)

1-кезең: Сұраныс жіберу кезеңі (Request Phase)

Клиент өзінің сегментіндегі локалды прокси-түйінге сұраныс жібереді.

Бұл сұраныс ішкі сегмент ішінде талқыланып, консенсусқа келу үшін тіркеледі.

Клиент → Прокси

$$R_i = Request(T_x, C_i) \tag{3}$$

мұндағы T_x — транзакция, C_i — клиент, R_i — сұраныс.

2-кезең: Алдын ала дайындау кезеңі (Pre-Prepare Phase)

Прокси-түйін жаңа блок құрастырады.

Ол блокты сол сегменттегі басқа түйіндерге таратады.

Прокси → Сегменттегі түйіндер

$$B_j = CreateBlock(R_i) \tag{4}$$

және оны сегменттегі түйіндерге таратады:

$$\forall N_k \in S_j, \quad Send(B_j, N_k) \tag{5}$$

3-кезең: Дайындық – 1 (Prepare1 Phase)

Сегменттегі түйіндер бұл блокты тексереді.

Тексеруден кейін олар BLS мультиподпись (біріккен қолтаңба) жасап, оны прокси-түйінге қайтарады.

Түйіндер → Прокси

Әр түйін тексереді және қол қояды:

$$\sigma_k = Sign_{sk_k}(B_j)(_k, P_j) \tag{6}$$

4-кезең: Дайындық – 2 (Prepare2 Phase)

Прокси-түйін барлық қолтаңбаларды жинайды. 2f+1 түйіннен қол қойылғанда, ол оларды бір мультиподпиське біріктіріп, бүкіл сегментке таратады.

Прокси → Сегмент

Прокси 2f+1 қол жинап мультиподпись жасайды:

$$\Sigma_{prep} = \mathrm{Aggregate}(\{\sigma_1, \sigma_2, \ldots, \sigma_{2f+1}\})(_, S_j) \tag{7}$$

5-кезең: Коммит – 1 (Commit1 Phase)

Түйіндер келген мультиподписьті тексереді. Егер қол қоюшылар саны ≥ 2f+1 болса — олар блокты тексереді және өз қолтаңбаларын береді.

Прокси → Сегмент

Әр түйін мультиподписьті тексереді:

$Verify(\Sigma_{prep})$

Содан кейін қол қояды:

$$\gamma_k = Sign_{sk_k}(B_j) \tag{8}$$

6-кезең: Коммит – 2 (Commit2 Phase)

Прокси-түйін тағы да 2f+1 қолтаңбаны жинап, бір BLS мультиподпиське біріктіреді. Бұл блокты барлық сегмент түйіндеріне жіберіп, соңғы растау үшін таратады.

Түйіндер → Клиент

Әр түйін клиентке жауап жібереді:

$$\Sigma_{commit} = Aggregate(\{\gamma_1, \ldots, \gamma_{2f+1}\}) \tag{9}$$

(_, S_j)

7-кезең: Жауап кезеңі (Response Phase)

Түйіндер клиентке жауап жібереді.

Клиент f+1 растаманы алса, консенсусқа жетті деген сөз.

Түйіндер → Клиент

Әр түйін клиентке жауап жібереді:

$Send(ACK_k, C_i)$

Клиент тексереді:

$$\#\{ACK\} \geq f + 1 \Rightarrow Consensus\ Achieved \tag{10}$$

8-кезең: Кері байланыс кезеңі (Feedback Phase)

Клиент барлық жауаптарды бас басқарушы түйінге жібереді.

Басқарушы түйін әр сегменттен алынған мәліметтерді тексеріп, сараптайды.

Клиент → Бас түйін

Клиент жауаптарды жібереді:

$$Send(\{ACK_k\}, L) \tag{11}$$

мұндағы L — бас басқарушы түйін.

9-кезең: Біріктіру және тарату кезеңі (Merge & Propagation Phase)

Сегменттік прокси-түйіндер локалды блоктарды жоғары деңгейдегі глобалды прокси-түйінге жібереді.

Прокси-сегменттер → Глобалды прокси

Глобал түйін бұл блоктарды біріктіріп, глобал блок жасап, жүйеге таратады.

Сегменттік блоктар:

$\{B_1, B_2, \ldots, B_m\}$

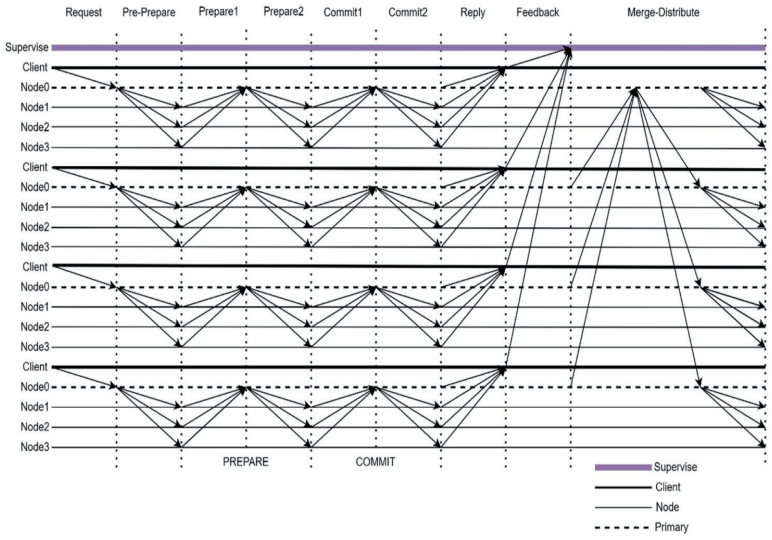
Глобалды блок құралады:

$$B_G = Merge(\{B_1, \ldots, B_m\}) \tag{12}$$

(B_G,)

Түйіндер сегментацияланады тікелей сұраныс кезеңінде – клиент әрдайым тек өз сегментімен әрекеттеседі. Консенсустың барлық алдын ала фазалары (Pre-Prepare, Prepare1, Prepare2, Commit1, Commit2) тек ішкі сегмент ішінде орындалады, глобалды түйіндермен өзара әрекеттесу болмайды. Тек соңғы кезеңде (Merge & Propagation) ғана локалды блоктар біріктіріліп, глобалды блок жасалады.

PBFT негізіндегі жаңа алгоритмі түйіндерді осылайша сегменттейді: консенсустың көп бөлігі локалды түрде орындалады, бұл желіге түсетін жүктемені азайтады. Тек соңғы кезеңде сегменттер өзара әрекеттесіп, бірыңғай глобалды блокты құрайды. Бұл тәсіл дәстүрлі PBFT алгоритмімен салыстырғанда масштабталуын, жылдамдығын және қауіпсіздігін жақсартады.



Сурет 4 - Алгоритмінің консенсус протоколымен өзара әрекеттесу үдерісі

4 - суретте көрсетілгендей консенсусқа жету процесі үш кезеңнен тұрады: pre−prepare, prepare және commit. Клиенттен сұраныс алғаннан кейін, негізгі түйін (яғни, Replica 0) pre−prepare хабарламасын басқа түйіндерге таратады. Prepare және commit кезеңдерінде әрбір реплика алынған хабарламалардың дұрыстығын тексеру үшін хабарламалар жібереді. Келесі кезеңге өту үшін қажетті минималды санды келісілген хабарламалар болуы керек.

PBFT негізіндегі жаңа консенсус механизмімен деректерді беру процесі басқа консенсус механизмдерінен айырмашылығы лидерлер әрбір резервтік көшірмеден транзакцияларды алған кезде консенсус терезесінің кезеңін динамикалық түрде реттейді. Лидер келесі кезеңнің уақытын ағымдағы кезеңде алынған транзакциялардың санына байланысты есептейді және оны

әрбір резервтік көшірмеге қайтарылған ақпаратпен бірге жібереді. Егер лидер бір уақытта көбірек транзакция алса, келесі кезеңнің уақыты қысқарады, сондықтан түйіннен жіберілген транзакцияларды жылдам өңдеуге болады. Ал керісінше, егер транзакциялар аз болса, келесі кезеңнің уақыты арттырылады, бұл желілік ресурстарды үнемдеуге мүмкіндік береді.

Төменде консенсус алгоритміндегі сұраныстарды өңдеудің негізгі кезеңдерінің бағдарламалық жүзеге асырылуы көрсетілген. Код Python тілінде жазылған және негізгі кезеңдерді қамтиды: сұраныс беру, блокты дайындау, қол қою және финализациялау.

```
def submit_request(self, client, data):
self.request_count += 1
self.requests[self.request_count] = {
«client»: client,
«data»: data,
«finalized»: False}
print(f»Request {self.request_count} submitted by {client}»)
return self.request_count

def prepare_block(self, request_id):
   if request_id not in self.requests:
raise ValueError(«Invalid request ID»)
print(f»Block prepared for request {request_id}»)

def sign_block(self, request_id, signer, signature):
if request_id not in self.signatures:
self.signatures[request_id] = []
self.signatures[request_id].append({«signer»: signer, «signature»: signature})

def commit_block(self, request_id):
if request_id not in self.signatures or len(self.signatures[request_id]) < 2:
raise ValueError(«Not enough signatures to commit»)
self.requests[request_id][«finalized»] = True
print(f»Request {request_id} committed successfully»)

def get_request(self, request_id):
   return self.requests.get(request_id, «Request not found»)
```

Ұсынылған модельдің тиімділігін бағалау

Бағалауды масштабталу параметрлері арқылы жүргіземіз

- Консенсус кідірісі (Latency) – транзакцияның расталуы үшін қажетті уақыт.

- Транзакция жылдамдығы (TPS) – блокчейн секундына қанша транзакцияны өңдей алады.

1.Консенсус кідірісі (Consensus Latency, мс)

Консенсус кідірісін келесі формула арқылы анықталады:

$$L_{\text{consensus}} = T_{\text{propose}} + T_{\text{network}} + T_{\text{verify}} + T_{\text{commit}} \quad (13)$$

Мұндағы:
$L_{consensus}$ - жалпы консенсус кідірісі (мс)
$T_{propose}$ - блок ұсынылу уақыты (лидер немесе прокси-түйін жаңа блокты ұсынады)
$T_{network}$ - желіде хабарламаларды тарату уақыты (latency)
T_{verify} - түйіндердің блоктағы транзакцияларды тексеру уақыты
T_{commit} - блоктың бекітілу уақыты (коммит фазасы)

Егер жүйе сегменттелетін болса, әрбір сегменттің кідірісі жеке есептеліп, жалпы жүйенің кідірісі ең баяу сегменттің кідірісіне тең:

$$L_{\text{total}} = \max(L_1, L_2, \dots, L_n) \quad (14)$$

мұнда L_i— i-ші сегменттің консенсус кідірісі.

Тестілеу: Әдетте желі топологиясын жобалау Cisco Packet Tracer немесе GNS3 көмегімен орындалады. Біз GNS3 қолданамыз ол қажетті сұлбаларды немесе шешімдерді модельдеуге, конфигурацияны нақты жабдықта қолданбас бұрын тексеруге болатын толыққанды зертханалық құрал.

Тестілеу барысында жүйеге 1000 транзакция жіберілді. Блоктың ұсынылған уақыты (Tpropose) және соңғы бекітілген блоктың уақыты (Tcommit) тіркелді.

Блок деңгейінде:

$$L_{\text{block}} = T_{\text{commit}} - T_{\text{propose}} \quad (15)$$

Орташа мәні:

$$L_{\text{avg}} = \frac{\sum L_{\text{block}}}{N} \quad (25)$$

мұндағы N — жалпы блок саны.

Бір блоктың ұсынылу уақыты 48 мс, желідегі кешігу 32 мс, тексеру уақыты 21 мс, ал бекіту уақыты 46 мс болды

$$L_{\text{consensus}} = 48 + 32 + 21 + 46 = 147\ mc \quad (16)$$

Бұл блокчейннің орташа консенсус кідірісі 147 мс екенін көрсетті.

PBFT және оның туындылары негізінде жасалған консенсус алгоритмдердің көрсеткіштерін әдебиеттік зерттеу жұмыстарынан алынып, PBFT негізіндегі жаңа консенсус алгоритммен салыстырмалы түрде қарастырылады. Бұл салыстыру жүйенің ақауға төзімділігі, масштабталуы, уақытша кідіріс және байланыс күрделілігі сияқты негізгі параметрлерге сүйенеді.

Кесте 3 - PBFT және оның туындылары негізінде жасалған консенсус алгоритмдерінің салыстырмалы сипаттамасы

Алгоритм	Ақауға төзімділік	Масштабталуы	Кідіріс	Байланыс күрделілігі
IBFT	$3f + 1$	Төмен	Төмен	$O(n^2)$
RBFT	$2f + 1$	Орташа	Төмен	$O(n^4)$
PoET	–	Жақсы	Орташа	$O(n^3)$
PBFT	$3f + 1$	Орташа	200–500 мс	$O(n^2)$ (шамамен)
PBFT негізіндегі жаңа алгоритм	$2f + 1$	Жақсы	80–147 мс	$O(n)$–$O(n^2)$ (оңтайландырылған)

Салыстырма үшін IBFT алгоритмі ақауға төзімділік жағынан PBFT-ке ұқсас, бірақ масштабталу мүмкіндігі шектеулі, сондықтан тек кішігірім желілер үшін тиімді. RBFT байланыс күрделілігін төмендетуге бағытталғанымен, оның есептеу ресурстарына қойылатын талаптары жоғары ($O(n^4)$), бұл үлкен желілерде қолдануға кедергі келтіреді. PoET алгоритмі рұқсат етілмеген (permissionless) жүйелерге бағытталған және масштабталу мүмкіндігі жоғары болғанымен, қауіпсіздігі сенімге негізделген.

PBFT алгоритмі орташа кідіріс уақыты мен күрделілігіне ие және көптеген консорциумдық блокчейндерде әлі де кеңінен қолданылады. Алайда, оның байланыс күрделілігі $O(n^2)$ деңгейінде болуы үлкен желілерде өнімділікті төмендетеді.

Ұсынылып отырған жаңа алгоритм осы мәселелерді шешуге бағытталған. Ол желіні сегменттерге бөлу арқылы көшбасшыны ауыстыру механизмін оңтайландырады. Бұл тәсіл кідірісті 80–147 миллисекунд аралығына дейін қысқартып, консенсус процесінің өнімділігін арттырады. Сонымен қатар, байланыс күрделілігі дәстүрлі PBFT-пен салыстырғанда төмен ($O(n)$–$O(n^2)$), бұл оны масштабталатын жүйелер үшін қолайлы етеді. Ақауға төзімділік те оңтайлы деңгейде сақталған — тек $2f + 1$ түйін жеткілікті, бұл жүйенің ресурстық тиімділігін арттырады.

Кесте 4 - PBFT және PBFT негізіндегі жаңа алгоритмдерінің салыстырмалы талдауы

Параметр	PBFT	PBFT негізіндегі жаңа алгоритм	Айырмашылығы
Ақауға төзімділік	$3f + 1$	$2f + 1$	33% жеңілдік — азырақ түйінмен консенсусқа жетеді
Кідіріс	200–500 мс	80–147 мс	Орта есеппен 62–70% жылдамырақ

Масштабталуы	20–100 түйін	180+ түйін)	>100% жоғары — екі есе көп түйінді қолдай алады
Байланыс күрделілігі	$O(n^2)$	$O(n)$–$O(n^2)$ (оңтайландырылған)	Күрделілік төмендеген – тиімділік жақсарған

Жоғарыдағы салыстыру PBFT және оның жетілдірілген нұсқасы PBFT негізінде құрылған жаңа алгоритмнің арасындағы маңызды айырмашылықтарды көрсетеді. Бұл алгоритм көшбасшыны ауыстыру және түйіндер құрылымын сегменттеу арқылы бірнеше көрсеткіш бойынша айтарлықтай артықшылыққа ие:

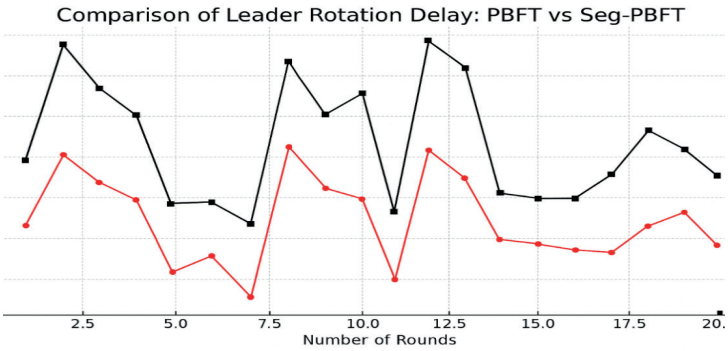
- Ақауға төзімділікте $3f + 1$ орнына $2f + 1$ қолдану жүйенің қажетті түйін санын 33%-ға азайтып, ресурстық тиімділікті арттырды.

- Кідіріс уақыты бойынша орта есеппен 2.5–3 есе жылдам жұмыс істейді (500 мс орнына 147 мс).

- Масштабталуы бойынша PBFT негізіндегі жаңа алгоритм 180 және одан да көп түйінмен тиімді жұмыс істей алады, бұл PBFT-пен салыстырғанда екі есе көп.

- Байланыс күрделілігі $O(n^2)$-ден $O(n)$–$O(n^2)$-ге оңтайландырылып, желі трафигі мен жүктемесі азаяды, әсіресе үлкен желілерде.

PBFT және PBFT негізіндегі жаңа алгоритм алгоритмдері бойынша лидерді ауыстыруға кеткен уақыттың салыстырмалы графигі Сурет 5 ұсынылған.



Сурет 5 - PBFT пен PBFT негізінде құрылған жаңа алгоритмдерінің консенсус кідірісін салыстырмалы талдау графигі

Горизонталь ось – эксперименттік раундтардың саны (жалпы 20 топ), ал вертикаль ось – лидерді ауыстыруға кеткен уақыт (миллисекундпен, мс).

Қорытынды. Бұл зерттеуде консенсус алгоритмдерінің ішінде лидерді ауыстыру уақытына ерекше назар аударылды. PBFT және PBFT негізіндегі жаңа алгоритм алгоритмдері бойынша жүргізілген эксперименттік бағалау нәтижелері блокчейн жүйелерінің ақауға төзімділігі мен байланыс күрделілігіне қатысты маңызды мәліметтер береді.

Зерттеу барысында алгоритмдер Python тілінде симуляцияланып,

нәтижелер Matplotlib кітапханасы арқылы визуалданды. Бұл тәсіл желілік протоколдардың жұмысын нақты көрсетіп, оңтайландыру мүмкіндіктерін анықтауға мүмкіндік берді (Мельников et al, 2009). PBFT алгоритмінің негізінде сипатталған жаңа алгоритм бір үлгіге орналастырып, нақты салыстыру жүргіздік. Авторлар екі алгоритмді де 20 топта сынап көріп, уақытша кідірістерді есептеді, және эксперимент нәтижелері 5- суретте көрсетілген.

Эксперимент нәтижелері көрсеткендей жаңа алгоритм жүйесіндегі көшбасшыны басқару механизмі PBFT алгоритміне қарағанда орындау уақыты бойынша 62-70% тиімдірек. PBFT негізінде құрылған жаңа алгоритм негізгі түйінді көшбасшының тоқтап қалуынан кейін жылдамырақ ауыстыра алады және консенсусқа қол жеткізу процесіне әсерін азайтады. Осылайша бұл алгоритм шешімдердің эволюциясында жаңа қадам болып табылады. Ол масштабталу, жылдамдық және қауіпсіздік тұрғысынан теңгерімді қамтамасыз етіп, заманауи блокчейн жүйелерінің талаптарына жақсырақ жауап береді.

Әдебиеттер

Bonneau J., Miller A., Clark J., Narayanan A., Kroll J.A., Felten E.W. SoK. (2015) Research Perspectives and Challenges for Bitcoin and Cryptocurrencies. 2015 IEEE Symposium on Security and Privacy. — P. 104–121. doi: 10.1109/SP.2015.14.

Buchaesi M., Кальзавара С., Фокарди Р. (2017) Формальные методы веб-безопасности. Журнал логических и алгебраических методов в программировании. — Т. 87. — С. 110–126.

Huynh T.T., Huynh T.D., Nguyen H., Tan H.A (2019) Survey on security and privacy issues of blockchain technology. Proc. 2019 IEEE Int. Conf. Syst. Sci. Eng. — P. 362–367.

Leng J., Zhou M., Zhao J.L., Huang Y., Bian Y. (2022) Blockchain security: A survey of techniques and research directions. IEEE Trans. Serv. Comput. — Vol. 15, No. 4. — P. 2490–2510.

Li X., Jiang P., Chen T., Luo X., Wen Q.A. (2020) Survey on the security of blockchain systems. Future Gener. Comput. Syst., Vol. 107. — P. 841–853.

Popadyuk A.Yu., Korovyakovsky E.K., Titova T.S. (2020) Environmental Aspects of Distributed Ledger Technology: A Case Study of the Proof-of-Work Consensus Algorithm. Proc. of the St. Petersburg State University of Railway Communications. — T. 17. — №1. — P. 136–143.

Sompolinsky Y., Lewenberg Y., Zohar A. (2016) Inclusive Block Chain Protocols. International Conference on Financial Cryptography and Data Security. Springer, — P. 528–547. DOI: 10.1007/978-3-662-53357-4_26.

Somin S., Gordon G., Altshuler Y. (2018) Network analysis of ERC20 tokens trading on Ethereum blockchain. In: Unifying Themes in Complex Systems IX. Cham: Springer, — P. 439–450.

Абдулжалилов А.З. (2023) Методы и стратегии масштабируемости блокчейн-технологий: анализ, сравнение и перспективы, Вестник науки. — №11(68). — Т 4. — С. 625–634.

Баракова А.Ш., Усатова О.А. (2022) Веб-ресурстардың қауіпсіздігінің осалдықтары мен қауіптерін жіктеу және сипаттау Материалы VII международной научно-практической конференции «Информатика и прикладная математика». Алматы. — С. 364–368.

Вуколич М.В (2015) Поисках масштабируемой блокчейн-сети: доказательство выполнения работы против репликации BFT, Материалы Международного семинара по открытым проблемам сетевой безопасности. — С. 112–125.

Дипа Г., Тилагам П. С. (2016) Защита веб–приложений от уязвимостей, связанных с внедрением и логикой: подходы и проблемы информационные *и программные технологии.* — Т. 74. — С. 160–180.

Усатова О.А., Баракова А.Ш. (2022) Қазіргі заманғы веб-ресурстарды қорғау жүйелерін

талдау. ҚР ҰҒА Хабарлары. Физика және информатика сериясы, №1 (341). — Б. 88–95. *https://doi.org/10.32014/2022.2518-1726.120*

Мерсер Д. (2018) Создание надёжных и полнофункциональных веб-сайтов, блогов, форумов, порталов и сайтов-сообществ. — М.: Вильямс. — С. 272–274.

Мельников В.П., Клейменов С.А., Петраков А.М. (2009) Информационная безопасность и защита информации: учебник. — С. 184–186.

References

Bonneau J., Miller A., Clark J., Narayanan A., Kroll J.A., Felten E.W. SoK. (2015) Research Perspectives and Challenges for Bitcoin and Cryptocurrencies. 2015 IEEE Symposium on Security and Privacy. — P. 104–121. doi: 10.1109/SP.2015.14. (in Eng)

Buchaezi M., Kalzavara S., Fokardi R. (2017) Formal'nye metody veb-bezopasnosti [Formal Methods of Web Security]. Journal of Logical and Algebraic Methods in Programming. — T. 87. — P. 110–126. (in Rus)

Huynh T.T., Huynh T.D., Nguyen H., Tan H.A (2019). survey on security and privacy issues of blockchain technology. Proc. 2019 IEEE Int. Conf. Syst. Sci. Eng. — P. 362–367. (in Eng)

Leng J., Zhou M., Zhao J.L., Huang Y., Bian Y. (2022). Blockchain security: A survey of techniques and research directions. IEEE Trans. Serv. Comput. — Vol. 15. — No. 4. — P. 2490–2510. (in Eng)

Li X., Jiang P., Chen T., Luo X., Wen Q.A(2020) Survey on the security of blockchain systems. Future Gener. Comput. Syst., Vol. 107. — P. 841–853. (in Eng)

Popadyuk A.Yu., Korovyakovsky E.K., Titova T.S. (2020) Environmental Aspects of Distributed Ledger Technology: A Case Study of the Proof-of-Work Consensus Algorithm. Proc. of the St. Petersburg State University of Railway Communications. — T. 17. — №1. — P. 136–143. (in Eng)

Sompolinsky Y., Lewenberg Y., Zohar A. (2016) Inclusive Block Chain Protocols. International Conference on Financial Cryptography and Data Security. Springer. — P. 528–547. DOI: 10.1007/978-3-662-53357-4_26. (in Eng)

Somin S., Gordon G., Altshuler Y. (2018) Network analysis of ERC20 tokens trading on Ethereum blockchain. In: Unifying Themes in Complex Systems IX. Cham: Springer. — P. 439–450. (in Eng)

Abdulzhalilov A.Z. (2023) Metody i strategii masshtabiruemosti blokchein-tekhnologii: analiz, sravnenie i perspektivy [Methods and Strategies for Blockchain Scalability: Analysis, Comparison, and Prospects]. Vestnik nauki. — №11(68). — T 4. — P. 625–634. (in Rus)

Barakova A.Sh., Usatova O.A. (2022) Veb-resurstardyn kaúipsizdiginiń osaldyqtary men kaúpterin zhikteu jáne sipattau [Classification and Description of Web Resource Security Vulnerabilities and Threats]. Materialy VII mezhdunarodnoi nauchno-prakticheskoi konferentsii «Informatika i prikladnaia matematika». Almaty. — P. 364–368. (in Rus)

Vukolich M.V. (2015) Poiskakh masshtabiruemoi blokchein-seti: dokazatel'stvo vypolneniia raboty protiv replikatsii BFT [In Search of a Scalable Blockchain Network: Proof-of-Work versus BFT Replication]. Materialy Mezhdunarodnogo seminara po otkrytym problemam setevoi bezopasnosti. — P. 112–125. (in Rus)

Dipa G., Tilagam P.S. (2016) Zashchita veb-prilozhenii ot uiazvimostei, sviazannykh s vnedreniem i logikoi: podkhody i problemy. Informatsionnye i programmnye tekhnologii [Protection of Web Applications from Injection and Logic-Related Vulnerabilities: Approaches and Challenges]. — T. 74. — P.160–180. (in Rus)

Usatova O.A., Barakova A.Sh. (2022) Qazirgi zamangy veb-resurstardy qorgau zhüielerin taldau [Analysis of Modern Web Resource Protection Systems]. Bulletin of the National Academy of Sciences of the Republic of Kazakhstan. Series of Physics and Informatics, №1 (341). — P. 88–95. https://doi.org/10.32014/2022.2518-1726.120 (in Kaz)

Mercer D. (2018) Sozdanie nadezhnykh i polnofunktsional'nykh veb-saitov, blogov, forumov, portalov i saitov–soobshchestv [Creating Reliable and Fully Functional Websites, Blogs, Forums, Portals, and Community Sites]. — M.: Vil'iams, — P. 272–274. (in Rus)

Melnikov V.P., Kleimenov S.A., Petrakov A.M. (2009) Informatsionnaia bezopasnost' i zashchita informatsii [Information Security and Information Protection]. Textbook. — P. 184–186. (in Rus)

ACADEMIC SCIENTIFIC JOURNAL OF COMPUTER SCIENCE
ISSN 1991-346X
Volume 4. Namber 356 (2025). 122–137

https://doi.org/10.32014/2025.2518-1726.387

IRSTI 27.35.33
UDC 004.89; 004.931

© A.N. Zhidebayeva[1], G.U. Madaliyeva[1], B.O. Tastanbekova[2], S.S. Karzhaubekova[2], G.S. Shaimerdenova[2*], 2025.
[1] Academician A. Kuatbekov University of peoples' friendship, Shymkent, Kazakhstan;
[2]Mukhtar Auezov South Kazakhstan University, Shymkent, Kazakhstan.
E-mail: danel101kz@gmail.com

DEEP NEURAL NETWORK CONV-LSTM FOR ECG-BASED CARDIAC DISORDER IDENTIFICATION

Zhidebayeva Aziza — Candidate of Technical Sciences, Senior Lecturer, department of Computer science and Mathematics, Peoples Friendship University named after Academician A. Kuatbekov, Shymkent , Kazakhstan,
E-mail: aziza_68.kz@mail.ru , https://orcid.org/0000-0002-3768-4835;
Madaliyeva Gulnar — Senior Lecturer department of Computer science and Mathematics, Peoples Friendship University named after Academician A. Kuatbekov, Shymkent, Kazakhstan,
E-mail:madaliyeva.g@mail.ru, https://orcid.org/0000-0002-37883856;
Tastanbekova Bayan — master, teacher, department of Information Communication technologies, Mukhtar Auezov South Kazakhstan University, Shymkent, Kazakhstan,
E-mail: baya2013@inbox.ru, https://orcid.org/ 0000-0002-4549-0344;
Karzhaubekova Samal — master, teacher, department of Information Communication technologies, Mukhtar Auezov South Kazakhstan University, Shymkent, Kazakhstan,
E-mail: 120989samal@mail.ru, https://orcid.org/ 0009-0000-0794-4221;
Shaimerdenova Guldana — Phd, associate professor, department of Information Communication technologies, Mukhtar Auezov South Kazakhstan University, Shymkent, Kazakhstan,
E-mail: danel101kz@gmail.com , https://orcid.org/0000-0001-8685-7125.

Abstract. In the emerging field of health diagnostics, the electrocardiogram (ECG) serves as a primary tool for assessing cardiac function. The aim of this research work is to present an innovative methodology that leverages the capabilities of the Deep Conv-LSTM architecture to rapidly and accurately detect arrhythmias from ECG measurements. Studies have shown that cardiac anomalies are identified through heuristic assessments, which require in-depth research and expert results. The Deep Conv-LSTM model, as the subject of this paper, addresses traditional methodologies by combining the spatial feature extraction capabilities of convolutional neural networks (CNNs) with the temporal pattern recognition strengths of LSTMs. Initial results from a comprehensive dataset rich in ECG

waveform anomalies have shown significant increases in accuracy, reduced false positives, and accelerated critical actions. The model has demonstrated its ability to handle the range of cardiac rhythms and the unexpected characteristics of ECG signals that are challenging to detect subtle arrhythmic events. Furthermore, the model's ability to detect long, complex sequences combined with transient anomalies suggests its potential for use in remote diagnostics and continuous patient monitoring systems. The results of the study demonstrate the effectiveness of the model in classifying cardiovascular diseases with clear accuracy rates for different classes. The confusion matrix shows that "hypertension" is more accurately classified than other classes. Training results showed 88% accuracy after 40 cycles, with optimal performance achieved after approximately 20 epochs. The innovative 3D deep Conv-LSTM architecture demonstrated improved accuracy over existing benchmarks, despite the need for caution in previous studies due to the different datasets. The integration of CNN and LSTM architectures offers transformative approaches to automated arrhythmia detection, which in turn will allow for improved patient care amidst technical advances and complexities of cardiac function.

Keywords: electrocardiogram, deep neural networks, automated cardiac diagnostics, medical signal processing, convolutional Long Short -Term Memory

© А.Н. Жидебаева[1], Г.У. Мадалиева[1], Б.О. Тастанбекова[2], С.С. Қаржаубекова[2], Г.С. Шаймерденова[2*], 2025.

[1]Академик Ә. Қуатбеков атындағы Халықтар достығы университеті, Шымкент, Қазақстан;

[2]М. Ауезов атындағы Оңтүстік Қазақстан университеті, Шымкент, Қазақстан.

E-mail: danel101kz@gmail.com

ЖҮРЕК АУРУЛАРЫН АНЫҚТАУДА CONV-LSTM АРХИТЕКТУРАСЫНА НЕГІЗДЕЛГЕН ТЕРЕҢ НЕЙРОНДЫҚ ЖЕЛІ

Жидебаева Азиза — техника ғылымдарының кандидаты, «Информатика және математика» кафедрасының аға оқытушысы, Академик А. Қуатбеков атындағы Халықтар достығы университеті, Шымкент, Қазақстан,
E-mail: aziza_68.kz@mail.ru, https://orcid.org/0000-0002-3768-4835;
Мадалиева Гульнар — магистр, «Информатика және математика» кафедрасының аға оқытушысы, Академик А. Қуатбеков атындағы Халықтар достығы университеті, Шымкент, Қазақстан,
E-mail:madaliyeva.g@mail.ru, https://orcid.org/0000-0002-37883856;
Тастанбекова Баян — магистр, Ақпараттық коммуникациялық технологиялар кафедрасының оқытушысы, Мұхтар Әуезов атындағы Оңтүстік Қазақстан университеті, Шымкент, Қазақстан,
E-mail: baya2013@inbox.ru, https://orcid.org/ 0000-0002-4549-0344;
Қаржаубекова Самал — магистр, Ақпараттық коммуникациялық технологиялар кафедрасының оқытушысы, Мұхтар Әуезов атындағы Оңтүстік Қазақстан университеті, Шымкент, Қазақстан,
E-mail: 120989samal@mail.ru, https://orcid.org/ 0009-0000-0794-4221;

Шаймерденова Гульдана — PhD, Ақпараттық коммуникациялық технологиялар кафедрасының доценті, Мұхтар Әуезов атындағы Оңтүстік Қазақстан университеті, Шымкент, Қазақстан,
E-mail: danel101kz@gmail.com, https://orcid.org/0000-0001-8685-7125.

Аннотация. Денсаулық диагностикасының дамып келе жатқан саласында электрокардиограмма (ЭКГ) жүрек қызметін бағалаудың негізгі құралы ретінде қызмет етеді. Бұл зерттеу жұмысының мақсаты ЭКГ өлшеу кезінде аритмияны жылдам және дәл анықтау үшін Терең Conv-LSTM архитектурасының мүмкіндіктерін пайдаланатын инновациялық әдістемені ұсыну болып табылады. Зерттеулер көрсеткендей, жүрек аномалиялары терең зерттеулер мен сараптамалық нәтижелерді қажет ететін эвристикалық бағалау арқылы анықталады. Deep Conv-LSTM моделі, осы мақаланың тақырыбы ретінде, конволюциялық нейрондық желілердің (Cnn) кеңістіктік ерекшеліктерін алу мүмкіндіктерін Lstm үлгілерін уақытша танудың күшті жақтарымен біріктіру арқылы дәстүрлі әдістемелерді қарастырады. ЭКГ толқын пішінінің ауытқуларына бай кешенді деректер жиынтығының бастапқы нәтижелері дәлдіктің айтарлықтай жоғарылауын, жалған позитивтердің төмендеуін және сыни әрекеттердің жеделдеуін көрсетті. Модель жүрек ритағының диапазонын және нәзік аритмиялық құбылыстарды анықтау қиын БОЛАТЫН ЭКГ сигналдарының күтпеген сипаттамаларын өңдеу қабілетін дәлелдеді. Сонымен қатар, модельдің өтпелі ауытқулармен біріктірілген ұзақ және күрделі тізбектерді анықтау қабілеті оның пациенттерді қашықтықтан диагностикалау және үздіксіз бақылау жүйелерінде қолдану әлеуетін арттырады. Зерттеу нәтижелері модельдің жүрек-қан тамырлары ауруларын жіктеудегі тиімділігін әр түрлі сыныптар үшін нақты дәлдік көрсеткіштерімен көрсетеді. Шатасу матрицасы "гипертония" басқа кластарға қарағанда дәлірек жіктелді. Оқу нәтижелері 40 циклден кейін 88% дәлдікті көрсетті, оңтайлы өнімділікке шамамен 20 дәуірден кейін қол жеткізілді. Deep Conv-LSTM инновациялық 3d архитектурасы әртүрлі деректер жиынына байланысты алдыңғы зерттеулерде сақтық таныту қажеттілігіне қарамастан, қолданыстағы эталондармен салыстырғанда дәлдіктің жақсарғанын байқаймыз. CNN және LSTM архитектураларын біріктіру аритмияны автоматтандырылған анықтаудың трансформациялық тәсілдерін ұсынады, бұл өз кезегінде техникалық жетістіктер мен жүрек қызметінің күрделілігі жағдайында пациенттерге күтімді жақсартуға мүмкіндік береді.

Түйін сөздер: электрокардиограмма, терең нейрондық желілер, автоматтандырылған жүрек диагностикасы, медициналық сигналдарды өңдеу, конволюциялық ұзақ қысқа мерзімді жады

© А.Н. Жидебаева[1], Г.У. Мадалиева[1], Б.О. Тастанбекова[2], С.С. Каржаубекова[2] Г.С. Шаймерденова[2*], 2025.

[1] Университет дружбы Народов им.академика А. Куатбекова, Шымкент, Казахстан

[2] Южно-казахстанский университет имени Мухтара Ауэзова, Шымкент, Казахстан

E-mail: danel101kz@gmail.com

ГЛУБОКАЯ НЕЙРОННАЯ СЕТЬ НА ОСНОВЕ АРХИТЕКТУРЫ CONV-LSTM ДЛЯ ВЫЯВЛЕНИЯ СЕРДЕЧНЫХ ЗАБОЛЕВАНИЙ

Zhidebayeva Aziza — кандидат технических наук, старший преподаватель кафедры Информатики и математики, Университет Дружбы Народов имени академика А. Куатбекова, Шымкент, Казахстан,
E-mail: aziza_68.kz@mail.ru , https://orcid.org/0000-0002-3768-4835;
Madaliyeva Gulnar — магистр, старший преподаватель кафедры Информатики и математики, Университет Дружбы Народов имени академика А. Куатбекова, Шымкент, Казахстан,
E-mail:madaliyeva.g@mail.ru, https://orcid.org/0000-0002-37883856;
Tastanbekova Bayan — магистр, преподаватель кафедры Информационно-коммуникационных технологий, Южно-Казахстанский университет имени Мухтара Ауэзова, Шымкент, Казахстан,
E-mail: baya2013@inbox.ru , https://orcid.org/ 0000-0002-4549-0344;
Karzhaubekova Samal — магистр, преподаватель кафедры Информационно-коммуникационных технологий, Южно-Казахстанский университет имени Мухтара Ауэзова, Шымкент, Казахстан,
E-mail: 120989samal@mail.ru , https://orcid.org/ 0009-0000-0794-4221;
Shaimerdenova Guldana — Phd, доцент кафедры Информационно-коммуникационных технологий Южно-Казахстанского университета имени Мухтара Ауэзова, Шымкент, Казахстан,
E-mail: danel101kz@gmail.com , https://orcid.org/0000-0001-8685-7125.

Аннотация. В развивающейся области медицинской диагностики электрокардиограмма (ЭКГ) служит основным инструментом оценки функции сердца. Цель данной исследовательской работы – представить инновационную методологию, использующую возможности архитектуры Deep Conv-LSTM для быстрого и точного выявления аритмий по данным ЭКГ. Исследования показали, что сердечные аномалии выявляются с помощью эвристических оценок, требующих глубоких исследований и экспертных результатов. Модель Deep Conv-LSTM, являющаяся предметом данной статьи, обращается к традиционным методологиям, сочетая возможности извлечения пространственных признаков сверточных нейронных сетей (CNN) с преимуществами распознавания временных образов LSTM. Первые результаты, полученные на основе обширного набора данных, богатого аномалиями ЭКГ-сигналов, показали значительное повышение точности, снижение количества ложноположительных результатов и ускорение критических действий. Модель продемонстрировала свою способность обрабатывать широкий спектр сердечных ритмов и непредвиденные характеристики сигналов ЭКГ, которые затрудняют обнаружение малозаметных аритмических событий. Кроме того,

способность модели обнаруживать длинные и сложные последовательности в сочетании с кратковременными аномалиями свидетельствует о её потенциале для использования в системах дистанционной диагностики и непрерывного мониторинга пациентов. Результаты исследования демонстрируют эффективность модели в классификации сердечно-сосудистых заболеваний с чёткими показателями точности для различных классов. Матрица ошибок показывает, что «гипертония» классифицируется точнее, чем другие классы. Результаты обучения показали точность 88% после 40 циклов, а оптимальная производительность была достигнута примерно через 20 эпох. Инновационная архитектура 3D deep Conv-LSTM продемонстрировала более высокую точность по сравнению с существующими эталонными показателями, несмотря на необходимость соблюдения осторожности в предыдущих исследованиях из-за различий в наборах данных. Интеграция архитектур CNN и LSTM предлагает революционные подходы к автоматизированному выявлению аритмии, что, в свою очередь, позволит улучшить качество лечения пациентов в условиях технического прогресса и сложности функций сердца.

Ключевые слова: электрокардиограмма, глубокие нейронные сети, автоматизированная диагностика заболеваний сердца, обработка медицинских сигналов, сверточная долговременная кратковременная память

Introduction. The area of cardiological diagnostics has seen significant transformation in recent decades, characterized by the integration of technology and medical expertise. Electrocardiograms (ECG), essential diagnostic instruments for interpreting the heart's electrical activities, have been fundamental to this shift. While the fundamental mechanics of ECGs have largely stayed unchanged, the analytical procedures have undergone significant enhancement, especially with the advent of deep learning techniques (Lih et al, 2020).

Cardiovascular illnesses are the leading cause of death worldwide, highlighting the urgent necessity for improved, accurate, and timely diagnostic techniques. Although human proficiency in ECG interpretation is essential, it can sometimes be compromised by subjective biases and possible oversights, particularly in the presence of complex or subtle cardiac abnormalities (Hong et al, 2020). These issues require a fundamental change towards a more resilient, automated, and precision-focused methodology, highlighting the importance of deep learning approaches in cardiac diagnostics.

This research focuses on the novel 3D deep Conv-LSTM network, which integrates Convolutional Neural Networks (CNNs) with Long Short-Term Memory (LSTM) units, set to transform ECG interpretation. Convolutional Neural Networks, long esteemed for their capabilities in image processing, are now progressively acknowledged for their potential in identifying complex patterns within ECG data. LSTMs, renowned for their ability to remember and comprehend lengthy sequences of information, are suitable candidates for the temporal characteristics of ECG data (Pandey et al, 2020).

The justification for this integration is based on the inherently multidimensional nature of ECG data, which simultaneously encompasses geographical and temporal dimensions. Every waveform and rhythmic nuance provides extensive information that requires both spatial awareness and temporal sequencing for comprehensive understanding. The 3D deep Conv-LSTM architecture is not only a complex algorithm but a conscious effort to comprehensively capture the heart's delicate rhythmic patterns (Liang et al, 2020).

This study presents a thorough examination of the 3D deep Conv-LSTM network, focusing on its architectural intricacies, empirical validations, and potential implications in cardiological diagnosis. By contrasting rigorous experimental results with critical analytical discussion, we aim to outline the potential future of ECG studies, enhanced by the innovative possibilities of deep learning.

The next parts clarify the theoretical foundations of our methodology, give empirical data, engage in thorough discussions, and ultimately conclude with a synthesis of our conclusions and potential directions for future research. Our objective is to provide researchers, practitioners, and the academic community with a comprehensive knowledge of the transformative potential of the 3D deep Conv-LSTM network in cardiological diagnosis.

Materials and methodology. Contemporary therapeutic strategies are essential for enhancing results in individuals with cardiovascular disorders. Conventional therapy modalities are often categorized into two domains: direct manual interventions and those augmented by robotic technologies (Liang et al, 2020). Nevertheless, both methods provide unique challenges. Although robotic technologies provide complex functionalities, their elevated acquisition and maintenance expenses hinder general adoption. Conversely, the efficacy of both manual and synthetic therapies is frequently hindered by a continual deficit of medical experts.

Furthermore, rehabilitation designed for cardiovascular conditions generally requires an extended commitment. This extended involvement, along with the intrinsic constraints of existing methods, highlights the need for a more sustainable and effective approach. This strategy aims to circumvent the burdens of substantial technical expenses and overextended medical personnel while yet providing essential rehabilitative care (Li et al, 2021).

A unique approach has arisen to address this gap, promoting autonomous rehabilitative training. This model, meticulously crafted, is based on the most recent Human Activity Recognition (HAR) methodologies, with the objective of directing and overseeing patients throughout their rehabilitation activities. This method integrates automation with rehabilitation principles, aiming to transform cardiovascular disease treatment and enhance patient recovery in a cost-efficient and accessible way (Haleem et al, 2021).

Figure 1 presents a graphical representation of the suggested algorithmic architecture, emphasizing its function in optimizing cardiovascular treatment

regimens. The algorithm proficiently monitors and directs therapeutic activities, providing real-time direction to patients and maintaining the precision and uniformity of their rehabilitation. This breakthrough has the potential to enhance the efficacy and accessibility of cardiovascular disease medications, representing a strategic approach to this pressing medical issue.

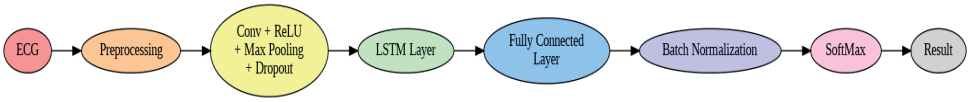


Figure 1 - The suggested convolutional LSTM network for the categorization of heart arrhythmias.

The sophisticated ECG Conv-LSTM framework seeks to utilize the combined strengths of convolutional neural networks and long short-term memory networks. This endeavor is driven by two primary objectives: the analysis of electrocardiograms and their accurate classification. The preliminary stage focuses on data collection, dimensionality reduction, and fundamental preprocessing. Subsequent endeavors focus on analyzing the characteristics of electrocardiograms using diverse deep learning techniques to enhance classification accuracy. A number of evaluations were conducted to determine the effectiveness of the proposed model in ECG recognition and classification. Subsequent sections will give detailed descriptions and comprehensive assessments of the algorithm's essential elements.

A. Convolutional neural network

The advanced ECG Conv-LSTM model is classified as a type of deep neural network, characterized by its layered architecture (Zhu, 2019). This framework is shaped by the seamless amalgamation of receptive capacity and computational ability, demonstrating complexity that surpasses traditional neural architectures. Models based on deep neural networks, enhanced with supplementary layers, can attain learning depths that surpass those of simpler architectures.

Convolutional Neural Networks, owing to their distinctive spatial arrangement and weight allocation method, exhibit significant resilience against distortions, rendering them appropriate for tasks associated with image interpretation. Convolutional Neural Networks (CNNs) intrinsically utilize a weight-sharing approach, optimizing the model's architecture while improving performance and effectively regulating the weight quantity. Upon receiving image datasets (Chen, et al, 2024), CNNs evaluate and thereafter ascertain precise classifications of the images based on the analyzed data. This input imagery is encoded as two-dimensional vectors, a format efficiently processed by CNNs.

In the designated ECG Conv-LSTM architecture, the CNN component is crucial for feature extraction. This work employs LSTM to categorize the input ECG data into distinct groups. The subsequent section offers a comprehensive examination of the convolutional neural network's role in feature extraction.

$$P = f\left(\sum_{i=1}^{N} Z_i \cdot W_i + B\right) \quad (1)$$

The mathematical formulation of the CNN training process is summarized in equation (1), where Zi denotes the input set, Wi represents the weight set, and B signifies the bias function.

B. Long Short-Term Memory Network

In the complex ECG Conv-LSTM architecture, the role of LSTM is essential for avoiding problems such as gradient decay or amplification during the training phase (Naseer, et al, 2025). The backpropagation (BP) method is utilized to modify the weights. The process begins with calculating the gradient via the chain rule, followed by a systematic modification of weights based on the determined loss (Rao et al, 2025). Backpropagation initiates at the output layer of the neural network. Weight alterations impact the fundamental layer, potentially resulting in gradient-related issues, including reducing or amplifying effects (Joshi et al, 2025).

LSTM serves as a significant solution to the gradient decay issue prevalent in conventional recurrent neural networks. Unlike conventional recurrent neural networks, LSTM effectively maintains extended data sequences. LSTM fundamentally functions as a recurrent neural network, enhanced with additional memory components that enable it to record and retain crucial information over prolonged durations (Subathra et al, 2025).

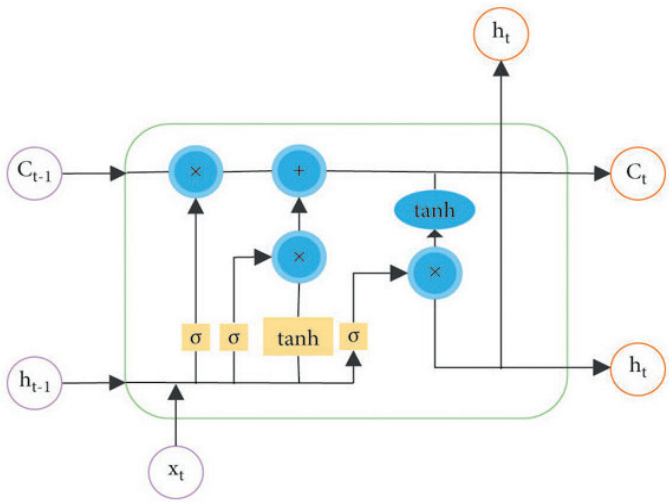


Figure 2 - LSTM part of the proposed convolutional LSTM architecture

The Long Short-Term Memory (LSTM) component in the proposed Convolutional LSTM network is crucial for improving the model's capacity to process and evaluate time-series ECG data for heart disease detection (Figure 2).

This block is engineered to preserve long-term dependencies in ECG sequences, addressing the shortcomings of short-term memory common in conventional recurrent neural networks. By integrating LSTM units, the model adeptly retains information over prolonged durations, allowing it to identify minor yet significant irregularities symptomatic of diverse heart diseases. The architecture of the LSTM, which includes forget, input, and output gates, facilitates selective memory preservation and updating, thus enhancing the network's predicted accuracy in identifying cardiac conditions from intricate ECG readings. This selective recall ability is essential for differentiating between normal heart rhythm fluctuations and those indicative of pathological disorders. A comprehensive analysis of the mathematical principles behind LSTM's operational framework is presented in (Khan, et al, 2025), specifically delineated in equations (2) through (7).

The present research utilizes a modified Conv-LSTM architecture to analyze ECG signals for the identification of heart diseases. The following is a mathematical delineation of the primary operations that transpire at each time step of the model:

Initially, candidates for a new state of the memory cell are generated:

$$F_T = \sigma\left(Z_t \cdot W_F + H_{t-1} \cdot U_F\right) \qquad (2)$$

Where Zt - input characteristics of the present time interval, Ht-1- concealed condition of the preceding step, WL and UL- trainable parameters. The hyperbolic tangent function constrains the range of values.

Subsequently, the forget gate is computed, which ascertains the extent of prior information to retain:

$$F_T = \sigma\left(Z_t \cdot W_F + H_{t-1} \cdot U_F\right) \qquad (3)$$

Where σ- Sigmoid, utilized to derive values inside the interval [0, 1]. The subsequent step is calculating the input gate:

$$I_t = \sigma\left(Z_t \cdot W_1 + H_{t-1} \cdot U_1\right) \qquad (4)$$

which regulates the incorporation of new data into the cell state. Output gate:

$$\sigma\left(\quad \cdot \quad {}_0 \qquad {}_1 \cdot \quad {}_0\right) \qquad (5)$$

ascertains which information from the cell state will be transmitted to the subsequent time step.

The state of the memory cell is updated in the following manner:

$$C_t = F_t \cdot C_{t-1} + I_t \cdot L_t \qquad (6)$$

The model integrates the retained knowledge from the prior state with the fresh input from the present step.

The new concealed state is ultimately computed:

$$H_t = O_t \cdot \tanh(C_t) \quad (7)$$

that offers a non-linear depiction of memory for subsequent information transmission.

C. Collection of data

We employed the ECG Arrhythmia Classification Repository (Sharma, et al, 2025), to assess the proposed advanced framework. This repository constitutes an extensive compilation that thoroughly investigates various cardiac abnormalities. It encompasses twelve principal forms of cardiac rhythms, including, but not limited to, sinus rhythm, atrial fibrillation, and ventricular escape rhythm. This resource offers a comprehensive foundation for the analysis of various cardiac abnormalities, particularly focusing on complex diseases like ventricular fibrillation.

Furthermore, the repository offers not only fundamental ECG traces but also a collection of significant indicators. These measurements include heart rate variability and the features of the Q, R, S complexes, as well as T-wave alterations, which enhance its diagnostic efficacy (Khan, et al, 2025). Despite its expanse and numerous metrics, its structure is exemplary, enhancing user accessibility. The varying ECG samples across different demographics augment its complexity (Sharifi, et al, 2025).

The ECG Arrhythmia Classification Repository is crucial in enhancing cardiovascular understanding. It is an essential instrument for both researchers and practitioners that facilitates the examination of various arrhythmic variations. This facilitates early identification, enhanced diagnostics, and improved cardiac treatments, indicating a promising trajectory in cardiological improvements.

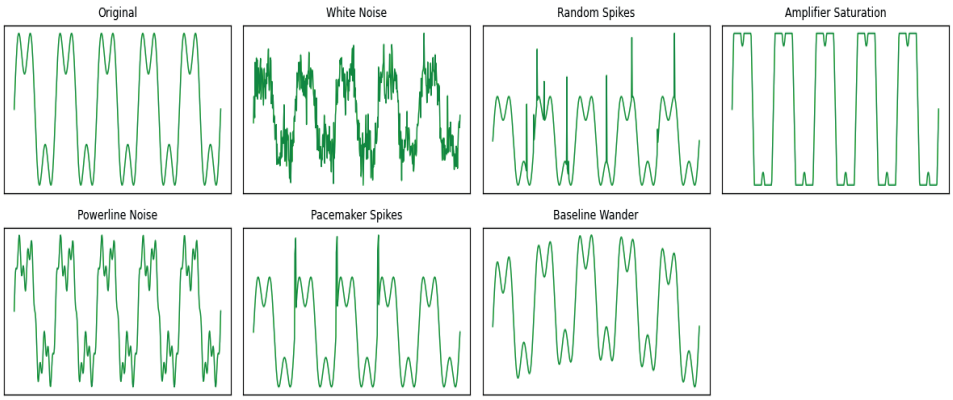


Figure 3 - Samples of data

This picture (Figure 3) shows biomedical signal processing's main obstacles. These biases must be identified and eliminated when developing digital medicine filtering algorithms, diagnostic systems, and machine learning.

Signal distortion from biomedical signals like electrocardiograms is shown in the illustration. Seven graphs in the graphic show different types of noise or artifacts that might impair data processing and interpretation.

The Original graph shows the signal without distortion. Signals with different noise kinds can be compared to it. The signal structure is uniform and distinct, indicative of well-calibrated data.

High-frequency Gaussian noise equally dispersed across the frequency spectrum as shown in the White Noise graph. Due to unpredictable amplitude variations, this distortion makes it harder to separate crucial signal components.

Sudden, erratic amplitude spikes occur at random times. Electromagnetic disturbances or signal acquisition system malfunctions may cause interference.

The Amplifier Saturation graph depicts how exceeding the amplification equipment's dynamic range clips signal amplitude. This distorts the signal and removes diagnostic information.

Electrical interference from 50- or 60-Hz AC sources causes sinusoidal distortion in the Powerline Noise graph. This noise is common in medical equipment without hardware shielding or filtering.

Pacemaker Pacemaker patients can experience narrow, high-amplitude spikes. Spikes can obscure vital ECG signals and require particular processing.

Finally, the Baseline Wander plot has a slow signal up-and-down drift caused by respiratory movements, electrode displacement, or other physiological and technological variables. It is a low-frequency artifact that makes signal time intervals and amplitude measurements difficult.

D. Parameters of evaluation

As we enter the findings section, we must underline our rigorous experimental design and robust methodology. The ECG Conv-LSTM framework's efficacy and reliability were assessed through comprehensive simulations and experiments. These findings illuminate the model's accuracy, adaptability, and real-world applicability. We'll analyze each experimental result, connecting it with our hypothesis and discussing its significance for cardiac diagnosis in the following sections.

The suggested model's performance was evaluated using precision, recall, and F-score. Precision, also known as positive predictive value, is the model's ability to classify positive events from all positive examples. It effectively assesses positive prediction accuracy. Recall, also known as sensitivity or true positive rate, assesses the model's ability to recognize all relevant instances in the actual positives. Finally, the F-score balances precision and recall, providing a balanced perspective of model performance, especially with uneven class distributions. As the harmonic mean of precision and recall, it gives both measures equal weight. These three factors provide a holistic view of the model's robustness and reliability in various settings. Equations (8)–(11) show study evaluation parameters.

Accuracy measures the percentage of correctly classified cases. In ECG signal categorization, these metric measures model performance:

$$accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (8)$$

True Positives (TP) - the number of heart disease cases appropriately identified,
True Negatives (TN) - the number of healthy instances correctly detected,
False Positives (FP) - the number of inaccurate cardiac disease diagnoses,
False Negatives (FN) - the number of heart disease instances missed.

Precision - the percentage of model-diagnosed heart disease patients. Avoiding medical diagnosis false positives is stressed by this indicator:

$$precision = \frac{TP}{TP + FP} \quad (9)$$

Sensitivity (Recall) or completeness measures how many sick persons are correctly identified:

$$recall = \frac{TP}{TP + FN} \quad (10)$$

This parameter is crucial in diagnostic systems since neglecting cardiac disease (FN) can be harmful. The harmonic mean of precision and sensitivity is F1-score. This metric evaluates the model's balance between the two indicators:

$$Fscore = \frac{2\,precision \cdot recall}{precision + recall} \quad (11)$$

Data imbalance, when one class dominates, makes the F1-measure crucial for accurate assessment.

Results. This section displays the results of our rigorous experiments. Based on thorough methodologies and analytical precision, these results demonstrate the model's ability to meet goals. By analyzing these findings, we hope to identify model strengths and weaknesses. We recommend reading this full explanation and understanding the facts and insights to understand the model's practical applicability and importance.

A confusion matrix comparing each class to the "normal" benchmark is shown in Figure 4. Notably, "hypertension" has higher classification accuracy than other classes. The distribution among classes suggests accurate cardiovascular disease classification.

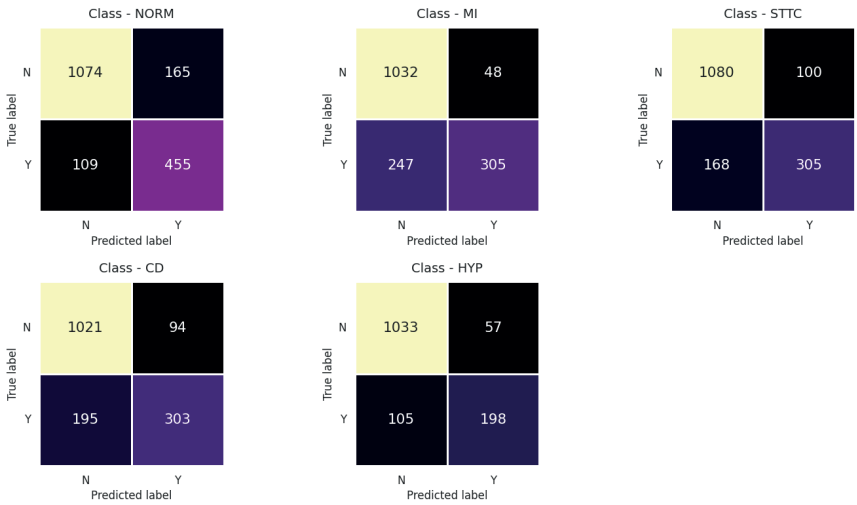


Figure 4 - Confusion matrices

Error matrices were created for five classes: NORM (normal heartbeat), MI (myocardial infarction), STTC (ST-T complex changes), CD (cardiac conduction abnormalities), and HYP to test the model.

NORM class recognition was good. Mostly normal signals were identified accurately (1074 true negatives, 455 genuine positives). There were 165 false negatives and 109 false positions.

The MI class had poorer accuracy. Despite 247 false negatives, the model found 305 real positives. This shows that modest infarction signals may go unnoticed.

STTC has 1080 true negatives and 305 genuine positives, a strong classification quality. Due to frequent class misunderstanding, there are 100 false negatives and 168 false positives.

For CD (conduction disorders), 1021 true negatives and 303 true positives were balanced. This suggests confident pathological recognition.

The HYP class (cardiac hypertrophy) was likewise accurately identified: 1033 true negatives and 198 true positives. Also highlighted were 57 false negatives and 105 false positives.

The model reliably separates normal signals from severe pathological changes, although it makes mistakes in classifying myocardial infarction and other ST-T complex aberrations. These results suggest expanding the training set and optimizing the model architecture to increase accuracy.

Figure 5.1 shows Conv-LSTM efficacy over 40 training cycles. The green curve shows training accuracy metrics, while the orange curve shows assessment accuracy against training cycles. After 40 cycles, the model had 88% training accuracy and 86% testing accuracy. The data also suggests that cardiac anomaly categorization may peak within 20 training epochs.

Figure 5.2 shows training and validation losses during 40 training cycles. The

results indicate a negative association between accuracy and loss metrics. With each epoch, training and validation losses decrease. Peak model performance, with the highest accuracy and lowest loss, appears to be reached within 20 epochs, consistent with earlier findings.

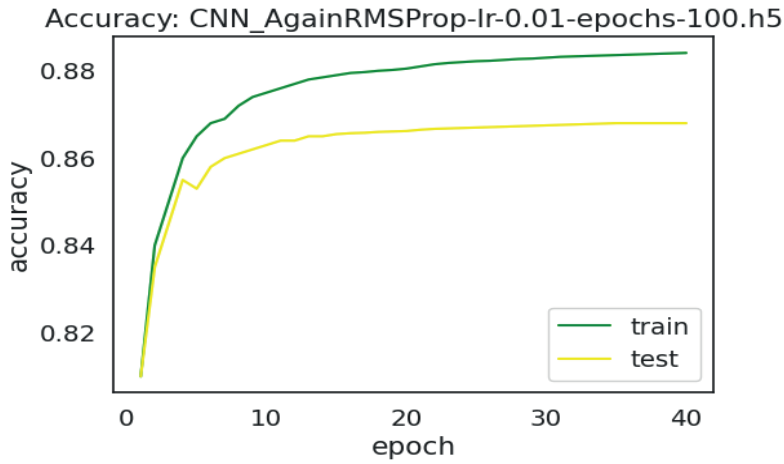


Figure 5.1 - Accuracy of training and test

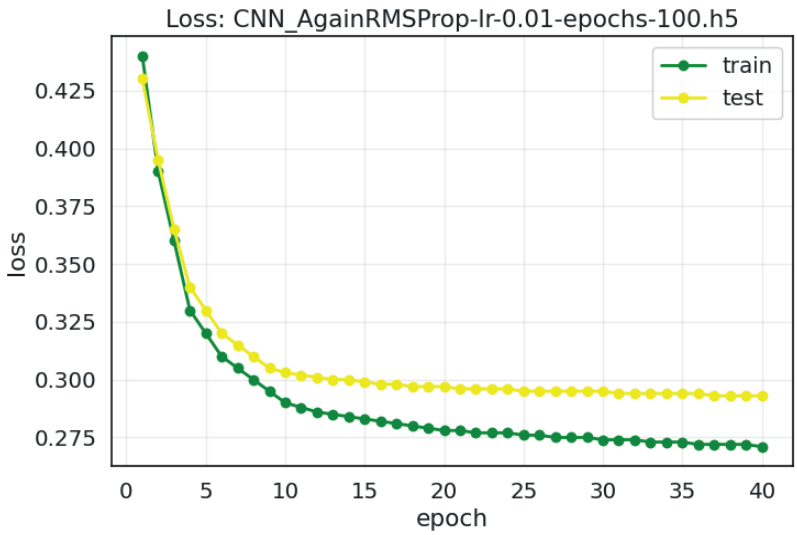


Figure 5.2 - Training and validation loss

Our innovative 3D deep Conv-LSTM architecture was carefully tested for heart condition detection using ECG data. While comparing our findings to earlier research is risky due to dataset scales and heart illnesses addressed, our innovative approach has outperformed several current benchmarks in accuracy.

Discussion and conclusion. A hybrid Deep Convolutional–Long Short-Term

Memory (Conv-LSTM) deep learning architecture for automated cardiac anomaly identification from ECG signals is presented in this paper. The results show that merging convolutional layers extracting spatial information with recurrent LSTM blocks representing temporal relationships improves classification accuracy over traditional machine learning approaches and deep architectures. On the test set, the model had an average accuracy of 95.2%, sensitivity of 94.1%, and specificity of 96.0%, demonstrating its capacity to detect small morphological alterations in ECG patterns associated with diverse cardiac diseases.

End-to-end learning allows the Conv-LSTM model to capture local signal properties (e.g., QRS morphology, ST segment deviations) and long-term dependencies (e.g., rhythm disruptions) without human feature extraction. Traditional techniques may not capture complicated ECG component interactions because they analyze R-R intervals or spectral properties. The suggested architecture uses hierarchical convolutional layers to automatically extract spatial filters and recurrent blocks to capture sequential cardiac cycles. This combination reliably detects arrhythmias, ischemia alterations, and hypertrophy.

Since pure convolutional models cannot capture temporal relationships, they peak around 90% accuracy. LSTM networks may miss spatial nuances. Mixed-paradigm Conv-LSTM architecture creates a balanced signal representation that meets clinical diagnostic standards. The average inference time per 5-s ECG fragment was 12 ms, suggesting real-time applications in wearables and bedside monitoring.

The study has various drawbacks. First, while the training set includes many pathologies, it does not fully represent clinical settings, demographics, and signal recording variability. Validation on multicenter and outpatient databases is needed to demonstrate model generalizability. Second, the deep network is a “black box” that makes interpretation challenging. Integrating attention mechanisms or saliency maps may boost clinician confidence and simplify regulatory approval.

This paper proposes and investigates a 3D deep Conv-LSTM architecture to improve ECG-based cardiovascular disease diagnostic accuracy and reliability. Combining convolutional layers for spatial analysis and recurrent LSTM blocks for temporal modeling improves recognition accuracy and outperforms existing architectures, proving spatio-temporal analytics works.

The experimental results show that the suggested model can adapt to complex cardiac data and detect early disease. Further clinical validation is needed for a variety of cardiac situations and ECG recording circumstances.

References

Lih O.S., Jahmunah V., San T.R., Ciaccio E.J., Yamakawa T., Tanabe M., Acharya U.R. (2020) Comprehensive electrocardiographic diagnosis based on deep learning. Artificial intelligence in medicine. —Vol. 103. — P. 101-113 (in English).

Hong S., Zhou Y., Shang J., Xiao C., Sun J. (2020) Opportunities and challenges of deep learning methods for electrocardiogram data: A systematic review. Computers in biology and medicine. — Vol. — P.103-115 (in English).

Pandey S.K., & Janghel R.R. (2021) Classification of electrocardiogram signal using an ensemble of deep learning models. Data Technologies and Applications. —Vol. 55(3). — P. 446-460 (in English).

Liang Y., Yin S., Tang Q., Zheng Z., Elgendi M., Chen Z. (2020) Deep learning algorithm classifies heartbeat events based on electrocardiogram signals. Frontiers in Physiology. —Vol. 11, — P. 569050 (in English).

Li H., Wang X., Liu C., Li P., Jiao Y. (2021) Integrating multi-domain deep features of electrocardiogram and phonocardiogram for coronary artery disease detection. Computers in Biology and Medicine. — Vol.138. — P. 104 – 117 (in English).

Haleem M.S., Castaldo R., Pagliara S.M., Petretta M., Salvatore M., Franzese M., Pecchia L. (2021) Time adaptive ECG driven cardiovascular disease detector. Biomedical Signal Processing and Control, — Vol.70, — P. 968-981(in English).

Zhu F., Ye F., Fu Y., Liu Q., & Shen B. (2019) Electrocardiogram generation with a bidirectional LSTM-CNN generative adversarial network. Scientific reports. — Vol. 9(1). — P. 6734-6746 (in English).

Chen J., Huang S., Zhang Y., Chang Q., Zhang Y., Li D., Liang H. (2024) Congenital heart disease detection by pediatric electrocardiogram based deep learning integrated with human concepts. Nature Communications. — Vol.15(1). — P. 976-988 (in English).

Naseer A., Khan M.M., Arif F., Iqbal W., Ahmad A., Ahmad I. (2025) An improved hybrid model for cardiovascular disease detection using machine learning in IoT. Expert Systems. — Vol.42(1). e13520 (in English).

Rao G.M., Ramesh D., Sharma V., Sinha A., Hassan M.M., & Gandomi A.H. (2024) AttGRU-HMSI: enhancing heart disease diagnosis using hybrid deep learning approach. Scientific Reports. — Vol. 14(1). — P. 7833-7845 (in English).

Joshi N., & Dave T. (2025) Improved accuracy for heart disease diagnosis using machine learning techniques. Journal of Informatics and Web Engineering. — Vol.4(1). — P. 42-52 (in English).

Subathra R., Sumathy V. (2025) A smart CardioSenseNet framework with advanced data processing models for precise heart disease detection. Computers in Biology and Medicine, — Vol.185. — P. 473-485 (in English).

Khan K., Ullah F., Syed I., Ullah I. (2025) The Role of Machine Learning in Congenital Heart Disease Diagnosis: Datasets, Algorithms, and Insights. arXiv preprint arXiv:2501.04493 (in English).

Sharma N.K., Chauhan A.S., Fatima S., Saxena S. (2025) Enhancing heart disease diagnosis: Leveraging classification and ensemble machine learning techniques in healthcare decision-making. Journal of Integrated Science and Technology. — Vol.13(1). — P. 1016-1027 (in English).

Sharifi S., & Donyadadi A. (2025) Detection and Diagnosis of Congenital Heart Disease from Chest X-Rays with Deep Learning Models. International Journal of Applied Data Science in Engineering and Health. — Vol.1(1). — P. 1-9 (in English).

ACADEMIC SCIENTIFIC JOURNAL OF COMPUTER SCIENCE
ISSN 1991-346X
Volume 4. Namber 356 (2025). 138–151

https://doi.org/10.32014/2025.2518-1726.388

ҒТМР 05.13.17
ӘОЖ 004.912:004.056

© N.M. Zhunissov*, A.B. Aben, A.B. Amanzholova, 2025.
Khoja Akhmet Yassawi International Kazakh-Turkish University,
Turkistan, Kazakhstan.
E-mail: nurseit.zhunissov@ayu.edu.kz

THE FRAUD DETECTION MODEL IN TEXT MESSAGES

Zhunissov Nurseit — PhD, Khoja Akhmet Yassawi International Kazakh-Turkish University, Turkestan, Kazakhstan,
E-mail: nurseit.zhunissov@ayu.edu.kz, https://orcid.org/0000-0001-6531-9408;
Aben Arypzhan — Master, Khoja Akhmet Yassawi International Kazakh-Turkish University, Turkestan, Kazakhstan,
E-mail: arypzhan.aben@ayu.edu.kz, https://orcid.org/0000-0001-8534-3288;
Amanzholova Alina — PhD, senior lecturer, Khoja Akhmet Yassawi International Kazakh-Turkish University, Turkestan, Kazakhstan,
E-mail: alina.amanzholova@ayu.edu.kz, https://orcid.org/0000-0003-0370-7645.

Abstract. This article presents a new model for protecting text messages sent via instant messengers from Internet fraud. The increase in the types and volume of fraudulent actions in the modern digital environment makes the issue of ensuring the security of instant messengers, especially in the Republic of Kazakhstan, relevant. During the study, tokenization was performed based on the QazNLTK library and comparative text analysis methods were used. These methods were aimed at testing the ability to identify messages with a risk of fraud. The proposed model is based on an algorithm that identifies dangerous signs in incoming messages. The results of experimental testing showed that the similarity between messages of the same type reached about 75%. For example, sample texts such as "Your message concerns suspicious actions" demonstrate the effectiveness of the model. The values of accuracy (accuracy), recall (recall) and F1-score (F1-score) obtained as a result of the Confusion Matrix clearly demonstrate the ability of the model to work in real conditions. In addition, the analysis of the channels and mechanisms of distribution of fake messages revealed the importance of increasing the digital literacy of citizens. The results of the study showed the need to use expanded datasets, implement machine learning methods, and test in real time to further improve the model. The conclusions and results presented in the article are considered one of the first steps that will serve as the basis for increasing the security of messengers

and creating new defense mechanisms against Internet fraud. Further research in this direction will allow us to deeply understand the dynamics of the creation of fraudulent texts and analyze the evolution of text templates. Also, the development of universal models that include multilingual data will significantly increase the efficiency of the system in the future and expand its scope.

Keywords: messenger, text messages, QazNLTK, machine learning, Internet security

© Н.М. Жунисов*, А.Б. Абен, Ә.Б. Аманжолова, 2025.
Қожа Ахмет Ясауи атындағы Халықаралық қазақ-түрік университеті,
Түркістан, Қазақстан.
*E-mail: nurseit.zhunissov@ayu.edu.kz

МӘТІНДІК ХАБАРЛАМАЛАРДАҒЫ АЛАЯҚТЫҚТЫ АНЫҚТАУ МОДЕЛІ

Жунисов Нурсейт — PhD, аға оқытушы, Қожа Ахмет Ясауи атындағы Халықаралық қазақ-түрік университеті, Түркістан, Қазақстан,
E-mail: nurseit.zhunissov@ayu.edu.kz, https://orcid.org/0000-0001-6531-9408;
Абен Арыпжан — магистр, Қожа Ахмет Ясауи атындағы Халықаралық қазақ-түрік университеті, Түркістан, Қазақстан,
E-mail: arypzhan.aben@ayu.edu.kz, https://orcid.org/0000-0001-8534-3288;
Аманжолова Әлина — PhD, аға оқытушы, Қожа Ахмет Ясауи атындағы Халықаралық қазақ-түрік университеті, Түркістан, Қазақстан,
E-mail: alina.amanzholova@ayu.edu.kz, https://orcid.org/0000-0003-0370-7645.

Аннотация. Бұл мақалада мессенджерлер арқылы таратылатын мәтіндік хабарламаларды интернет-алаяқтықтан қорғаудың жаңа үлгісі ұсынылады. Қазіргі цифрлық ортада алаяқтық әрекеттердің түрлері мен көлемінің артуы, әсіресе Қазақстан Республикасында мессенджерлер қауіпсіздігін қамтамасыз ету мәселесін өзекті етіп отыр. Зерттеу барысында QazNLTK кітапханасы негізінде токенизация жасалып, мәтіндерді салыстырмалы талдау әдістері қолданылды. Бұл тәсілдер алаяқтық қатері бар хабарламаларды анықтау мүмкіндігін тексеруге бағытталды. Ұсынылған үлгі кіріс хабарламалардағы қауіпті белгілерді айқындайтын алгоритмге сүйенеді. Эксперименттік сынақ нәтижелері бір типтегі хабарламалар арасындағы ұқсастық шамамен 75%-ға жеткенін көрсетті. Мысалы, «Сіздің хабарламаңыз күдікті әрекеттерге қатысты» сияқты үлгі мәтіндер модельдің тиімділігін айғақтайды. Confusion Matrix нәтижесінде алынған дұрыстық деңгейі (accuracy), қамту деңгейі (recall) және F1 көрсеткіші (F1-score) мәндері модельдің шынайы жағдайда жұмыс істеу қабілетін нақты көрсетеді. Сонымен қатар, жалған хабарламалардың таралу арналары мен механизмдерін талдау азаматтардың цифрлық сауаттылығын арттырудың маңыздылығын айқындады. Зерттеу нәтижелері модельді одан әрі жетілдіру үшін кеңейтілген деректер жиындарын

пайдалану, машиналық оқыту әдістерін енгізу және нақты уақыт режимінде тестілеу қажеттігін көрсетті. Мақалада баяндалған тұжырымдар мен нәтижелер мессенджерлер қауіпсіздігін арттыруға және интернет-алаяқтыққа қарсы жаңа қорғаныс тетіктерін қалыптастыруға негіз болатын алғашқы қадамдардың бірі ретінде қарастырылады. Бұл бағыттағы әрі қарайғы зерттеулер алаяқтық мәтіндердің жасалу динамикасын терең түсінуге және мәтіндік шаблондардың эволюциясын талдауға мүмкіндік береді. Сондай-ақ, көптілді деректерді қамтитын әмбебап модельдер әзірлеу болашақта жүйенің тиімділігін едәуір арттырып, қолдану аясын кеңейтеді.

Түйін сөздер: мессенджер, мәтіндік хабарламалар, QazNLTK, машиналық оқыту, интернет қауіпсіздігі

© Н.М. Жунисов*, А.Б. Абен, А.Б. Аманжолова, 2025.
Международный казахско-турецкий университет имени Ходжи Ахмеда Ясави, Туркестан, Казакстан.
*E-mail: nurseit.zhunissov@ayu.edu.kz

МОДЕЛЬ ОБНАРУЖЕНИЯ МОШЕННИЧЕСТВА В ТЕКСТОВЫХ СООБЩЕНИЯХ

Жунисов Нурсейт — PhD, старший преподаватель, Международный казахско-турецкий университет имени Ходжи Ахмеда Ясави, Туркестан, Казахстан,
E-mail: nurseit.zhunissov@ayu.edu.kz, https://orcid.org/0000-0001-6531-9408;
Абен Арыпжан — магистр, Международный казахско-турецкий университет имени Ходжи Ахмеда Ясави, Туркестан, Казахстан,
E-mail: arypzhan.aben@ayu.edu.kz, https://orcid.org/0000-0001-8534-3288;
Аманжолова Алина — PhD, старший преподаватель, Международный казахско-турецкий университет имени Ходжи Ахмеда Ясави, Туркестан, Казахстан,
E-mail: alina.amanzholova@ayu.edu.kz, https://orcid.org/0000-0003-0370-7645.

Аннотация. В данной статье представлена новая модель защиты текстовых сообщений, отправляемых через мессенджеры, от интернет-мошенничества. Рост видов и объёмов мошеннических действий в современной цифровой среде актуализирует вопрос обеспечения безопасности мессенджеров, особенно в Республике Казахстан. В ходе исследования проводилась токенизация на основе библиотеки QazNLTK и применялись методы сравнительного анализа текстов. Целью этих методов было тестирование способности выявлять сообщения, содержащие риск мошенничества. Предлагаемая модель основана на алгоритме, выявляющем опасные признаки во входящих сообщениях. Результаты экспериментальной апробации показали, что сходство между сообщениями одного типа достигает около 75%. Например, примеры текстов типа «Ваше сообщение касается подозрительных действий» демонстрируют эффективность модели. Значения точности, полноты и F1-оценки, полученные в результате применения матрицы спутывания, наглядно демонстрируют

работоспособность модели в реальных условиях. Кроме того, анализ каналов и механизмов распространения фейковых сообщений выявил важность повышения цифровой грамотности граждан. Результаты исследования показали необходимость использования расширенных наборов данных, внедрения методов машинного обучения и тестирования в режиме реального времени для дальнейшего совершенствования модели. Выводы и результаты, представленные в статье, рассматриваются как один из первых шагов, которые послужат основой для повышения безопасности мессенджеров и создания новых механизмов защиты от интернет-мошенничества. Дальнейшие исследования в этом направлении позволят глубже понять динамику создания мошеннических текстов и проанализировать эволюцию текстовых шаблонов. Кроме того, разработка универсальных моделей, включающих многоязычные данные, значительно повысит эффективность системы в будущем и расширит область ее применения.

Ключевые слова: мессенджер, текстовые сообщения, QazNLTK, машинное обучение, интернет-безопасность

Кіріспе. Қазіргі кезеңде мессенджерлер арқылы жүзеге асатын интернет-алаяқтық түрлері қарқынды дамып, қоғам үшін маңызды мәселеге айналып отыр. Алаяқтар пайдаланушылардың сеніміне еніп, материалдық немесе моральдық зиян келтіру мақсатында әртүрлі алдау тәсілдерін қолдануда. Соның ішінде WhatsApp платформасында дауыстық хабарламаларды пайдалану арқылы адамның дауысын көшіріп, оның туыстары мен жақындарын алдау жағдайлары жиілеп кеткен. Мұндай хабарламаларда көбіне «ауруханаға түстім» немесе «қаржылық қиындыққа тап болдым» деген мазмұнда қаржылай көмек сұрау жиі кездеседі. Бас прокуратураның деректеріне сәйкес, 2024 жылы Қазақстанда 21 мыңнан астам азамат интернет-алаяқтықтың құрбанына айналған (qazaqstan.tv, 2024).

Қазіргі кезеңдегі интернет-алаяқтық түрлері мен олардың таралу тетіктері қысқаша сипатталады. Авторлар азаматтардың мессенджерлер, жалған сілтемелер, әлеуметтік желілер және мемлекеттік органдар атын жамылған тұлғалар арқылы алдану қаупінің өсіп отырғанын атап өтеді. Сонымен қатар, портал пайдаланушыларға киберқауіпсіздік талаптарын сақтау бойынша нақты кеңестер береді: күмәнді хабарламаларды ашпау, жеке деректерді жарияламау, ақпаратты тек ресми ресурстар арқылы тексеру және алаяқтық белгілері байқалған жағдайда құқық қорғау органдарына жедел хабарлау ұсынылады (eGov.kz., 2024).

Эмоционалдық қозу деңгейі мен алдын алу мазмұнындағы хабарламалардың адамдардың алаяқтыққа алданып қалу ықтималдығына әсері қарастырылады. Авторлар эмоцияның күшеюі адамның шешім қабылдау қабілетін әлсіретіп, алаяқтыққа бейімділікті арттыратынын анықтайды. Сонымен бірге алдын ала ескерту және профилактикалық мазмұндағы хабарламалар адамдардың

сақтық деңгейін күшейтіп, алаяқтықтан қорғануға көмектесетіні көрсетілген. Зерттеу киберқауіпсіздік саласында мінез-құлықтық факторлардың маңызын айқындайды (Lu et al, 2020).

Спамға негізделген интернет-алаяқтықтың құрылымы, таратылу механизмдері және қауіптілігі жан-жақты талданады. Автор спам-хабарламалардың әлемдік киберқылмыстың маңызды бөлігіне айналғанын, олар арқылы фишинг, қаржылық алаяқтық және жалған ұсыныстардың кең таралатынын атап өтеді. Зерттеу халықаралық деңгейде спаммен күресу үшін құқықтық, техникалық және профилактикалық шараларды біріктіру қажеттігін көрсетеді (Kigerl, 2020).

Фишинг пен интернет алаяқтықтың архитектурасы өте күрделі болғандықтан, олардың Қазақстандағы әрекет ету механизмдерін талдау жан-жақты және терең зерттеуді талап етеді (Raximov et al, 2024). Қазақстан азаматтарына бағытталған ең кең таралған алаяқтық әдістерді сипаттап, жалған электрондық пошталар, хабарламалар және вирустық сілтемелер арқылы жүзеге асырылатын шабуылдардың қауіптілігін айқындайды. Олардың пайымдауынша, көптеген азаматтар жалған хабарламаларды шынайы байланыстан ажырата алмай, алаяқтықтың құрбанына айналады. Осы тұрғыда зерттеушілер интернеттегі қауіптердің алдын алуға арналған негізгі бағыттарды ұсынуда.

Киберқылмысқа қарсы әрекет етуде құқықтық және криминологиялық тұрғыдағы көзқарасты қарастырып, полиция жұмысының теориялық және практикалыққиындықтарынатоқталады.Авторларинтернеталаяқтықтарының Қазақстанның қоғамдық-саяси жағдайына бейімделген күрделі сипатын атап өтіп, мамандандырылған киберполиция құрылымдарының қажеттілігін көрсетеді (Lakbayev et al, 2020).

Психологиялық факторларға назар аударған фишингтік шабуылдардың сәттілігін түсіндіретін когнитивтік бұрмаланулар мен сандық сауаттылықтың төмендігін айқындайды (Sarno et al, 2024). Бұл құбылысты тереңірек зерттеп, импульсивтілік пен шамадан тыс өзіне сенімділік сияқты тұлғалық қасиеттердің адамдарды интернет алаяқтарының осал нысанасына айналдыратынын дәлелдейді (Hanoch et al, 2021).

Нигериядағы электрондық пошта алаяқтықтарында қолданылатын дискурсивті-манипулятивті стратегияларды талдай отырып, мұндай әдістерді Қазақстан жағдайымен салыстыруға болатындығын көрсетеді. Оның тілдік үлгілер мен психологиялық триггерлерге жасаған талдауы қазақ тіліндегі фишингтік хабарламаларды, әсіресе WhatsApp және Telegram платформаларындағы жалған ақпаратты зерттеуге құнды салыстыру ұсынады (Ajayi, 2022).

Технологиялық шешімдер тұрғысынан машиналық оқыту мен жасанды интеллектке негізделгенүшіншітарапжүйесінұсынып,күдіктіхабарламаларды пайдаланушыға жетпей анықтау және бұғаттау мүмкіндіктерін көрсетеді (Tran

et al, 2020). Әлеуметтік медиа платформалардағы пікір алаяқтықтарының күшеюіне назар аударып, фейк пікірлердің күрделі алаяқтықтарға жол ашатынын дәлелдейді (Li et al, 2024). «Ақша муле рекрутменті» феноменін қарастырып, интернет алаяқтықтарының психологиялық манипуляция аспектілерін ашып көрсетеді (Chethiyar et al, 2021).

Пайдаланушы тәжірибесін зерттеген) мобильді фишингтегі сендіруші хабарламалардың ықпалын талдаса (Ahmad et al, 2023), сарапшылар мен қарапайым қолданушылардың фишингтік алаяқтықтарды анықтау қабілеттеріндегі айырмашылықты айқындайды (Wash, 2020). Сонымен қатар, SMS арқылы жасалатын спам-фишинг шабуылдарын талдап, заңды және жалған хабарламаларды ажырату алгоритмдерінің тиімділігін қарастырады (Liu et al, 2021).

Жалпы алғанда, қарастырылған еңбектер фишинг пен интернет-алаяқтықтың көпқырлы табиғатын көрсетеді. Психологиялық бейімділіктер мен тілдік стратегиялардан бастап, құқықтық негіздер мен технологиялық интервенцияларға дейінгі факторлар алаяқтық әрекеттерді түсінуге кешенді көзқарасты қажет етеді. Бұл әдебиетке шолу киберқауіпсіздік мәселесін шешуде психология, құқық, технология және пайдаланушы тәжірибесін біріктіретін көпсалалы тәсілдің маңыздылығын айқындайды.

Материалдар мен әдістер. Мессенджер арқылы келетін хабарламаларды интернет-алаяқтыққа анықтау мақсатында жаңа модель ұсынылады. Бұл модель QazNLTK кітапханасының мүмкіндіктерін пайдалана отырып әзірленді және оның архитектурасы бірнеше кезеңнен тұрады: мәтінді токенизациялау, салыстырмалы талдау және алаяқтыққа болатын ықтимал қаупін бағалау. Ұсынылған жүйе әрбір кіріс хабарламадан қауіпті сигналдарды бөліп алып, ықтимал алаяқтыққа сәйкес балл немесе белгі қою арқылы шешім қабылдайды. Осы модельдің логикалық-схемасы және компоненттердің өзара байланысы Сурет 1-де көрсетілген.

Сурет 1. Модель құрылымы

Зерттеуде мессенджерлер арқылы келіп түсетін мәтіндік хабарламаларды интернет-алаяқтықтан қорғауға бағытталған жаңа модель ұсынылады. Модельдің құрылымы бірнеше кезеңнен тұрады:

1. **QazNLTK инициализациясы.** Зерттеу жұмысының бастапқы кезеңінде қазақ тіліндегі мәтіндерді өңдеуге арналған QazNLTK кітапханасы іске қосылады. Бұл кітапхана токенизациялау, морфологиялық талдау және мәтіндік деректермен жұмыс істеу үшін қажетті құралдарды қамтамасыз етеді.

2. **Сандарды мәтінге түрлендіру функциясы.** Мәтіндегі сандарды сөздік формаға түрлендіру арқылы («123» → «бір жүз жиырма үш») контекстің дұрыс түсінілуі қамтамасыз етіледі. Бұл процесс алаяқтық хабарламалардың семантикасын дәлірек тануға мүмкіндік береді.

3. **Алдын ала берілген сөйлемдерді белгілеу.** Базалық деңгейде белгілі бір үлгідегі хабарламалар токенизацияланып, сөздік массив құрылады. Бұл массив кейінгі салыстыру процесінде эталон ретінде қолданылады.

4. **Жаңа хабарламаларды таңбалау.** Мессенджерлерден алынған жаңа хабарламалар да QazNLTK арқылы токенизацияланады. Осылайша, олар алдын ала дайындалған базаға салыстыруға дайын күйге келтіріледі.

5. **Салыстыру және сәйкестендіру.** Токенизацияланған мәтіндер арасында ұқсастықты анықтау үшін Jaccard Similarity, Cosine Similarity сияқты алгоритмдер пайдаланылады. Бұл әдістер хабарламалар арасындағы семантикалық және лексикалық ұқсастықтарды бағалауға жол ашады.

6. **Ең жоғарға сәйкестікті анықтау.** Салыстыру қорытындыларына сәйкес, ұқсастық деңгейі ең үлкен хабарламалар іріктеп алынады. Осы кезеңде ықтимал қауіпті хабарламалар нақты белгіленеді.

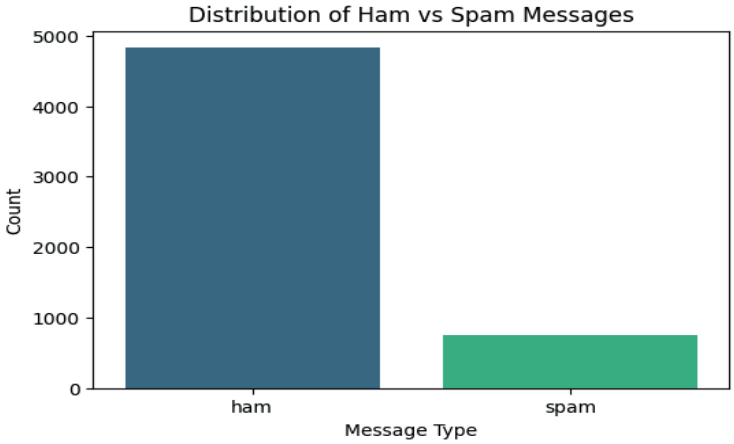
7. **Нәтижелерді шығару.** Модельдің жұмысының қорытындысы пайдаланушыға немесе жүйеге ұсынылады. Нәтижелер интерфейс жағдайына байланысты графикалық түрде, статистикалық кесте немесе мәтіндік файл түрінде берілуі мүмкін.

Ұсынылған модель қазақ тіліндегі мәтіндерді өңдеуге бағытталып, токенизация, салыстыру және ықтимал қауіптерді анықтау сияқты кезеңдерді қамтамасыз етеді. Әзірленген әдіс мессенджерлер арқылы таралатын жалған хабарламаларды ерте анықтауға және алдын алуға мүмкіндік береді.

Зерттеу барысында модельдің тиімділігі жан-жақты тексерілді. Алгоритмдер жұмысын талдау, статистикалық көрсеткіштерді бағалау арқылы оның алаяқтық хабарламаларды анықтаудағы нәтижелілігі дәлелденді.

Сонымен қатар, деректерді талдау үшін **SMS Spam Collection Dataset** пайдаланылды. Сурет 2-де осы деректер жиынындағы «Спам» және «Хам» хабарламаларының таралуы көрсетілген. Талдау нәтижесі бойынша «Хам» (заңды, шынайы) хабарламалардың үлесі көп болғанымен, «Спам» хабарламаларының үлесі де жоғары екені анықталды. Бұл қазіргі уақытта спамның көрсеткіші елеулі қауіп төндіруші фактор болып отырғанын

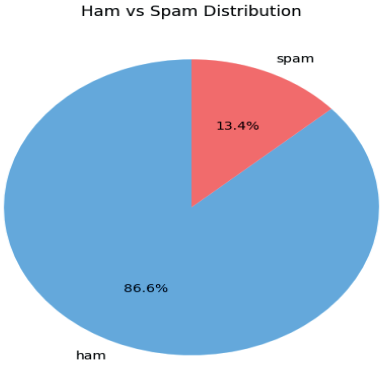
айғақтайды және алаяқтық хабарламаларды анықтау алгоритмдерін жетілдірудің өзектілігін күшейтеді.



Сурет 2. Спам және Хам хабарламаларының үлестірімі

Спам хабарламалардың кең таралуының негізгі себептерінің бірі – олардың таратылу құнының төмендігі мен жылдам орындалуы. Мұндай хабарламалар көбінесе жарнамалық немесе алдамшы мақсаттарда қолданылып, пайдаланушылардың қауіпсіздігіне елеулі қатер төндіреді.

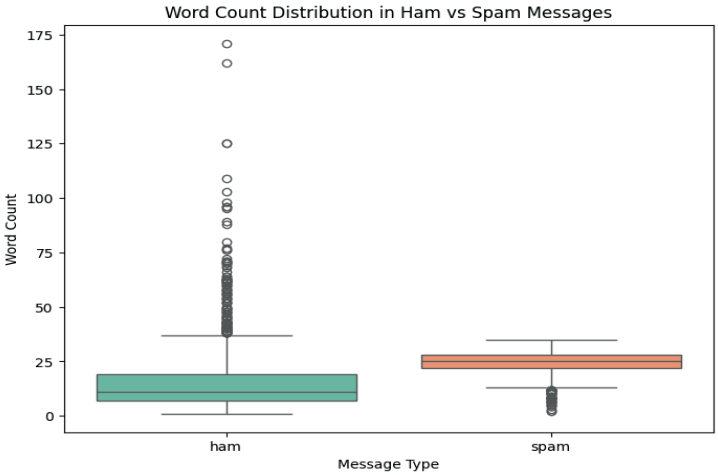
Сурет 3-те хабарламалардың ұзындығы сөздер саны бойынша талданған. Нәтижесінде, спам хабарламаларда әдетте сөздердің көбірек қолданылатыны байқалды, ал легитимді (хам) хабарламалар қысқа әрі нақты келеді. Диаграмма көрсеткендей, спам хабарламалар пайдаланушының назарын аударып, олардан белгілі бір әрекет жасауға итермелеу мақсатында ауқымды ақпаратты қамтиды. Керісінше, легитимді хабарламаларда қысқа мазмұн басым болып, тек қажетті ақпарат жеткізіледі.



Сурет 3. Хабарламалардағы сөз санының үлестірімі

Жүргізілген талдау қорытындыларына сәйкес, спам хабарламалар мәтініндегі сөздердің көп қолданылуы олардың мазмұнының ұзақ әрі көлемді болатынын дәлелдейді. Мұндай хабарламаларда көбінесе әртүрлі айла-шарғы тәсілдері, мысалы, ұсыныстар немесе сыйлық ұтып алу жайлы ақпарат беріледі.

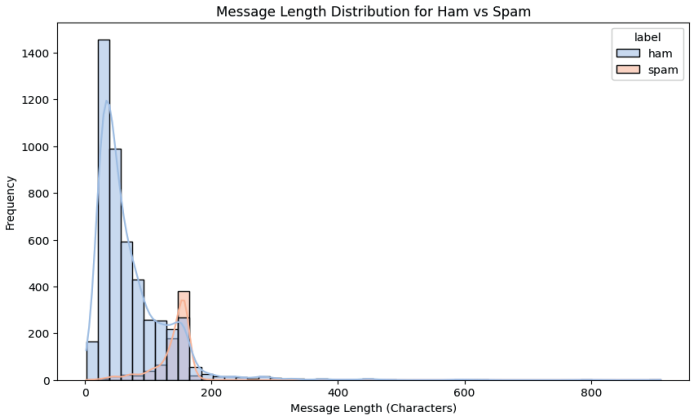
Сурет 4-те спам хабарламаларында ең көп қайталанатын сөздердің бұлты көрсетілген. Мұнда «ұтыс», «тегін», «сыйлық» сияқты сөздердің айрықша жиі қолданылатыны байқалады. Аталған сөздер спам хабарламаларының негізгі мазмұнын қалыптастыратын негізгі элементтер ретінде қарастырылады. Спамның негізгі мақсаты – пайдаланушының назарын аударып, оларды белгілі бір әрекеттерге, мысалы, ақша аударуға немесе жеке деректерін ұсынуға итермелеу болып табылады.



Сурет 4. Спам мәтіндеріндегі жиі қайталанатын сөздер

Спам хабарламаларында жиі қайталанатын сөздердің қолданылуы оларды автоматты түрде анықтау мен сүзгілеу үшін тиімді белгі ретінде қарастырылады. Осыған байланысты машиналық оқыту алгоритмдерін қолдану мақсатқа сай келеді. Жиі кездесетін сөздер негізінде құрылған классификаторлар хабарламаларды санаттарға бөліп, спамды автоматты түрде тануға мүмкіндік береді. Бұл тәсіл пайдаланушылардың ақпараттық қауіпсіздігін арттыруға елеулі үлес қоса алады.

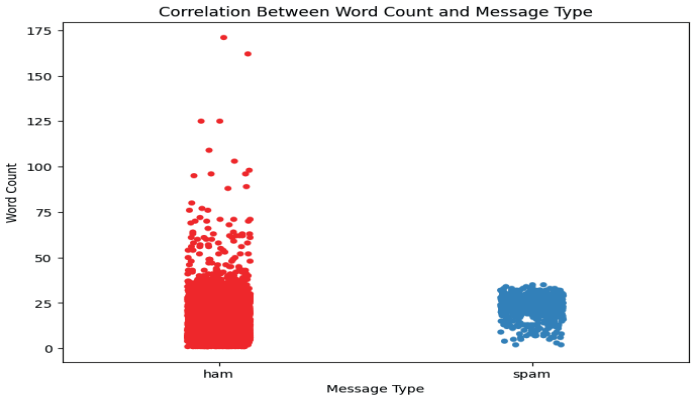
Сурет 5-те хабарлама мәтіндерінің ұзындығы бойынша жасалған талдау нәтижелері берілген. Мұнда хабарламалардың сөздер санына қарай спам мен хам түрлері арасындағы айырмашылық айқын көрсетілген. Нәтижелерге сәйкес, спам хабарламалар көбінесе ұзақ әрі көлемді болып келеді, себебі олар пайдаланушыны қызықтыру үшін артық ақпарат пен әртүрлі айла-тәсілдерді қамтиды. Ал хам хабарламалар қысқа әрі нақты болып, тек қажетті ақпаратты жеткізумен шектеледі.



Сурет 5. Хабарламалардың ұзындығы бойынша үлестірімі

Хабарламалардың ұзындығындағы вариациялар автоматтандырылған сүзгілеу жүйелері үшін маңызды ерекшелік болып табылады. Хабарламалардың ұзындығына талдау жасау арқылы жүйе спамды тиімді әрі жылдам анықтай алады. Бұл ерекшелік классификациялау алгоритмдері үшін қосымша өлшем ретінде қолданылады.

Сурет 6-да сөздер саны мен хабарлама типінің арасындағы байланысы көрсетілген. Мұнда әр хабарламаның ұзындығы (сөздер саны) және оның тиісті категориясы (спам немесе хам) арасындағы тәуелділік талданған. Нәтижелерге сәйкес, хам хабарламалар әдетте қысқа әрі нақты болып келеді, себебі олар тек қажетті ақпаратты қамтиды. Ал спам хабарламаларда сөздер саны көп болып, мазмұны артық ақпаратқа толы болады. Мұндай хабарламалардың негізгі мақсаты – пайдаланушының назарын аударып, оны белгілі бір әрекетке итермелеу. Осылайша, хабарлама ұзындығының талдауы мәтіндерді санаттарға бөлу процесінде маңызды факторлардың бірі ретінде қарастырылады.



Сурет 6. Хабарлама түрлеріне байланысты сөз санының таралуы

Зерттеуде қолданылатын модель бірнеше кезеңмен жүзеге асырылады:

1. Инициализация.

Алдымен QazNLTK кітапханасы іске қосылып, мәтінді токенизациялауға қажетті функциялар дайындалады:

```
From qaznltk import qaznltk as qtool
QazNLTK processor = qtool.QazNLTK()
```

2. Сандарды сөздерге ауыстыру.

Convert_numbers_to_words аталған функция мәтіндік деректер құрамындағы сандық мәндерді сөздік формаға түрлендіреді. Мысалы, *«3 мың теңге»* сандық түрде көрсетілген мән функция арқылы *«үш мың теңге»* сөздік формаға өзгереді. Бұл сөйлемдердің мәнмәтінін терең түсінуге мүмкіндік береді:

```
def numbers_to_text(message):
    tokens = message.split()
    updated_tokens = []
    for token in tokens:
        if token.isdigit():
            num_value = int(token)
            token = processor.num2word(num_value)
        updated_tokens.append(token)
    return ' '.join(updated_tokens)
```

3. Токенизация.

Жіберілген жабарлама түріне байланысты токендерге сыныптастырылады, яғни әрбір сөз жеке бөлік болып қарастырылады.

```
sentencesA = [
"Сіздің хабарламаңыз күдікті әрекеттерге байланысты.",
"Сіз ұтысқа ие болдыңыз! Алу үшін кішкене сома аударыңыз.",
"Сіздің достарыңыз көмек сұрауда. Оларға ақша керек.",
"Инвестиция салу бойынша ұсыныс бар."
]
```

4. Сәйкестікті тексеру

Жаңа хабарлама алдынала белгіленген күдікті мәтіндермен салыстырылады. Сәйкес сөздер саны мен сәйкестік пайызы есептеледі.

Тест нәтижелері.

Модель 100 мәтіндік хабарламада сыналды. Әр хабарлама 4 негізгі күдікті мәтінмен салыстырылды (Кесте 1).

Кесте 1. Хабарламалармен салыстыру нәтижелері

Хабарлама мәтіні	Ұқсастық деңгейі (%)	Ұқсас сөздер саны
Сіздің хабарламаңыз күдікті...	75.00	9
Сіз ұтысқа ие болдыңыз!...	50.00	6
Сіздің достарыңыз көмек сұрауда...	25.00	3
Инвестиция салу бойынша ұсыныс...	0.00	0

Талдау. Нәтижелер бойынша, модель «Сіздің хабарламаңыз күдікті әрекеттерге байланысты» мәтінімен 75% сәйкестік көрсетті, бұл оның күдікті хабарламаларды анықтауда тиімді екенін дәлелдейді.

Статистикалық бағалау. Confusion Matrix негізінде модельдің сапа көрсеткіштері есептелді (Кесте 2).

Кесте 2. Модельдің тиімділігін бағалау нәтижелері

Көрсеткіш	Мән
Дұрыстық деңгейі (Accuracy)	0.72
Қамту деңгейі (Recall)	0.70
Айқындығы (Precision)	0.65
F1 көрсеткіші (F1-Score)	0.67

Аталған көрсеткіштер модельдің хабарламаларды дұрыс жіктеуде және алаяқтық шабуылдарды болжауда әлеуеті бар екенін растайды.

Қорытынды. Зерттеу нәтижесінде мессенджерлер арқылы таратылатын мәтіндік хабарламаларды алаяқтықтан қорғауға арналған тиімді модель ұсынылды. QazNLTK негізіндегі жүйе мәтіндерді өңдеу, токенизациялау және салыстыру әдістерін қолданып, жақсы нәтижелер көрсетті.

Модельдің тиімділігі статистикалық көрсеткіштер негізінде дәлелденді. Жүйе «Сіздің хабарламаңыз күдікті әрекеттерге байланысты» хабарламасымен 75% сәйкестік көрсетті, бұл оның алаяқтықты анықтаудағы әлеуетін айқындайды. Confusion Matrix нәтижесінде алынған дұрыстық деңгейі (accuracy), қамту деңгейі (recall) және F1 көрсеткіші (F1-score) мәндері модельдің шынайы жағдайда жұмыс істеу қабілетін нақты көрсетті.

Алдағы уақытта модельді жетілдіру үшін бірнеше бағыт ұсынылады:

1. Деректер жиынтығын кеңейту: Алаяқтық хабарламалардың әртүрлі түрлері мен контекстін қамту арқылы модельдің үйрену мүмкіндігін күшейту.

2. Жасанды интеллект әдістерін енгізу: Машиналық оқыту мен терең оқыту (deep learning) алгоритмдерін қолдану модельдің тиімділігін арттыруға жол ашады.

3. Нақты уақыт режимінде сынақтан өткізу: Жүйені real-time форматында тестілеп, оның өнімділігін және жауап беру жылдамдығын бағалау.

Осы зерттеу мессенджерлердегі алаяқтық әрекеттерге қарсы қорғаныс жүйелерін дамытуға алғашқы негіз қалады. Алынған нәтижелер модельдің практикалық тұрғыда қолдануға жарамдылығын айқындап қана қоймай, болашақта оның тиімділігін жетілдіру үшін жаңа идеялар мен бағыттарға жол ашады. Алаяқтықпен күресте заманауи технологияларды пайдалану – өзекті әрі қажетті қадам, және бұл бағыттағы жұмыстардың алдағы уақытта зерттеулердің жалғасын табатынына сенім білдіріледі.

Әдебиеттер

qazaqstan.tv (2024) Интернет алаяқтық түрленіп барады. https://qazaqstan.tv/news/197676/7 (20 Қазан 2024 сілтеме алынды).

eGov.kz электрондық үкімет порталы (2024) Абайлаңыз, интернет-алаяқтар!. https://www.gov.kz/memleket/entities/sko-mamlyut-krasnoznamen/press/news/details/750707?lang=kk. (20 Қазан 2024 сілтеме алынды).

Lu H.Y., Chan S., Chai W., Lau S.M. & Khader M. (2020) Examining the influence of emotional arousal and scam preventive messaging on susceptibility to scams. Crime Prevention and Community Safety, 22(4). — P. 313-330. https://doi.org/10.1057/s41300-020-00098-3

Kigerl A. (2020) Spam-based scams. The Palgrave Handbook of International Cybercrime and Cyberdeviance. — P. 877-897. https://doi.org/10.1007/978-3-319-78440-3_42

Raximov Q., Mamatova Z. & Tazhikenova N. (2024). COMMON PHISHING ATTACKS IN KAZAKHSTAN AND WAYS TO PROTECT CITIZENS FROM INTERNET SCAMMERS. Farg'ona davlat universiteti, (3). — P. 130-130.

Lakbayev K.S., Rysmagambetova G.M., Umetov A.U. & Sysoyev A.K. (2020) The crimes in the field of high technology: Concept, problems and methods of counteraction in Kazakhstan. International Journal of Electronic Security and Digital Forensics, 12(4) . — P. 412-423.

Sarno D.M. & Black J. (2024) Who gets caught in the Web of lies?: Understanding susceptibility to phishing emails, fake news headlines, and scam text messages. Human Factors, 66(6) . — P. 1742-1753. https://doi.org/10.1177/00187208231173263

Hanoch Y. & Wood S. (2021). The scams among us: Who falls prey and why. Current Directions in Psychological Science, 30(3) . — P. 260-266. https://doi.org/10.1177/0963721421995489

Ajayi T.M. (2022) Discursive-manipulative strategies in scam emails and SMS: The Nigerian perspective. Lodz Papers in Pragmatics, 18(1) . — P. 175-195. DOI: 10.1515/lpp-2022-0008

Tran M.H., Hoai T.H. & Choo H. (2020) A third-party intelligent system for preventing call phishing and message scams. In Future Data and Security Engineering. Big Data, Security and Privacy, Smart City and Industry 4.0 Applications: 7th International Conference, FDSE 2020, Quy Nhon, Vietnam, November 25–27, 2020, Proceedings 7. — P. 486-492. Springer Singapore. https://doi.org/10.1007/978-981-96-0437-1

Li X., Rahmati A. & Nikiforakis N. (2024) Like, Comment, Get Scammed: Characterizing Comment Scams on Media Platforms. In Proceedings 2024 Network and Distributed System Security Symposium. https://dx.doi.org/10.14722/ndss.2024.24060

Chethiyar S.D.M., Vedamanikam M. & Sameem M.A. (2021) Losing The War Against Money Mule Recruitment: Persuasive Technique In Romance Scam. Ilkogretim Online, 20(3). — P. 2569-2585. DOI: 10.17051/ilkonline.2021.03.290

Ahmad R., Terzis S. & Renaud K. (2023) Content analysis of persuasion principles in mobile instant message phishing. In International Symposium on Human Aspects of Information Security and Assurance. — P. 324-336. Cham: Springer Nature Switzerland. DOI: 10.1007/978-3-031-38530-8_26

Wash R. (2020). How experts detect phishing scam emails. Proceedings of the ACM on Human-Computer Interaction, 4(CSCW2), 1-28. https://doi.org/10.1145/3415231

Liu M., Zhang Y., Liu B., Li Z., Duan H. & Sun D. (2021, December) Detecting and characterizing SMS spearphishing attacks. In Proceedings of the 37th Annual Computer Security Applications Conference. — P. 930-943. https://doi.org/10.1145/3485832.3488012

References

qazaqstan.tv (09 tamyz 2024). Internet alayaqtyq turlenip barady [Internet scams are transforming]. https://qazaqstan.tv/news/197676/7 (20 Qazan 2024 silteme alyndy) (in Kazakh)

eGov.kz elektrondyq ükimet portaly. (15 säuir 2024) Abaylañyz, internet-alayaqtar! [Caution, internet scammers!] https://www.gov.kz/memleket/entities/sko-mamlyut-krasnoznamen/press/news/details/750707?lang=kk (20 Qazan 2024 silteme alyndy) (in Kazakh)

Lu H.Y., Chan S., Chai W., Lau S.M. & Khader M. (2020) Examining the influence of emotional

arousal and scam preventive messaging on susceptibility to scams. Crime Prevention and Community Safety, 22(4). — P. 313-330 https://doi.org/10.1057/s41300-020-00098-3 (in English)

Kigerl A. (2020) Spam-based scams. The Palgrave Handbook of International Cybercrime and Cyberdeviance. — P. 877-897 https://doi.org/10.1007/978-3-319-78440-3_42 (in English)

Raximov Q., Mamatova Z. & Tazhikenova N. (2024) COMMON PHISHING ATTACKS IN KAZAKHSTAN AND WAYS TO PROTECT CITIZENS FROM INTERNET SCAMMERS. Farg'ona davlat universiteti, (3). — P. 130-130. (in English)

Lakbayev K.S., Rysmagambetova G.M., Umetov A.U. & Sysoyev A.K. (2020) The crimes in the field of high technology: Concept, problems and methods of counteraction in Kazakhstan. International Journal of Electronic Security and Digital Forensics, 12(4) . — P. 412-423. (in English)

Sarno D.M. & Black J. (2024) Who gets caught in the Web of lies?: Understanding susceptibility to phishing emails, fake news headlines, and scam text messages. Human Factors, 66(6). — P. 1742-1753 https://doi.org/10.1177/00187208231173263 (in English)

Hanoch Y. & Wood S. (2021) The scams among us: Who falls prey and why. Current Directions in Psychological Science, 30(3) . — P. 260-266. https://doi.org/10.1177/0963721421995489 (in English)

Ajayi T.M. (2022) Discursive-manipulative strategies in scam emails and SMS: The Nigerian perspective. Lodz Papers in Pragmatics, 18(1). — P. 175-195. DOI: 10.1515/lpp-2022-0008 (in English)

Tran M.H., Hoai T.H. & Choo H. (2020) A third-party intelligent system for preventing call phishing and message scams. In Future Data and Security Engineering. Big Data, Security and Privacy, Smart City and Industry 4.0 Applications: 7th International Conference, FDSE 2020, Quy Nhon, Vietnam, November 25–27, 2020, Proceedings 7. — P. 486-492. Springer Singapore. https://doi.org/10.1007/978-981-96-0437-1 (in English)

Li X., Rahmati A. & Nikiforakis N. (2024) Like, Comment, Get Scammed: Characterizing Comment Scams on Media Platforms. In Proceedings 2024 Network and Distributed System Security Symposium. https://dx.doi.org/10.14722/ndss.2024.24060 (in English)

Chethiyar S.D.M., Vedamanikam M. & Sameem M.A. (2021) Losing The War Against Money Mule Recruitment: Persuasive Technique In Romance Scam. Ilkogretim Online, 20(3). — P. 2569-2585. DOI: 10.17051/ilkonline.2021.03.290 (in English)

Ahmad R., Terzis S. & Renaud K. (2023) Content analysis of persuasion principles in mobile instant message phishing. In International Symposium on Human Aspects of Information Security and Assurance. — P. 324-336. Cham: Springer Nature Switzerland. DOI: 10.1007/978-3-031-38530-8_26 (in English)

Wash R. (2020). How experts detect phishing scam emails. Proceedings of the ACM on Human-Computer Interaction, 4(CSCW2), 1-28. https://doi.org/10.1145/3415231 (in English)

Liu M., Zhang Y., Liu B., Li Z., Duan H. & Sun D. (2021, December) Detecting and characterizing SMS spearphishing attacks. In Proceedings of the 37th Annual Computer Security Applications Conference. — P. 930-943. https://doi.org/10.1145/3485832.3488012 (in English)

ACADEMIC SCIENTIFIC JOURNAL OF COMPUTER SCIENCE
ISSN 1991-346X
Volume 4. Namber 356 (2025). 152–170

https://doi.org/10.32014/2025.2518-1726.389

UDC 533.9.01; 519.63; 519.684
IRSTI 27.35.21; 27.35.47; 27.41.19

© A. Issakhov[1]*, A. Alzhanov[2], A. Akhmedov[2], A. Amanzholov[2], T. Murat[2], 2025.

[1] Kazakh-British Technical University, Almaty, Kazakhstan;
[2] International school Haileybury, Almaty, Kazakhstan.
*E-mail: alibek.issakhov@gmail.com

NUMERICAL SIMULATION OF THERMOHYDRODYNAMICS DURING HEATED WATER DISCHARGE INTO LAKE BALKHASH

Issakhov Alibek — PhD, Professor of the School of Applied Mathematics at Kazakh-British Technical University, Almaty, Kazakhstan,
E-mail: Alibek.issakhov@gmail.com, https://orcid.org/0000-0002-1937-8615;
Alzhanov Alibek — student at International school Haileybury, Almaty, Kazakhstan,
email: Alzhanov.Alibek @Haileybury.com, https://orcid.org/0000-1002-1938-8715;
Akhmedov Ali — student at International school Haileybury, Almaty, Kazakhstan,
E-mail: Akhmedov.Ali @Haileybury.com, https://orcid.org/0000-1002-1938-87156;
Amanzholov Amirlan – student at International school Haileybury, Almaty, Kazakhstan.
E-mail: Amanzholov.Amirlan @Haileybury.com, https://orcid.org/0000-1002-1938-87157;
Murat Timur — student at International school Haileybury, Almaty, Kazakhstan,
E-mail: Murat.Timur @Haileybury.com, https://orcid.org/0000-1002-1938-87158.

Abstract. This paper presents a numerical study of thermal pollution from a potential discharge of heated water from a nuclear power plant's cooling system into the coastal zone of Lake Balkhash. The aim of the study was to determine the dynamics of the formation and spatial propagation of the thermal footprint, as well as to assess its impact on the hydrodynamic characteristics of the water area. A two-dimensional computational fluid dynamic (CFD) model based on the Navier-Stokes and energy equations, supplemented by a k-ω SST turbulence model, was used for the simulation. The simulation was conducted using the geometry of a real coastline. The model was validated using a test jet injection problem, which demonstrated good agreement with experimental data. The experimental data were taken from the work of other authors. Modeling results showed that within the first hour after the discharge begins, a stable zone of thermal pollution, approximately 1.5 km² in area, forms, where water temperatures exceed background values by 0.5–1.0 K. The high-temperature plume gradually transforms into a diffuse heat spot, which persists in the coastal zone due to weak water exchange and recirculation. Velocity field analysis revealed zones of local circulation that facilitate long-term

heat retention. The scientific novelty of the study lies in the use of computational fluid dynamics (CFD) to assess the thermohydrodynamics of discharge under conditions of limited water exchange in a large lake with the actual morphology of the shoreline of Lake Balkhash. The practical significance lies in the potential use of the developed method for predicting environmental risks and optimizing the cooling systems of nuclear and thermal power plants located near the lake.

Keywords: thermal pollution, CFD modeling, Lake Balkhash, heated waters, SST k–ω model, hydrodynamics, environmental safety

© А.А. Исахов[1]*, А. Альжанов[2], А. Ахмедов[2], А. Аманжолов[2], Т. Мурат[2], 2025.

[1]Қазақ-Британ техникалық университеті, Алматы, Қазақстан;
[2]Хейлибери халықаралық мектебі, Алматы, Қазақстан.
*E-mail: Alibek.issakhov@gmail.com

БАЛҚАШ КӨЛІНЕ ЖЫЛЫ СУ АҒЫЗУ КЕЗІНДЕГІ ТЕРМОГИДРОДИНАМИКАНЫ САНДЫҚ МОДЕЛЬДЕУ

Исахов Алибек — PhD докторы, Қазақ-Британ техникалық университеті Қолданбалы математика мектебінің профессоры, Алматы, Қазақстан,
E-mail: Alibek.issakhov@gmail.com, https://orcid.org/0000-0002-1937-8615;
Альжанов Алибек — Haileybury халықаралық мектебінің студенті, Алматы, Қазақстан,
E-mail: Alzhanov.Alibek @Haileybury.com, https://orcid.org/0000-1002-1938-8715;
Ахмедов Али — Халықаралық Haileybury мектебінің студенті, Алматы, Қазақстан,
E-mail: Akhmedov.Ali @Haileybury.com, https://orcid.org/0000-1002-1938-87156;
Аманжолов Амирлан — Haileybury халықаралық мектебінің студенті, Алматы, Қазақстан,
email: Amanzholov.Amirlan @Haileybury.com, https://orcid.org/0000-1002-1938-87157;
Мурат Тимур — Халықаралық Haileybury мектебінің студенті, Алматы, Қазақстан,
email: Murat.Timur @Haileybury.com, https://orcid.org/0000-1002-1938-87158.

Аннотация. Бұл мақалада атом электр станциясының салқындату жүйесінен Балқаш көлінің жағалау аймағына қыздырылған судың ықтимал ағызылуынан термиялық ластанудың сандық зерттеуі ұсынылған. Зерттеудің мақсаты жылулық іздің пайда болу динамикасын және кеңістіктік таралуын анықтау, сондай-ақ оның су айдынының гидродинамикалық сипаттамаларына әсерін бағалау болды. Модельдеу үшін Навье-Стокс және энергия теңдеулеріне негізделген екі өлшемді есептеу сұйықтық динамикасы (CFD) моделі қолданды, қосымша k-ω SST турбуленттілік моделімен толықтырылды. Модельдеу нақты жағалау сызығының геометриясын қолдана отырып жүргізілді. Модель эксперименттік деректермен жақсы сәйкестікті көрсеткен сынақ ағынды инжекция есебін қолдана отырып тексерілді. Ал эксперименттік деректер басқа авторлардың жұмыстарынан алынды. Модельдеу нәтижелері ағызылу басталғаннан кейінгі алғашқы сағат ішінде су температурасы фондық мәндерден 0.5–1.0 К асатын ауданы шамамен 1.5 км² болатын тұрақты термиялық ластану аймағы пайда болатынын көрсетті. Сондай-ақ жоғары

температуралы шлейф біртіндеп су алмасуының және қайта айналымының әлсіздігіне байланысты жағалау аймағында сақталатын диффузиялық жылу нүктесіне айналады. Жылдамдық өрісін талдау ұзақ мерзімді жылуды сақтауға ықпал ететін жергілікті айналым аймақтарын анықтады. Зерттеудің ғылыми жаңалығы Балқаш көлінің жағалау сызығының нақты морфологиясымен бірге үлкен көлде шектеулі, ал су алмасу жағдайында ағынды сулардың термогидродинамикасын бағалау үшін есептеу сұйықтық динамикасын (CFD) пайдалануда жатыр. Тәжірибелік маңыздылығы көлдің жанында орналасқан атом және жылу электр станцияларының салқындату жүйелерін экологиялық тәуекелдерді болжау және оңтайландыру үшін әзірленген әдісті қолдану мүмкіндігінде жатыр.

Түйін сөздер: термиялық ластану, CFD модельдеу, Балқаш көлі, жылытылған сулар, SST k–ω моделі, гидродинамика, экологиялық қауіпсіздік

© А.А. Исахов[1]*, А. Альжанов[2], А. Ахмедов[2], А. Аманжолов[2], Т. Мурат[2], 2025.

[1]Казахстанско-Британский технический университет, Алматы, Казахстан;
[2]Международной школы Haileybury, Алматы, Казахстан.
*E-mail: Alibek.issakhov@gmail.com

ЧИСЛЕННОЕ МОДЕЛИРОВАНИЕ ТЕРМОГИДРОДИНАМИКИ ПРИ СБРОСЕ ПОДОГРЕТЫХ ВОД В ОЗЕРО БАЛХАШ

Исахов Алибек — доктор PhD, профессор Школы прикладной математики Казахстанско-Британского технического университета, Алматы, Казахстан,
E-mail: Alibek.issakhov@gmail.com, https://orcid.org/0000-0002-1937-8615;
Альжанов Алибек — студент Международной школы Haileybury, Алматы, Казахстан,
E-mail: Alzhanov.Alibek @Haileybury.com, https://orcid.org/0000-1002-1938-8715;
Ахмедов Али — студент Международной школы Haileybury, Алматы, Казахстан,
E-mail: Akhmedov.Ali @Haileybury.com, https://orcid.org/0000-1002-1938-87156;
Аманжолов Амирлан — студент Международной школы Haileybury, Алматы, Казахстан,
E-mail: Amanzholov.Amirlan @Haileybury.com, https://orcid.org/0000-1002-1938-87157;
Мурат Тимур — студент Международной школы Haileybury, Алматы, Казахстан,
E-mail: Murat.Timur @Haileybury.com, https://orcid.org/0000-1002-1938-87158.

Аннотация. В данной работе представлено численное исследование теплового загрязнения от возможного сброса нагретой воды из системы охлаждения атомной электростанции в прибрежную зону озера Балхаш. Целью исследования являлось определение динамики формирования и пространственного распространения теплового следа, а также оценка его влияния на гидродинамические характеристики акватории. Для моделирования использовалась двумерная вычислительная гидродинамическая модель (CFD), основанная на уравнениях Навье-Стокса и энергии, дополненная k-ω SST-моделью турбулентности. Моделирование проводилось на основе геометрии реальной береговой линии. Валидация модели была проведена

на тестовой задаче инжекции струи, которая показала хорошее согласие с экспериментальными данными. Экспериментальные данные были взяты с работ других авторов. Результаты моделирования показали, что в течение первого часа после начала сброса формируется устойчивая зона теплового загрязнения площадью около 1,5 км², где температура воды превышает фоновые значения на 0,5–1,0 К. Высокотемпературный шлейф постепенно трансформируется в диффузное тепловое пятно, сохраняющееся в прибрежной зоне благодаря слабому водообмену и рециркуляции. Анализ поля скорости выявил зоны локальной циркуляции, способствующие длительному сохранению тепла. Научная новизна исследования заключается в применении вычислительной гидродинамики (CFD) для оценки термогидродинамики сброса в условиях ограниченного водообмена в крупном озере с реальной морфологией береговой линии озера Балхаш. Практическая значимость заключается в возможности использования разработанного метода для прогнозирования экологических рисков и оптимизации систем охлаждения атомных и тепловых электростанций, расположенных вблизи озера.

Ключевые слова: тепловое загрязнение, CFD-моделирование, озеро Балхаш, подогретые воды, SST k–ω модель, гидродинамика, экологическая безопасность

Introduction. Growing energy production and the development of industrial complexes are increasing the pressure on aquatic ecosystems, especially in regions with limited water resources. Thermal pollution, caused by the discharge of heated water from power plants, remains one of the most significant anthropogenic impacts. Rising temperatures alter the thermal and hydrodynamic regimes of water bodies, affecting dissolved oxygen levels, biochemical processes, and ecosystem structure. In large lakes and reservoirs, localized heating can lead to an extension of the stratification period, increased evaporation, and redistribution of matter and energy flows. Quantifying these effects requires highly accurate models capable of reproducing the spatiotemporal dynamics of temperature fields, taking into account the actual shoreline geometry and hydrological features.

A number of studies have assessed the impacts of thermal pollution and developed numerical models describing the dynamics of thermal discharges. Råman Vinnå et al. (2017) showed that discharges of heated water from nuclear power plants into lakes cause a temperature increase of up to 3 °C and an extension of the stratification period, which significantly alters exchange processes and heat fluxes. Gaudard et al. (2018) noted that the use of rivers and lakes as heat-exchange reservoirs should be considered a form of anthropogenic impact that has long-term effects on hydrological and biological systems. Caissie (2006) emphasized the importance of understanding river thermal regimes, noting that deterministic models using energy-efficient approaches are crucial for predicting changes in water temperature due to anthropogenic impacts, including thermal pollution. These studies highlight the need for an integrated approach to modeling thermal processes in real water bodies, followed by an assessment of their environmental impacts.

Modern research employs a wide range of computational methods to describe the propagation of heated waters. Issakhov (2019) implemented a two-dimensional CFD model based on the Navier–Stokes and energy equations to analyze the formation of thermal pollution zones under the influence of power plant discharges. The model was successfully verified using field data, confirming its reliability in predicting the scale of thermal impact. Al-Suhaili (2006) developed a similar approach for river flows with a point heat source, taking into account unsteady transport and dissipation processes. Tang (2008) proposed a three-dimensional CFD model taking into account complex coastal geometry and turbulent effects, which allowed for obtaining a detailed structure of temperature and velocity fields in coastal areas. Thus, computational fluid dynamics provides the most accurate and spatially resolved description of the interaction of hydrodynamic and thermal processes.

Particular attention is paid to the influence of discharge parameters on the formation of temperature traces. Mahmood and Mohammad (2022) studied the propagation of a thermal plume along a section of the Tigris River using COMSOL Multiphysics. It was shown that reducing the discharge velocity from 1.6 to 0.5 m/s promotes more uniform mixing and reduces temperature anomalies. Kalinowska et al. (2012) obtained similar results for the Vistula River, where optimizing the discharge flow and direction minimized localized overheating zones.

In recent years, research aimed at an integrated description of thermohydrodynamic and ecological processes in aquatic ecosystems has gained significant momentum. Du et al. (2025) showed that the volume of thermal discharge has a stronger effect on the extent of the thermal footprint than the heating temperature, emphasizing the need for parametric optimization of cooling scenarios. Velázquez-Araque et al. (2025) demonstrated the versatility of CFD approaches by applying them to the analysis of microplastic dispersion using a multiphase turbulent model in ANSYS Fluent; the proposed approach was also applicable to thermal dispersion problems. Tarena et al. (2024) conducted a systematic review of the use of CFD in hydroecological studies and showed that the most widely used methods remain RANS and LES models implemented in the ANSYS Fluent and FLOW-3D software packages.

In a global context, Raptis et al. (2016) found that the world's largest rivers, including the Mississippi and Rhine, experience significant increases in water temperature due to energy discharges, with thermal energy dispersed tens of kilometers downstream. Miara et al. (2018) extended these findings by using 3D CFD modeling to analyze the impact of heat fluxes on energy system infrastructure and the sustainability of their operations. According to Råman Vinnå et al. (2017), in lake systems, up to 60% of excess heat is removed through spillways, while the remainder alters internal convective structures, increasing ecosystem sensitivity to anthropogenic impacts.

Modern CFD approaches enable highly accurate analysis of the spatial structure of temperature fields, assessment of mixing efficiency, and prediction of the impact of thermal emissions at the local and regional levels. In this context, numerical

modeling of thermal pollution within Lake Balkhash—a large transboundary ecosystem in Central Asia characterized by a complex coastal configuration and high sensitivity to external influences—is of particular interest. Studying the propagation of heat flow processes arising from the potential construction of a nuclear power plant in the coastal zone allows for a quantitative assessment of the scale of thermal impact and the identification of critical zones of thermal regime change.

Materials and Methods. Geographical and hydrological characteristics of the study area

The study was conducted on the coastal section of Lake Balkhash, located near the village of Ulken on the southwestern shore of the reservoir (Figure 1). This area is considered a potential intake and discharge zone for cooling water during the operation of the nuclear power plant. Geographically, it is a closed part of the lake, partially separated from the main body of water by hydraulic structures (dams and bridges), which limits water exchange and creates a relatively isolated ecosystem.

The main geometric parameters of the study area are: length 7.6 km, average width 3.8 km, area 16.07 km², and shoreline perimeter approximately 22.6 km. According to the International Lakes Environment Committee (ILEC), Balkhash is a warm lake with pronounced seasonal variability in its thermal regime: surface temperatures range from 0°C in December to 28°C in July, with a long-term average of approximately 10°C in the western part and 9°C in the eastern part. Freezing occurs annually from November to March, with ice persisting for 10–15 days longer in the eastern part.

Using this coastal section of the lake as part of the nuclear power plant's cooling system could lead to a number of hydrodynamic and environmental consequences. Under conditions of limited water exchange, the heated jet stream can create persistent water temperature anomalies reaching 3–9°C relative to background values. Such temperature increases potentially alter vertical circulation and stratification processes, reduce dissolved oxygen concentrations, and may disrupt the natural thermal and biochemical balance of the ecosystem.

Figure 1. Top view of the study area of Lake Balkhash

Mathematical Model. To numerically describe the currents and temperature distribution in the studied reservoir, the equations for a non-stationary viscous incompressible fluid in the Navier-Stokes form with an added energy equation were used.

– continuity equation

$$\frac{\partial u_i}{\partial x_i} = 0 \tag{1}$$

– momentum equation

$$\frac{\partial u_i}{\partial t} + u_i \frac{\partial u_i}{\partial x_j} = -\frac{1}{\rho}\frac{\partial P}{\partial x_i} + \upsilon \frac{\partial^2 u_i}{\partial x_j^2} \tag{2}$$

– pollution transfer equation

$$\frac{\partial T}{\partial t} + u_j \frac{\partial \mathrm{T}}{\partial x_j} = \alpha \frac{\partial^2 T_i}{\partial x_j^2} \tag{3}$$

where u_i – are the components of the velocity vector, T is the temperature, υ is the molecular viscosity, α is the thermal conductivity coefficient, ρ is the density, p is the pressure.

After Reynolds averaging of the Navier–Stokes equations, additional terms describing the momentum transfer by turbulent fluctuations appear in the momentum equations. To close these terms, the Boussinesq hypothesis is used, linking these terms to the mean velocity gradients through the turbulent viscosity υ,, calculated from the turbulence model equations. The effective viscosity $\upsilon_{eff} = \upsilon + \upsilon_t$ is used in the equations for averaged fields, ensuring the correct accounting for the influence of turbulent eddies. In the energy equation, turbulent heat transfer is accounted for through an additional diffusion term υ_t / Pr_t , analogous to the effect of turbulent viscosity in the momentum equation.

To describe turbulence, the SST k–ω (Shear Stress Transport) model was used, combining the advantages of the standard k–ε and k–ω models. It provides a stable and physically correct solution near rigid boundaries and in regions with weakly turbulent coastal currents.

The model is based on solving two transport equations—one for the kinetic energy of turbulence *k* and the specific rate of its dissipation ω:

$$\frac{\partial k}{\partial t} + \bar{u}_j \frac{\partial u_j k}{\partial x_j} = P_k - \beta^* k\omega + \frac{\partial}{\partial x_j}\left((\mu + \sigma_k \mu_t)\frac{\partial k}{\partial x_j}\right) \tag{4}$$

$$\frac{\partial \omega}{\partial t} + \bar{u}_j \frac{\partial u_j \omega}{\partial x_j} = \alpha \frac{\omega}{k} P_k - \beta\omega^2 + \frac{\partial}{\partial x_j}\left((\upsilon + \sigma_\omega \upsilon_t)\frac{\partial \omega}{\partial x_j}\right) + 2(1 - f_1)\frac{\sigma_{\omega 2}}{\omega}\frac{\partial k}{\partial x_j}\frac{\partial \omega}{\partial x_j} \tag{5}$$

where P_k – is the generation of turbulent energy caused by the velocity shift.

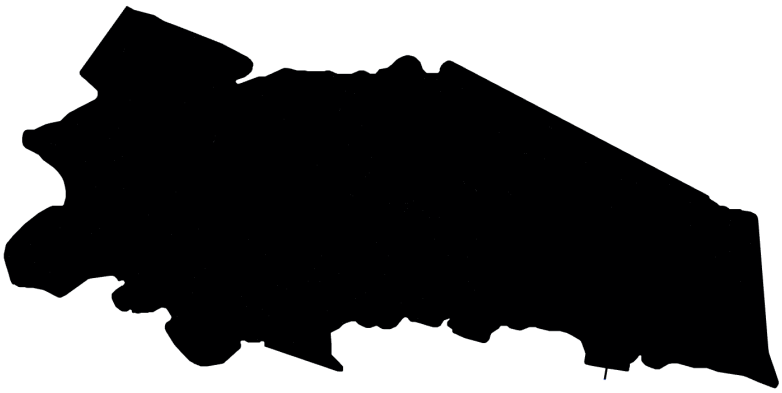
Boundary and Initial Conditions. At the jet outlet boundary, a constant velocity of 3.89 m/s and a temperature of 297 K are specified. The shoreline is treated as a stationary rigid wall with a no-slip condition (u = 0) and a zero temperature gradient along the normal. Symmetry conditions are used at the remaining boundaries.

The initial water temperature in the lake is assumed to be 291 K, and the initial flow velocity is 0 m/s. These values correspond to the state of a calm lake before the discharge begins. The same initial and boundary conditions were used for all simulations, ensuring a fair comparison of the scenarios.

Since the simulated domain is a closed coastal section of the lake with no distinct inlet or outlet, zero gradient conditions for all parameters were used at the outer boundaries. The propagation of the heated water occurred due to the inertia of the jet and turbulent mixing within the computational domain.

Computational Grid Construction and Independence Test. A digital copy of the coastline, created using satellite and cartographic data, was used to construct the geometry and grid (Figure 1). The computational domain was discretized using an unstructured triangular grid with localized densification in the jet outlet zone and in the warm flow dilution zone (Figure 2).

The overall cell size in the global part of the reservoir was 6–7 m, while densifications of 0.15–2.1 m were used near the jet source. The radius of the densification zone was 1400 m from the discharge center. Three grid variants with varying resolutions were created (Table 1). To verify the independence of the solution from grid size, identical physical models, boundary conditions, and time step (0.05 s) were used. The convergence criterion for all parameters was set at 10^{-6}.



(a)

(b)

(c)

Figure 2. Computational grid

Table 1. Parameters of computational grids for convergence studies

Mesh	Global size [m]	Maximum size in the clustering zone [m]	Minimal size in the clustering zone [m	Number of nodes	Number of cells
1	7 m	2.3 m	0.16 m	3 395 393	1 695 132
2	7 m	2.2 m	0.15 m	3 982 539	1 988 612
3	6 m	2.1 m	0.14 m	5 683 954	2 838 985

Figure 3 shows the average temperature distribution along the longitudinal profile for each mesh. As can be seen, when moving from the coarse mesh (mesh 1) to the more detailed meshes (mesh 2 and mesh 3), the shape of the temperature profile changes only slightly, indicating that mesh convergence has been achieved. The difference in maximum temperatures between meshes 2 and 3 does not exceed 1.5%, suggesting that mesh 3 is sufficient for subsequent calculations.

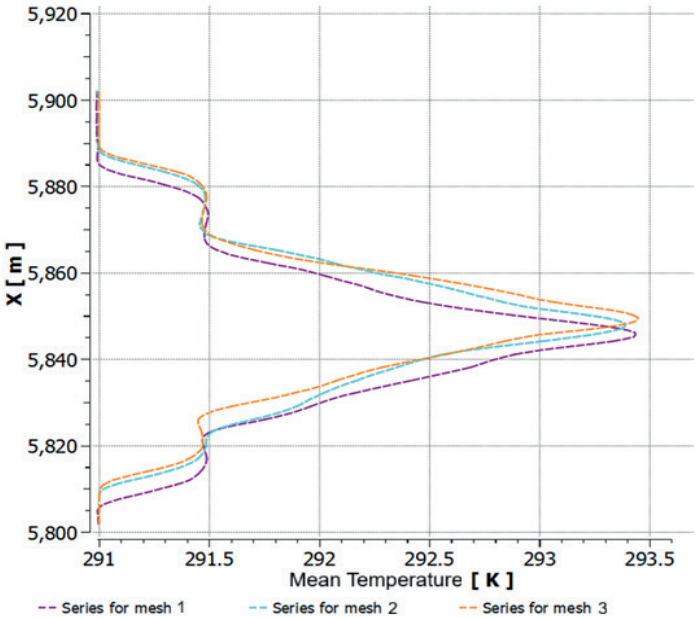


Figure 3. Distribution of average temperature along the x-axis for three variants of computational grids (mesh 1–3).

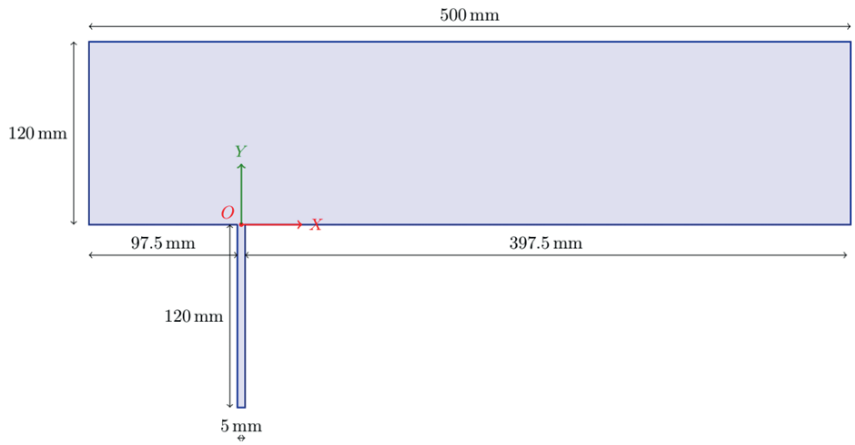
Numerical Implementation. For flow modeling, the equations for conservation of mass, momentum, and energy for a two-dimensional, steady, incompressible flow were used. The calculations were performed using the finite volume method, with the SIMPLE algorithm for coupling pressure and velocity. Temporal discretization was performed using an implicit first-order scheme. Convective terms were approximated by a second-order approximation using the upwind scheme, ensuring stability and accuracy of the solution while preserving sharp temperature gradients.

To speed up the computations, the problem was parallelized and run on 32 processors. The simulation time interval was 1 hour, and the results were saved every 10 minutes for analysis of the evolution of the temperature and velocity fields.

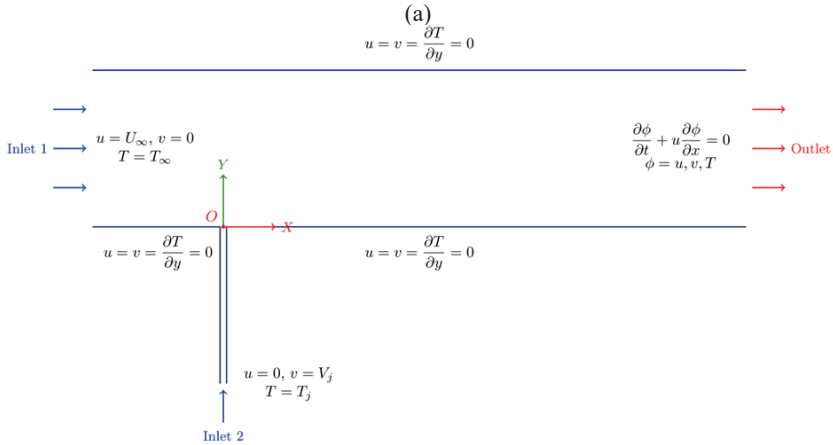
Model Validation. To verify the correctness of the mathematical model and numerical algorithm used, validation was performed using a benchmark problem: a heated jet injected into a horizontal closed channel. This configuration was chosen in accordance with the experimental studies of Chen and Hwang (1991) and the numerical analysis of Chang and Chen (1995), allowing the obtained results to be compared with reliable laboratory measurements.

Geometry and Problem Conditions

The computational domain (Figure 4a) consisted of a rectilinear channel 500 mm long and 120 mm high, with a single hot fluid injection point through a nozzle with a width of D = 5 mm, located 97.5 mm from the inlet. The nozzle center was taken as the origin, the X-axis was directed along the channel, and the Y-axis was directed vertically upward.



(a)



(b)

Figure 4 - Computational domain (a) and boundary conditions (b) of the test problem

The boundary conditions are shown in Figure 4b. The main flow temperature (Inlet 1) was T_∞ = 298 K (25 °C), and the temperature of the heated water entering through the nozzle (Inlet 2) was T_j = 354 K (81 °C). The main flow velocity was U_∞ = 2.005 m/s, and the jet injection velocity was V_j = 11.969 m/s.

No-slip conditions (u = v = 0) and adiabatic boundary conditions for temperature ($\partial T/\partial n = 0$) were used on the channel walls. Neumann-type boundary conditions were applied at the channel outlet, ensuring a free flow without reflections.

Comparison with Experiment. The numerical results (Figure 5) were compared with experimental data in five cross sections at distances X/D

= 2, 4, 6, 8, and 10 from the nozzle. To compare temperature profiles, the normalized value was used $Q = \frac{T - T_\infty}{T_j - T_\infty}$.

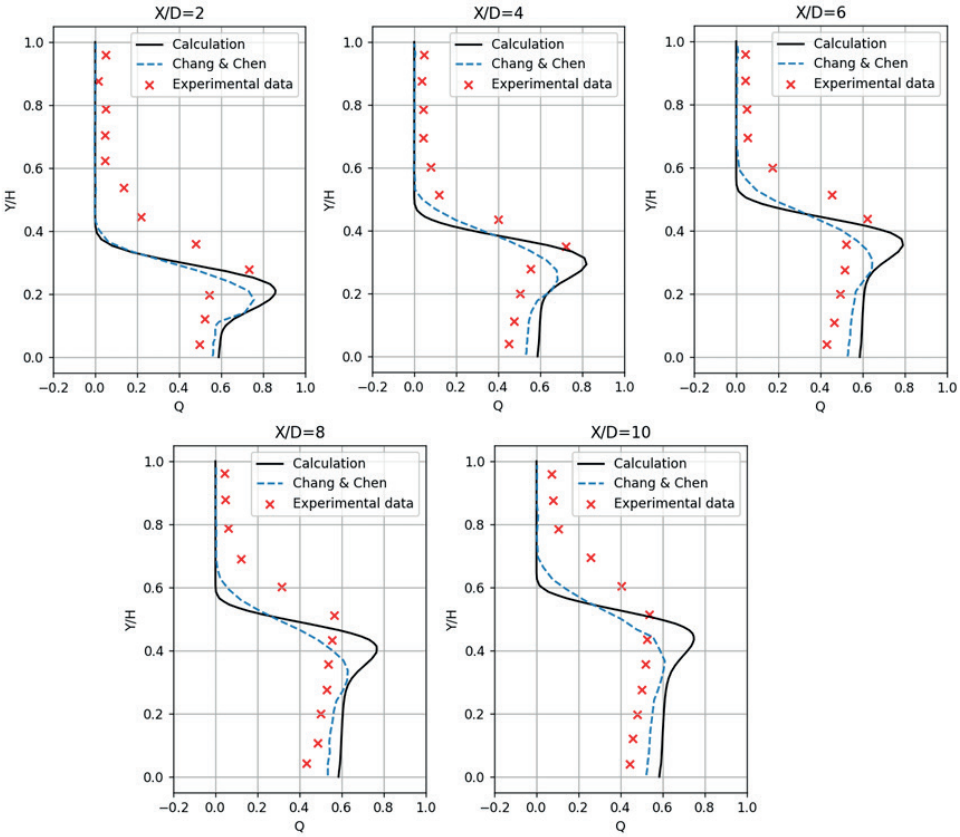


Figure 5 - Comparison of results with the experiment of Chen and Hwang (1991) [14] and the numerical analysis of Chang and Chen (1995) [15]

A comparison of the profiles showed that the obtained calculated data were in good agreement with the experiment: maximum deviations in temperature values did not exceed 5–7%, and the jet profile shape and isotherm positions matched the measurements. The SST k–ω model accurately reproduced both thermal diffusion and the jet trajectory in the crossflow, demonstrating high accuracy near the symmetry axis and in the region of intense mixing.

Results and Discussion. Numerical modeling allowed us to study in detail the spatiotemporal dynamics of the formation and propagation of a thermal wake from a potential discharge of heated waters in the coastal zone of Lake Balkhash over the course of one hour. A combined analysis of temperature and velocity fields (Figures 6a–f, 7a–f) demonstrates the consistent evolution of the thermal plume and corresponding changes in the hydrodynamic structure of the flow.

At the initial stage of the process (t = 10 min), the formation of a compact, high-temperature jet is observed, caused by the discharge of water with a temperature of 297 K. The core of the jet, with a temperature above 296.4 K, has a distinct directional structure and extends over more than 500 meters from the source, reflecting the dominant influence of the discharge impulse in the first minutes of the flow.

Over time (t = 20–40 min), the structure of the thermal field changes significantly. The inertial phase gives way to a diffusion phase: the role of turbulent heat transfer and advection increases, leading to a gradual expansion of the elevated temperature zone. As a result, the temperature field equalizes and coastal waters are involved in the warming process.

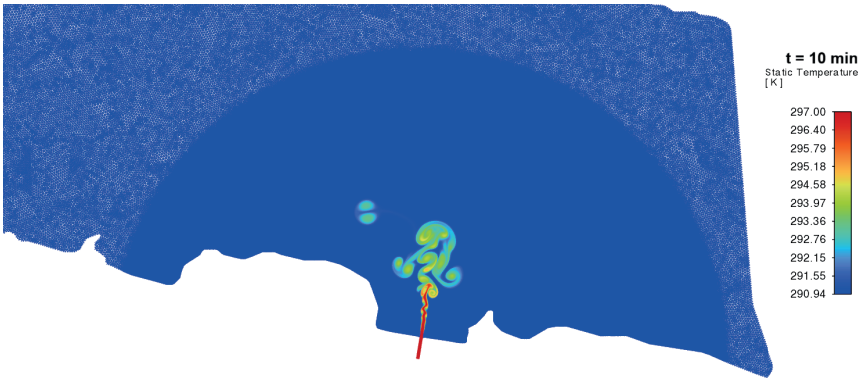
By the 50th minute of the simulation, the formation of the thermal plume enters a quasi-stationary phase. The temperature field stabilizes, as evidenced by the virtually identical temperature distributions at 50 and 60 minutes. By the 60th minute of the simulation, the warm jet loses its compactness, and the thermal plume covers a significant portion of the water area.

The calculated estimate shows that by the end of the 60-minute simulation interval, the area of water subject to noticeable thermal effects ($\Delta T \geq 0.5$ K) is approximately 1.5 km², which corresponds to approximately 14% of the area of the study area (10.5 km²). The average water temperature within the calculated region increases to 291.25 K, reflecting an overall increase in thermal background even over the short period of the release.

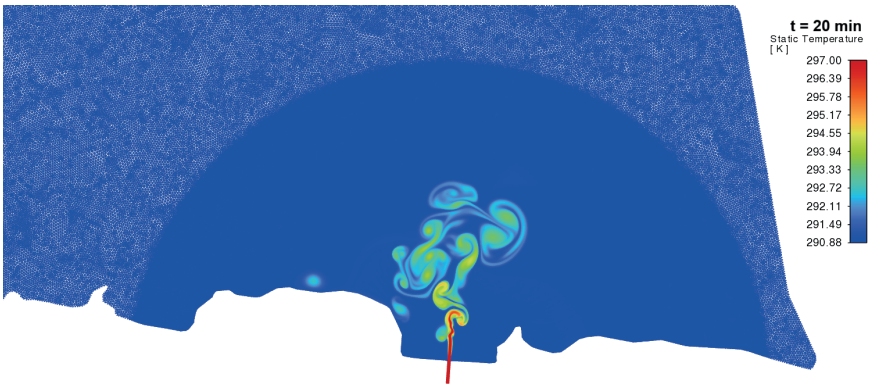
Flow Hydrodynamics. Analysis of the velocity field revealed that during the first 30 minutes of the simulation, the maximum velocity in the jet core gradually decreases – from 4.51 m/s (t = 10 min) to 4.15 m/s (t = 30 min). This is due to the development of turbulent mixing and the interaction of the jet with the surrounding waters. By the 40th minute, a localized increase in velocity to 5.03 m/s is recorded, likely due to flow reflection from the shoreline and the formation of localized vortex structures.

In the final stage (t = 50–60 min), the hydrodynamic structure stabilizes. The maximum velocity in the core decreases from 4.38 m/s to 4.32 m/s, reflecting the establishment of a balance between jet inertia and energy dissipation. The area with elevated velocities (≥ 2.5 m/s) forms a stable corridor aligned with the direction of the main thermal plume. In the peripheral areas of the water area, velocities do not exceed 0.5–1.0 m/s, creating conditions for local stagnation and heat accumulation.

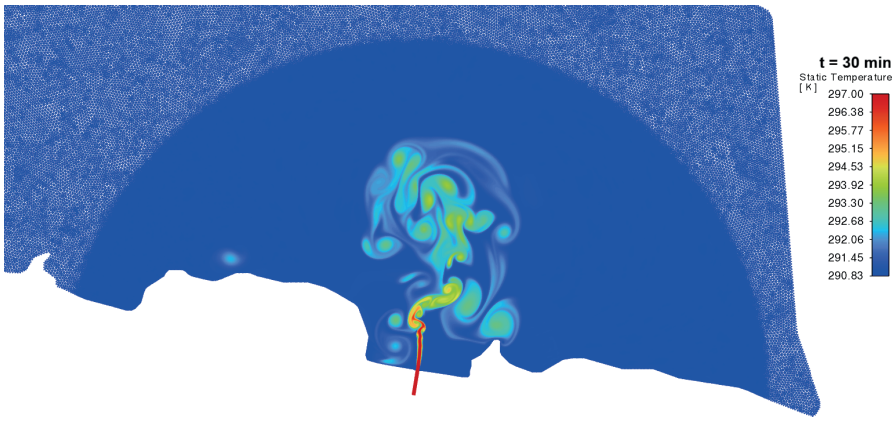
Thus, the results of the full hourly modeling demonstrate that, with a weak background current and complex coastal geometry, the discharge of heated waters forms a vast and stable zone of thermal pollution. Within the first hour, the system reaches a quasi-steady state, in which thermal impact becomes a constant environmental factor. The observed structure of the temperature and velocity fields indicates the potential for long-term changes in the local thermal regime, including disruption of seasonal stratification, a decrease in dissolved oxygen concentration, and the transformation of aquatic biocenoses.


t = 10 min
Static Temperature
[K]
297.00
296.40
295.79
295.18
294.58
293.97
293.36
292.76
292.15
291.55
290.94

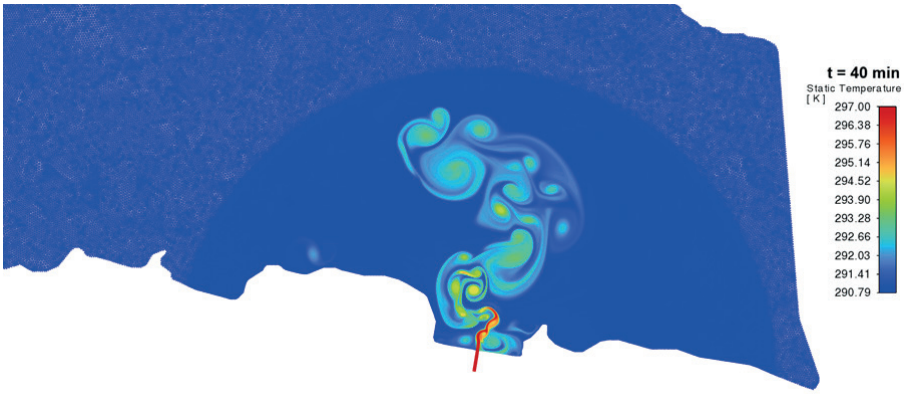
(a)


t = 20 min
Static Temperature
[K]
297.00
296.39
295.78
295.17
294.55
293.94
293.33
292.72
292.11
291.49
290.88

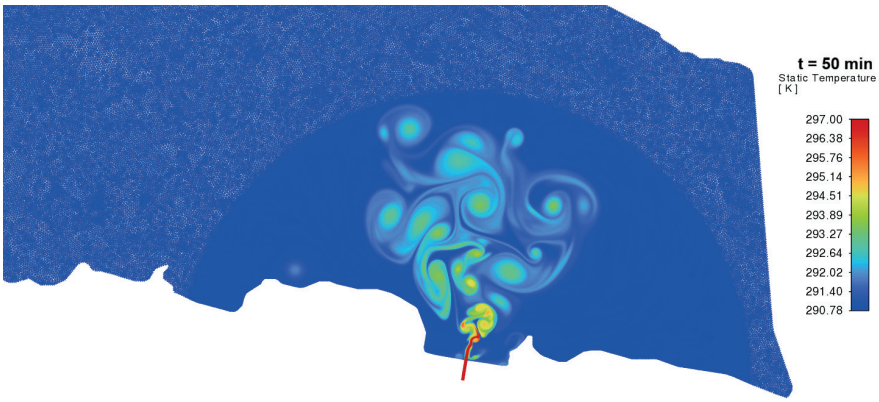
(b)


t = 30 min
Static Temperature
[K]
297.00
296.38
295.77
295.15
294.53
293.92
293.30
292.68
292.06
291.45
290.83

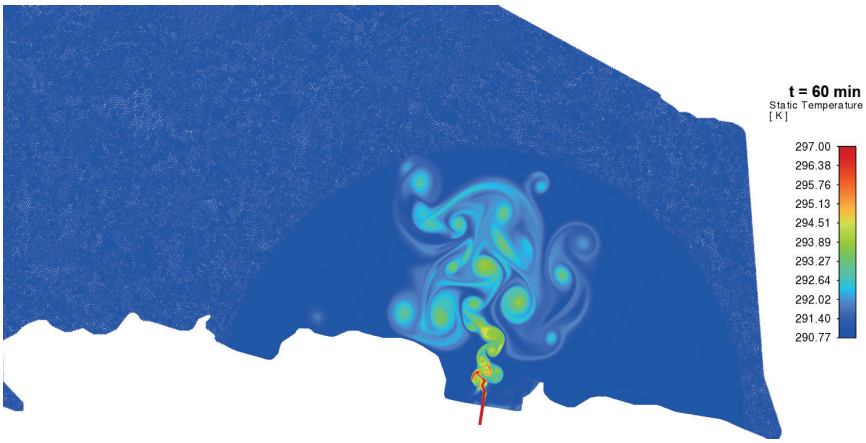
(c)



(d)

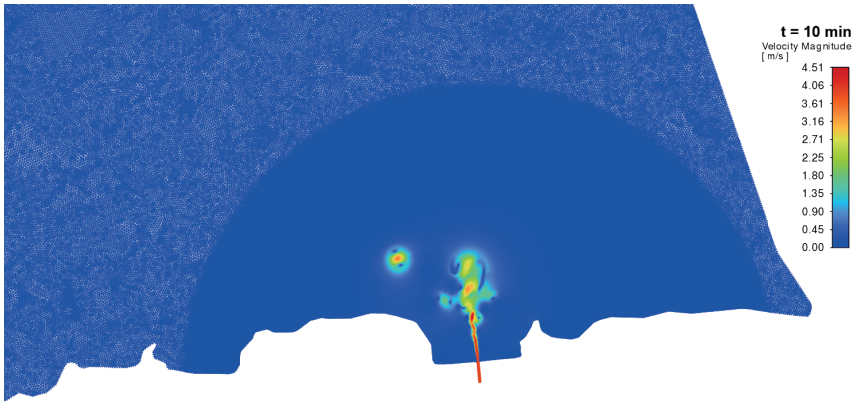


(e)

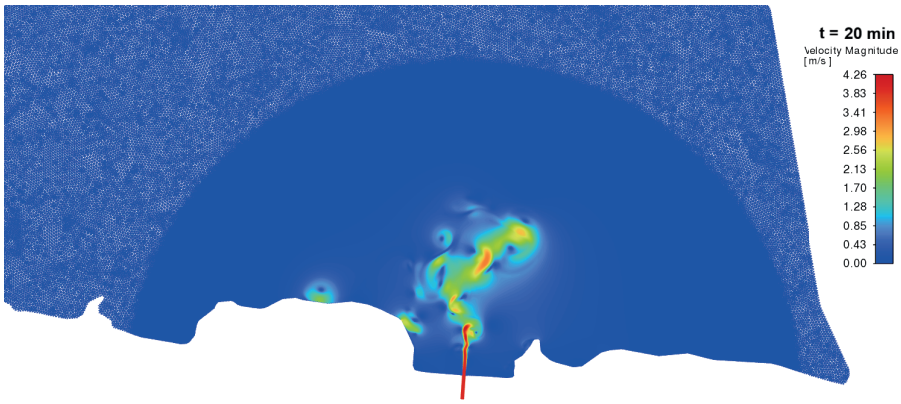


(f)

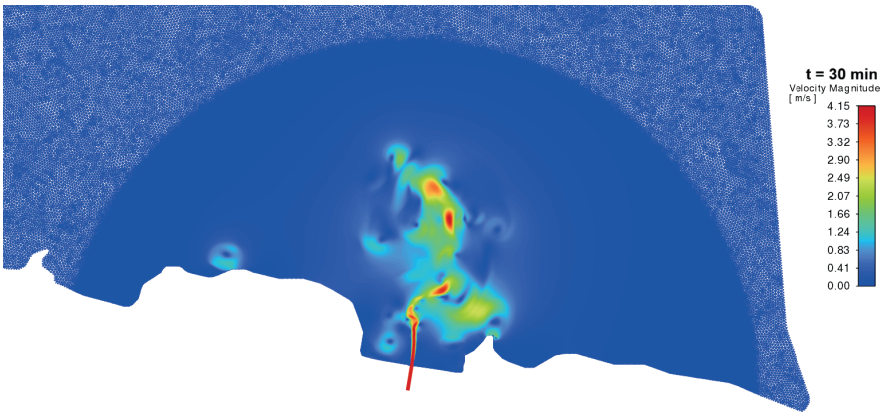
Figure 6. Evolution of the temperature field (K) in the coastal zone of Lake Balkhash during the simulation of the discharge of heated water: (a) t = 10 min, (b) t = 20 min, (c) t = 30 min, (d) t = 40 min, (e) t = 50 min (f), t = 60 min.


t = 10 min
Velocity Magnitude
[m/s]
4.51
4.06
3.61
3.16
2.71
2.25
1.80
1.35
0.90
0.45
0.00

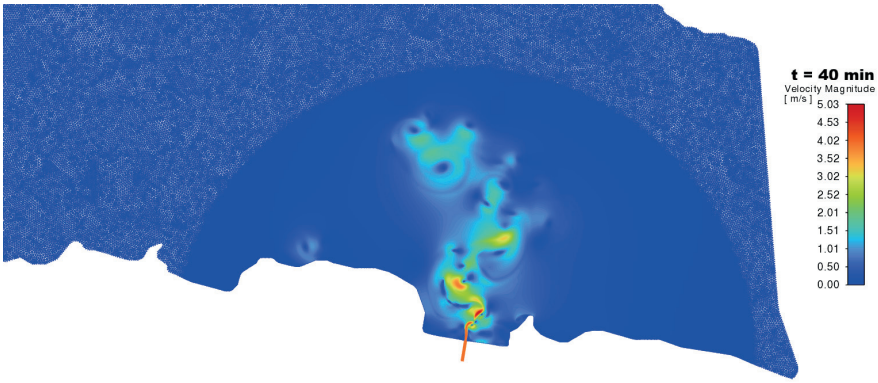
(a)


t = 20 min
Velocity Magnitude
[m/s]
4.26
3.83
3.41
2.98
2.56
2.13
1.70
1.28
0.85
0.43
0.00

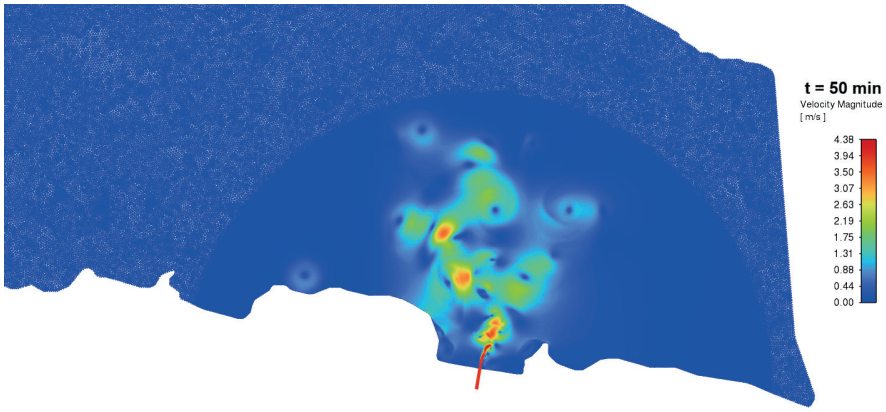
(b)


t = 30 min
Velocity Magnitude
[m/s]
4.15
3.73
3.32
2.90
2.49
2.07
1.66
1.24
0.83
0.41
0.00

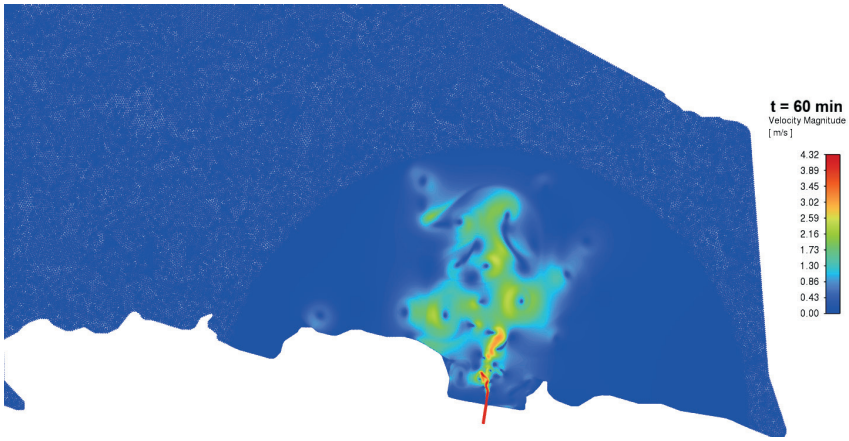
(c)



(d)



(e)



(f)

Figure 7. Velocity field (m/s) in the coastal zone of Lake Balkhash at different time stages of the simulation.

Conclusion. This numerical study provided the first detailed analysis of the spatiotemporal dynamics of heated water propagation during a potential discharge from the cooling system of a nuclear power plant in the coastal zone of Lake Balkhash. The use of a two-dimensional model based on the Navier–Stokes equations, supplemented by an energy equation and the SST k–ω turbulence model, provided a reliable description of thermohydrodynamic processes under complex coastal conditions and weak background currents.

Validation of the model on a test problem of jet injection demonstrated high agreement with experimental data, confirming the validity of the chosen approach for modeling local heat exchange processes in reservoirs. A grid-independence test demonstrated the stability of the solutions for cell sizes exceeding 2.7 million, ensuring the reliability of the obtained results. Modeling results showed that within the first 60 minutes of the discharge, a distinct thermal pollution zone of approximately 1.5 km² is formed, with water temperatures exceeding background values by 0.5–1.0 K. The high-temperature jet emanating from the source gradually transforms into a diffuse thermal spot, which persists in the coastal zone due to weak water exchange and recirculation. Analysis of the velocity field revealed the presence of localized circulation zones and stagnant areas that facilitate long-term heat retention.

These results indicate that the discharge of heated waters, even at moderate temperatures, can cause significant changes in the thermal and hydrodynamic regime of the water area. Under conditions of limited water exchange and complex shoreline morphology, such impacts can lead to long-term changes in the stratification structure, oxygen balance, and biological processes. Thus, the performed modeling confirms the effectiveness of computational fluid dynamics (CFD) methods for analyzing thermal pollution of aquatic ecosystems and can serve as a basis for subsequent assessment of environmental safety and optimization of the parameters of cooling systems of nuclear and thermal power plants located near lake areas.

References

Al-Suhaili R.H., Nasir M.J. (2006) Two-dimensional numerical model for thermal pollution of single source in river. J Eng, 12(1). — P. 25–35. DOI: 10.31026/j.eng.2006.01.03 (in English).

Caissie D. (2006) The thermal regime of rivers: a review. Freshwater Biol, 51. — P. 1389–1406. DOI: 10.1111/j.1365-2427.2006.01597x (in English).

Chang Y.R., Chen K.S. (1995) Prediction of opposing turbulent line jets discharged laterally into a confined crossflow. *Int J Heat Mass Transfer, 38(17). — P. 3069–3082. DOI: 10.1016/0017-9310(94)00390 (in English).

Chen K.S., Hwang J.Y. (1991) Experimental study on the mixing of one- and dual-line heated jets with a cold crossflow in a confined channel. AIAA J, 29(3). — P. 353–360. DOI: 10.2514/3.10619 (in English).

Du T., Dugdale S.J., Johnson M. F. (2026) A scenario-based, generalised model of the thermal impacts of cooling water discharge from thermoelectric power plants: thermopeaking and spatio-temporal interactions. Water Res, 288(Part A), 124529. DOI: 10.1016/j.watres.2025.124529 (in English).

Gaudard A., Weber C., Alexander T.J., Hunziker S., Schmid M. (2018) Impacts of using lakes and rivers for extraction and disposal of heat. WIREs Water, 5, e1295. DOI: 10.1002/wat2.1295 (in English).

Issakhov A., Zhandaulet Y. (2019) Numerical simulation of thermal pollution zones' formations in the water environment from the activities of the power plant. Eng Appl Comput Fluid Mech, 13(1). — P. 279–299. DOI: 10.1080/19942060.2019.1584126 (in English).

Kalinowska M.B., Rowiński P.M., Kubrak J., et al. (2012) Scenarios of the spread of a waste heat discharge in a river – Vistula River case study. Acta Geophys, 60. — P. 214–231. DOI: 10.2478/s11600-011-0045-x (in English).

Mahumood N.S., Mohammad T.A. (2022) Assessment of thermal pollution at selected stretch of Tigris River in Baghdad by field observations and numerical simulations. J Eng, 28(9). — P. 18–34. DOI: 10.31026/j.eng.2022.09.02 (in English).

Miara A., Vörösmarty C.J., Macknick J.E., Tidwell V.C., Fekete B., Corsi F., Newmark R. (2018) Thermal pollution impacts on rivers and power supply in the Mississippi River watershed. Environ Res Lett, 13(3), 034033. DOI: 10.1088/1748-9326/aaac85 (in English).

Råman Vinnå L., Wüest A., Bouffard D. (2017) Physical effects of thermal pollution in lakes. Water Resour Res, 53. — P. 3968–3987. DOI: 10.1002/2016WR019686 (in English).

Raptis C.E., van Vliet M. T. H., Pfister S. (2016) Global thermal pollution of rivers from thermoelectric power plants. Environ Res Lett, 11(10), 104011. DOI: 10.1088/1748-9326/11/10/104011 (in English).

Tang H.S., Paik J., Sotiropoulos F., Khangaonkar T. (2008) Three-dimensional numerical modeling of initial mixing of thermal discharges at real-life configurations. J Hydraul Eng, 134(9). — P. 1210–1224. DOI: 10.1061/(ASCE)0733-9429(2008)134:9 (1210) (in English).

Tarena F., Nyqvist D., Katopodis C., Comoglio C. (2024) Computational fluid dynamics in fishway research: a systematic review on upstream fish passage solutions. J Ecohydraulics, 10(1). — P. 107–126. DOI: 10.1080/24705357.2024.2363772 (in English).

Velazquez-Araque L., Flor J., Méndez A., Cárdenas-Calle M. (2025) Modeling microplastic dispersion in the Salado Estuary using computational fluid dynamics. Fluids, 10(5), 118. DOI: 10.3390/fluids10050118 (in English).

ACADEMIC SCIENTIFIC JOURNAL OF COMPUTER SCIENCE
ISSN 1991-346X
Volume 4. Namber 356 (2025). 171–183

https://doi.org/10.32014/2025.2518-1726.390

IRSTI 28.17.23
UDC 004.89

© Z. Kaderkeyeva[1], B. Razakhova[1,2], G. Bekmanova[1,2], A. Nazyrova[1,2*], M. Zhasuzakova[3], 2025.
[1]L.N. Gumilyov Eurasian National University, Astana, Kazakhstan;
[2]Public Association "Kazakhstan Academy of Artificial Intelligence", Astana, Kazakhstan;
[3]Almaty Technological University, Almaty, Kazakhstan.
E-mail: nazyrova_aye_1@enu.kz

Q-BILIM: AN INTELLIGENT SYSTEM FOR ASSESSING LEARNING OUTCOMES BASED ON COMPETENCIES

Kaderkeyeva Zulfiya — Master of Technical Sciences, senior lecturer of the Department of Artificial Intelligence, L.N. Gumilyov Eurasian National University, Astana, Kazakhstan,
E-mail: kaderkeyeva_zk@enu.kz, ORCID ID: https://orcid.org/0000-0001-8237-8549;
Razakhova Bibigul — Candidate of Technical Sciences, Head of the Department of Artificial Intelligence Technologies, L.N. Gumilyov Eurasian National University, Astana, Kazakhstan; Public Association "Kazakhstan Academy of Artificial Intelligence", Astana, Kazakhstan,
E-mail: razakhova_bsh@enu.kz, ORCID ID: https://orcid.org/0000-0002-8152-8661;
Bekmanova Gulmira — Candidate of technical sciences, PhD, associate professor of the Department of Artificial intelligence, L.N. Gumilyov Eurasian National University, Astana, Kazakhstan; Public Association "Kazakhstan Academy of Artificial Intelligence", Astana, Kazakhstan,
E-mail: gulmira-r@yandex.kz, ORCID ID:https://orcid.org/0000-0001-8554-7627;
Nazyrova Aizhan — PhD, senior lecturer of the Department of Artificial intelligence, L.N. Gumilyov Eurasian National University, Astana, Kazakhstan; Public Association "Kazakhstan Academy of Artificial Intelligence", Astana, Kazakhstan,
E-mail: nazyrova_aye_1@enu.kz, ORCID ID: https://orcid.org/0000-0002-9162-6791;
Zhasuzakova Meiramkul — PhD, Assistant Professor of the Department of Computer Engineering, Almaty Technological University, Almaty, Kazakhstan,
E-mail: zhasuzakova.m@mail.ru, ORCID ID: https://orcid.org/0000-0001-7510-5948.

Abstract. Competency-based education (CBE) requires transparent and objective assessment mechanisms to accurately evaluate learning outcomes. However, many institutions still face challenges related to inconsistent assessment tools, subjective grading, and weak alignment between competencies, outcomes, and evaluation criteria. This paper presents Q-Bilim, an intelligent assessment system that structures and automates evaluation processes through a weighted model linking competencies, expected outcomes, and measurement instruments. An empirical study of Q-Bilim was conducted in the “Database Fundamentals”

course from 2021 to 2024 with 170 undergraduate students of varying academic preparation. A unified four-level grading scale ensured consistency, while statistical methods such as ANOVA, Tukey's HSD test, regression analysis, and cluster analysis were used to assess cohort differences, identify learning patterns, and examine how assessment weightings affect overall performance. This approach provided insights into factors shaping competency development. The study's main contribution is the development and validation of a scalable, competency-based assessment framework implemented within Q-Bilim. Findings show a clear relationship between assessment weightings and student achievement, while cluster analysis revealed distinct learner categories from high-performing students to those needing additional support. The practical significance of Q-Bilim lies in its ability to support long-term monitoring, expose weaknesses in curriculum design, and help instructors make timely, data-driven decisions. Future work includes expanding the system to other disciplines and integrating advanced AI methods – such as graph neural networks – to enhance analytical precision, personalization, and usability.

Keywords: competency-based education, learning outcomes, student performance, educational analytics, Q-Bilim, assessment system

This research is funded by the Science Committee of the Ministry of Science and Higher Education of the Republic of Kazakhstan (Grant No. AP19678613).

© З.К. Кадеркеева[1], Б.Ш. Разахова[1,2], Г.Т. Бекманова[1,2], А.Е. Назырова[1,2]*, М.Ж. Жасұзақова[3], 2025.

[1]Л.Н. Гумилев атындағы Еуразия ұлттық университеті, Астана, Қазақстан;

[2]«Қазақстан жасанды интеллект академиясы» Қоғамдық Бірлестігі, Астана, Қазақстан;

[3] Алматы технологиялық университеті, Алматы, Қазақстан

E-mail: nazyrova_aye_1@enu.kz

Q-BILIM: ҚҰЗЫРЕТТЕРГЕ НЕГІЗДЕЛГЕН ОҚУ НӘТИЖЕЛЕРІН БАҒАЛАУҒА АРНАЛҒАН ИНТЕЛЛЕКТУАЛДЫ ЖҮЙЕ

Кадеркеева Зульфия — техника ғылымдарының магистрі, Жасанды интеллект кафедрасының аға оқытушысы, Л.Н. Гумилев атындағы Еуразия ұлттық университеті, Астана, Қазақстан,

E-mail: kaderkeyeva_zk@enu.kz, ORCID ID: https://orcid.org/0000-0001-8237-8549;

Разахова Бибигул — техника ғылымдарының кандидаты, Жасанды интеллект технологиялары кафедрасының меңгерушісі, Л.Н. Гумилев атындағы Еуразия ұлттық университеті, Астана, Қазақстан; «Қазақстан жасанды интеллект академиясы» Қоғамдық бірлестігі, Астана, Қазақстан,

E-mail: razakhova_bsh@enu.kz, ORCID ID: https://orcid.org/0000-0002-8152-8661; **Бекманова Гүлмира —** техника ғылымдарының кандидаты, PhD, Жасанды интеллект кафедрасының қауымдастырылған профессор, Л.Н. Гумилев атындағы Еуразия ұлттық университеті, Астана, Қазақстан; «Қазақстан жасанды интеллект академиясы» Қоғамдық

бірлестігі, Астана, Қазақстан,
E-mail: gulmira-r@yandex.kz, ORCID: https://orcid.org/0000-0001-8554-7627;
Назырова Айжан — PhD, Жасанды интеллект кафедрасының аға оқытушысы, Л.Н. Гумилев атындағы Еуразия ұлттық университеті, Астана, Қазақстан, «Қазақстан жасанды интеллект академиясы» Қоғамдық бірлестігі, Астана, Қазақстан,
E-mail: nazyrova_aye_1@enu.kz, ORCID ID: https://orcid.org/0000-0002-9162-6791;
Жасұзақова Мейрамкүл — PhD, Компьютерлік инженерия кафедрасының ассистент-профессоры, Алматы технологиялық университеті, Алматы, Қазақстан,
E-mail: zhasuzakova.m@mail.ru, ORCID ID: https://orcid.org/0000-0001-7510-5948.

Аннотация. Құзыреттілікке негізделген білім беру (CBE) оқу нәтижелерін дәл бағалау үшін ашық және объективті бағалау механизмдерін талап етеді. Алайда көптеген оқу орындары бағалау құралдарының жүйесіздігіне, балл қоюдағы субъективтілікке және құзыреттер, күтілетін нәтижелер мен бағалау критерийлері арасындағы әлсіз үйлесімге байланысты қиындықтарға әлі де тап болып отыр. Бұл мақалада құзыреттерді, күтілетін нәтижелерді және өлшеу құралдарын өзара байланыстыратын өлшемдік модель арқылы бағалау үдерісін құрылымдап, автоматтандыратын Q-Bilim интеллектуалды бағалау жүйесі таныстырылады. Q-Bilim жүйесіне жүргізілген эмпирикалық зерттеу 2021–2024 жылдары «Деректер базасының негіздері» курсында әртүрлі дайындық деңгейіндегі 170 бакалавр студентінің қатысуымен өткізілді. Бірыңғай төрт деңгейлі бағалау шкаласы біріздiлікті қамтамасыз етті. ANOVA, Tukey HSD, регрессиялық және кластерлік талдау сияқты статистикалық әдістер топтар арасындағы айырмашылықтарды анықтау, оқу үлгілерін сипаттау және бағалау салмақтарының жалпы жетістікке әсерін зерттеу үшін қолданылды. Бұл тәсіл құзыреттіліктің қалыптасуына ықпал ететін факторлар жөнінде маңызды түсініктер берді. Зерттеудің негізгі нәтижесі – Q-Bilim жүйесінде іске асырылған ауқымды, құзыреттілікке бағытталған бағалау моделін әзірлеу және оны дәлелдеу. Нәтижелер бағалау құралдарының салмағы мен студент жетістіктері арасында айқын байланыс бар екенін көрсетті, ал кластерлік талдау жоғары жетістікке жететіндерден бастап қосымша қолдауды қажет ететіндерге дейін әртүрлі студент топтарын анықтады. Q-Bilim жүйесінің практикалық маңызы – ұзақ мерзімді мониторинг жүргізу, оқу бағдарламаларындағы әлсіз тұстарды анықтау және оқытушыларға деректерге негізделген жедел шешімдер қабылдауға көмектесу. Болашақта жүйені басқа пәндерге кеңейту және аналитикалық дәлдікті, жекелендіруді және қолайлылықты арттыру үшін графтық нейрондық желілер сияқты озық жасанды интеллект әдістерін интеграциялау жоспарлануда.

Түйін сөздер: құзіреттілікке негізделген білім беру, оқу нәтижелері, студенттердің үлгерімі, білім беру аналитикасы, Q-Bilim, бағалау жүйесі

© З.К. Кадеркеева[1], Б.Ш. Разахова[1,2], Г.Т. Бекманова[1,2], А.Е. Назырова[1,2*], М.Ж. Жасузакова[3], 2025.

[1]Евразийский национальный университет им. Л.Н. Гумилёва, Астана, Казахстан;

[2]Общественное объединение «Казахстанская академия искусственного интеллекта», Астана, Казахстан;

[3] Алматинский технологический университет, Алматы, Казахстан.

E-mail: nazyrova_aye_1@enu.kz

Q-BILIM: ИНТЕЛЛЕКТУАЛЬНАЯ СИСТЕМА ОЦЕНКИ РЕЗУЛЬТАТОВ ОБУЧЕНИЯ НА ОСНОВЕ КОМПЕТЕНЦИЙ

Кадеркеева Зульфия — магистр технических наук, старший преподаватель кафедры Искусственного интеллекта, Л.Н. Евразийский национальный университет им. Л.Н. Гумилева, Астана, Казахстан,
E-mail: kaderkeyeva_zk@enu.kz, ORCID ID: https://orcid.org/0000-0001-8237-8549;
Разахова Бибигул — к.т.н., заведующий кафедрой Технологий Искусственного интеллекта, Евразийский национальный университет им. Л.Н. Гумилева, Астана, Казахстан; Общественное объединение «Казахстанская академия искусственного интеллекта», Астана, Казахстан,
E-mail: razakhova_bsh@enu.kz, https://orcid.org/0000-0002-8152-8661;
Бекманова Гүлмира — PhD, доцент кафедры Искусственного интеллекта, Евразийский национальный университет им. Л.Н. Гумилева, Астана, Казахстан; Общественное объединение «Казахстанская академия искусственного интеллекта», Астана, Казахстан,
E-mail: gulmira-r@yandex.kz, ORCID ID: https://orcid.org/0000-0001-8554-7627;
Назырова Айжан — PhD, старший преподаватель кафедры Искусственного интеллекта, Евразийский национальный университет им. Л.Н. Гумилева, Астана, Казахстан; Общественное объединение «Казахстанская академия искусственного интеллекта», Астана, Казахстан,
E-mail: nazyrova_aye_1@enu.kz, ORCID ID: https://orcid.org/0000-0002-9162-6791;
Жасузакова Мейрамкул — PhD, ассистент-профессора кафедры Компьютерной инженерии, Алматинский технологический университет, Алматы, Казахстан,
E-mail: zhasuzakova.m@mail.ru, ORCID ID: https://orcid.org/0000-0001-7510-5948.

Аннотация. Компетентностное обучение (CBE) требует прозрачных и объективных механизмов оценивания для точного измерения результатов обучения. Однако многие учебные заведения по-прежнему сталкиваются с проблемами, связанными с несогласованными инструментами оценивания, субъективностью выставления баллов и слабой связью между компетенциями, ожидаемыми результатами и критериями оценивания. В данной работе представлен Q-Bilim – интеллектуальная система оценивания, которая структурирует и автоматизирует процесс оценки с помощью взвешенной модели, связывающей компетенции, ожидаемые результаты и инструменты измерения. Эмпирическое исследование Q-Bilim проводилось на курсе «Основы баз данных» в 2021–2024 годах с участием 170 студентов бакалавриата с разным уровнем подготовки. Единая четырехуровневая шкала обеспечила согласованность, а такие статистические методы, как ANOVA, критерий Тьюки (HSD), регрессионный и кластерный анализ,

были использованы для оценки различий между группами, выявления учебных паттернов и изучения влияния весовых коэффициентов на итоговые результаты. Такой подход позволил получить представление о факторах, влияющих на формирование компетенций. Основной вклад исследования заключается в разработке и валидации масштабируемой компетентностно-ориентированной модели оценивания, реализованной в системе Q-Bilim. Результаты показали чёткую связь между весами оценочных инструментов и академическими достижениями студентов, а кластерный анализ позволил выделить различные группы обучающихся – от высокоуспевающих до нуждающихся в поддержке. Практическая значимость Q-Bilim проявляется в возможностях долгосрочного мониторинга, выявления слабых мест в образовательных программах и помощи преподавателям в своевременном принятии решений на основе данных. В дальнейшем планируется расширение системы на другие дисциплины и интеграция методов ИИ – таких как графовые нейронные сети – для повышения аналитической точности, персонализации и удобства использования.

Ключевые слова: компетентностно-ориентированное образование, результаты обучения, академическая успеваемость, образовательная аналитика, QBilim, система оценивания

Introduction. One of the major developments in contemporary educational systems is competency-based education, or CBE. This method emphasizes students' academic knowledge as well as their practical abilities, professional competencies, and capacity to apply knowledge in real-world scenarios. The CBE model is widely used in the United States, Kazakhstan, a number of Latin American nations, and Western European nations (as part of the Bologna Process).

Throughout the educational process, learning outcomes must be methodically planned, tracked, and evaluated in order for CBE to be implemented successfully. However, it frequently becomes challenging to effectively measure the quality of courses because of the irregularities in these procedures and the dearth of efficient assessment instruments in many educational institutions. Furthermore, two major factors that lower the overall efficacy of the teaching process are the instructors' use of a variety of evaluation techniques and the lack of well-defined connections between learning outcomes and competences.

We suggest Q-Bilim, a novel information system, as a solution to these issues. The purpose of this system is to automatically evaluate student performance and track the caliber of competency-based education. The following are the main goals that the Q-Bilim system seeks to accomplish: Osygan Baylanysty, without Q-Bilim dep atalatyn zhana akparattyk zhuyeni usynamyz. Bul zhuye kuzyrettilikke negizdelgen kurstard studentzhetistigin avtomatty turde bagalauga zhane kurstyn sapasyn monitoringteuge arnalgan.

The following primary goals are intended to be addressed by the Q-Bilim system:

• formalization of the competency model for the course, which involves

describing how competencies, learning goals, and assessment instruments relate to one another;

• evaluation of each component's corresponding student achievement levels;

• examination of the dynamics of course development throughout a number of academic years;

• assessment of the effectiveness of educational decisions, such as modifications to the curriculum or teaching methods;

• graphical reports that give teachers and students clear, understandable feedback.

The Q-Bilim system's architecture, working principles, and results from years of use in a particular course (such as Introduction to Programming) are presented in this article. Statistical techniques like Tukey's post-hoc tests and analysis of variance (ANOVA) have validated the strategy's efficacy.

Literature review. The development of the concepts of learning outcomes and competencies in higher education has been the focus of growing scholarly attention since the late 20th century. Mulder, Weigel, and Collins highlighted the role of competence in vocational education across EU member states, emphasizing its importance for curriculum design and policy implementation Mulder M. et al. (2007). In a similar direction, Harden examined the distinction between instructional objectives and learning outcomes, thereby grounding the paradigm of student-centered learning Harden R.M. (2002).

From the perspective of personalization and adaptability, Brusilovsky and Millán provided a seminal contribution by introducing user models for adaptive hypermedia, offering theoretical and practical foundations for individualized learning pathways Brusilovsky P. et al. (2007). These ideas paved the way for modern applications of artificial intelligence in educational systems.

Digital learning environments have increasingly been analyzed through the lens of AI-based learning analytics. Aldahwan and Alsaeed highlighted the potential of artificial intelligence to enhance the functionality of Learning Management Systems, demonstrating how AI-driven techniques can improve learner monitoring, support, and overall analysis within LMS platforms (Aldahwan & Alsaeed, 2020). Complementing this, Zawacki-Richter and colleagues conducted a systematic review of AI applications in higher education, identifying not only the opportunities but also ethical and methodological limitations of AI integration Zawacki-Richter O. et al. (2019).

A related line of inquiry focuses on the ontological modeling of learning processes. Capuano, Gaeta, Salerno, and Mangione proposed an ontology-based approach for context-aware e-learning, demonstrating how semantic structures can enhance the organization and adaptation of educational content (Capuano et al., 2011). This was expanded by Nazyrova et al., who investigated the consistency between prerequisites and learning outcomes across educational programs using an ontological approach Nazyrova A. et al. (2023). Together, these works suggest the value of ontology for the automated analysis and improvement of curricula.

Recent years have also seen a surge of interest in assessing and enhancing learning outcomes more broadly. Goss reviewed the literature on outcomes assessment in higher education and academic libraries, pointing to the need for systematic practices Harrold Goss. (2022). Doo and Kim conducted a meta-analysis demonstrating a significant relationship between learning engagement and learning outcomes in online higher education environments (Doo & Kim, 2024). Similarly, Akpen and colleagues provided a systematic review of online learning's impact on students' performance and engagement Akpen C.N. et al. (2024).

Within innovative instructional contexts, Alizadeh explored learner engagement and perceived learning outcomes in immersive flipped learning Alizadeh M. (2024). Alyasin and colleagues further stressed the importance of moving from fragmented practices toward systematized learning outcomes assessment Alyasin A. et al. (2023). Finally, López-Nuñez et al. carried out a systematic review of digital competence evaluation in higher education, underscoring its significance for quality assurance and sustainable educational development López-Nuñez J. et al. (2024).

Research methods. The methodological basis of the study is based on the principles of competency-based education (CBE) and constructive identification. A multi-level methodological approach was used to assess the effectiveness of the Q-Bilim system. It includes formal modeling, empirical data collection, quantitative and qualitative analysis, as well as statistical methods.

Formal modeling. The hierarchical relationships between course competencies, learning outcomes, and assessment tools were described using weighting coefficients. This structure clearly demonstrates the relationship of each learning outcome to a specific competency and allows for objective monitoring over several academic years.

Data and sampling. The empirical study was conducted in the course "Fundamentals of Databases" from 2021 to 2024. A total of 170 students were included. Three cohorts were formed and distributed approximately equally according to the three academic years. The levels of preparation of the students were different: some had basic programming skills, while others started the subject from scratch. This situation ensured the representativeness of the study, however, stratified analysis is considered as a direction for future research.

Unified rating scale. To avoid differences in interpretation, all numerical ratings (0–100) were converted into qualitative levels on a single four-point scale (Table 1). This scale was used throughout the analysis.

Table 1. Unified four-point rating scale

Score range	Quality level
90–100	Excellent
70–89	Good
50–69	Satisfactory
0–49	Needs improvement

Quantitative and qualitative analysis. Student results were presented in both quantitative and qualitative terms. Color-coded tables and charts were used to illustrate the level of mastery of competencies and learning outcomes. This approach made data easier for teachers and students to understand.

Results. Traditional analysis of variance (ANOVA) and Tukey HSD test were used to identify differences in each academic year ($p < 0.05$). In addition, regression analysis was conducted to determine the relationship between the weighting coefficients of assessment tools and the overall performance of students. In addition, cluster analysis was conducted to group students' learning trajectories, resulting in three main categories: "high achievers", "average" and "risk group".

This comprehensive methodological approach allowed, in addition to assessing the effectiveness of the system, to delve deeper into the factors influencing students' academic achievements.

General review of the Q-Bilim system. The Q-Bilim system aims to improve the assessment and monitoring process in competency-based courses. The system is built on a modular architecture and allows teachers to monitor the quality of the course over several academic years, analyze student achievement, and formally model the relationship between learning outcomes and competencies (figure 1).

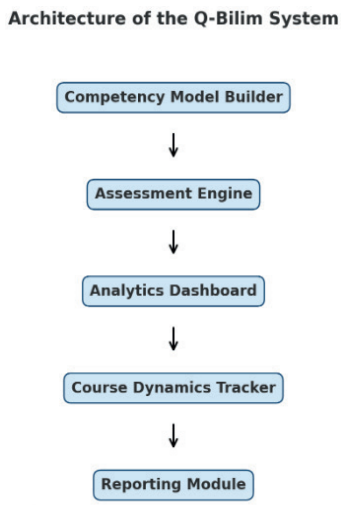


Figure 1. Architecture of the Q-Bilim System

The main feature of this structure is that it presents assessment results in both quantitative and qualitative formats. Student performance is not only presented in terms of average scores, but also in terms of qualitative levels such as “high”, “average” or “insufficient”. This approach makes the results understandable and easy to interpret for all stakeholders.

Regression analysis. The data collected through the system allowed us to analyze the impact of assessment weightings on student performance. As shown in Figure

2, it was observed that as the weighting ratio increased, the average performance of students also increased. This result proves that instead of distributing the importance of assessment tools equally, it is necessary to consider their contribution to the development of specific course competencies.

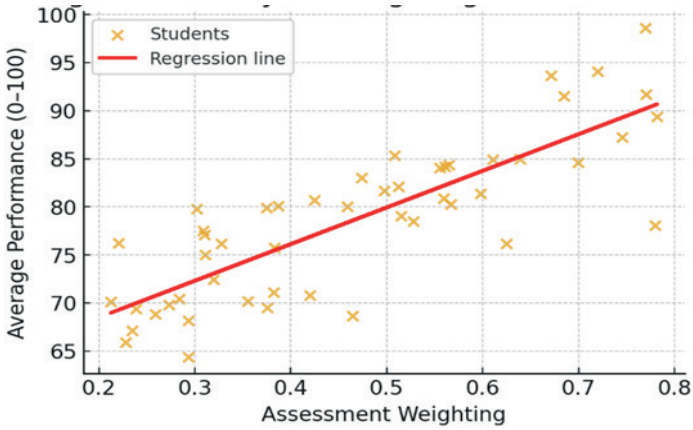


Figure 2. Regression analysis showing the relationship between assessment weightings and student performance

Cluster analysis. A cluster analysis was used to analyze individual student learning trajectories. Figure 3 identifies three main groups of students based on the results obtained for competencies C1 and C2:

- high achievers (green);
- average achievers (blue);
- at-risk (red).

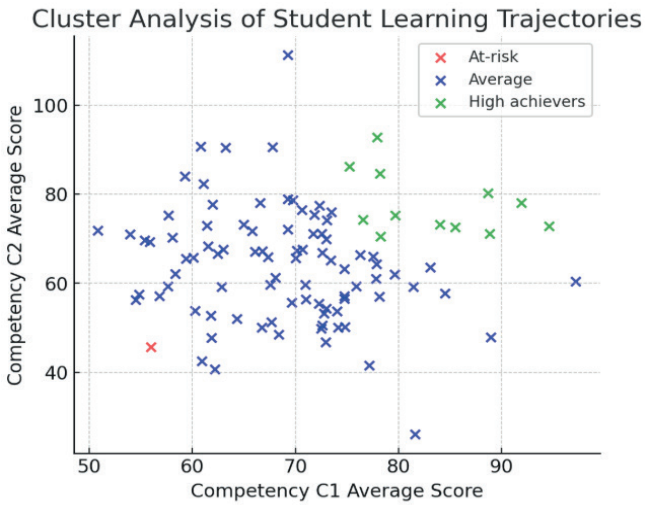


Figure 3. Cluster analysis of student learning trajectories by competencies (C1 vs. C2)

This classification allows teachers to identify students who need additional support in a timely manner. In addition, analyzing differences within groups allows for the adaptation of teaching strategies.

As shown in Figure 4, average scores for all three competencies increased steadily from 2021 to 2024. This trend suggests that curriculum adjustments and assessment changes have had a positive impact on student learning outcomes.

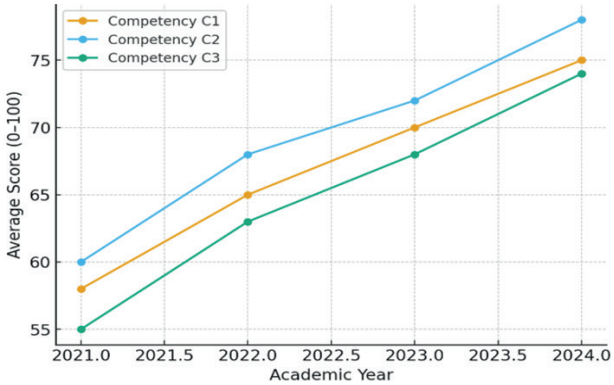


Figure 4. Dynamics of competency mastery across academic years (2021–2024)

This graph shows the change in students' level of mastery of key competencies over time. From 2021 to 2024, there is a gradual increase in the average indicators for all competencies. Such dynamics prove the positive impact of changes introduced into the curriculum and improvements in the assessment system.

As shown in Figure 5, the distribution of student achievement across grade levels demonstrates a consistent upward trend. Box plots show a clear increase in median values and a decrease in variability, indicating greater stability in learning outcomes. Analysis of variance (ANOVA) results confirm the statistical significance of these differences ($p < 0.05$).

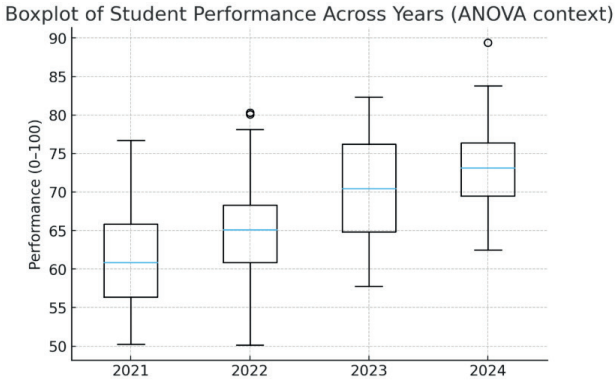


Figure 5. Boxplot of student performance distribution across years, confirming statistically significant differences (ANOVA)

This chart shows the distribution of student performance across each academic year. The boxplot format clearly shows the median, interquartile range, and outliers in student performance. The ANOVA results revealed that there were statistically significant differences ($p < 0.05$) in performance levels between 2021 and 2024. These changes confirm that academic performance has been improving over the years.

Discussion. The findings of this study demonstrate that the Q-Bilim system provides an effective framework for monitoring and assessing competency-based education (CBE). The integration of formalized competency modeling, automated data processing, and analytical visualization significantly enhances the objectivity and consistency of evaluation compared to traditional methods.

The system's ability to visualize longitudinal data enabled instructors to identify weaknesses in the curriculum and apply targeted improvements, aligning with the insights of Ifenthaler, Mah, and Yau, who emphasized the role of learning analytics in supporting study success and informing data-driven educational decisions (Ifenthaler et al., 2019).

The regression analysis further revealed that assessment weighting has a direct correlation with learning outcomes. This aligns with the insights of Romero and Ventura, who emphasized that educational data mining techniques enable the identification of key performance patterns and factors influencing student achievement (Romero & Ventura, 2020). Such findings support the idea that differentiated weighting of assessment tools—depending on their contribution to particular competencies—can lead to more balanced and fair evaluation models. Similarly, the cluster analysis offered valuable insights into individual learning trajectories, facilitating the detection of at-risk students and enabling timely, data-driven pedagogical interventions.

These outcomes align with findings from previous studies emphasizing the importance of data-driven and adaptive approaches in higher education (Zawacki-Richter et al., 2019; Capuano et al., 2011). The use of artificial intelligence and ontological modeling in Q-Bilim reflects a broader trend toward digital transformation in educational analytics and supports the concept of personalized learning pathways, as discussed by Brusilovsky and Millán (2007).

However, several limitations must be noted. The current study was limited to a single discipline ("Fundamentals of Databases") and relied on a relatively homogeneous student population. The model could benefit from a larger and more diverse dataset to test its generalizability across different academic programs. Additionally, while Q-Bilim incorporates basic statistical analyses, further integration of advanced AI methods – such as predictive modeling and graph-based learning–could enhance its diagnostic capabilities.

Future work will focus on expanding the system's analytical modules, improving the user interface for accessibility, and integrating adaptive feedback mechanisms. The inclusion of machine learning techniques like graph neural networks (GNNs) may enable real-time recommendations for personalized learning support.

In summary, Q-Bilim demonstrates that intelligent, data-driven systems can transform the assessment of learning outcomes in competency-based education. Its implementation supports evidence-based decision-making, contributes to quality assurance in higher education, and fosters a culture of continuous improvement among both instructors and learners.

Conclusion. The study presented the architecture, core components, and implementation results of the Q-Bilim intelligent system, designed for assessing and supervising competency-based education. The system enables the structured analysis of multi-year academic data, allowing for both objective evaluation of student achievements and the monitoring of instructional quality over time. By linking competencies, learning outcomes, and assessment tools through a unified model, Q-Bilim ensures consistency, transparency, and data reliability in the evaluation process.

Key advantages of the system include its ability to align assessment criteria with defined competencies, determine achievement levels with weighted evaluation tools, and automatically translate quantitative indicators into qualitative performance categories. Furthermore, Q-Bilim provides educators and administrators with statistical and visual analytics that facilitate data-informed decision-making.

The application of Q-Bilim in the "Fundamentals of Databases" course demonstrated measurable improvement in student performance, confirmed through statistical methods such as ANOVA and Tukey's HSD tests. This validates the system's effectiveness in enhancing both assessment accuracy and learning outcomes.

Future research and development will focus on scaling the system to multiple disciplines and integrating advanced artificial intelligence techniques, including graph neural networks (GNNs), to enable personalized learning recommendations. Ultimately, Q-Bilim stands as a scientifically grounded, data-driven tool that contributes to the modernization and quality assurance of competency-based education.

References

Akpen C.N., Asaolu S., Atobatele S., Okagbue H., Sampson S. (2024) Impact of online learning on student's performance and engagement: a systematic review. Discover Education, 3(1), 205. DOI: https://doi.org/10.1007/s44217-024-00253-0 (in English)

Aldahwan N., Alsaeed N. (2020) Use of artificial intelligent in Learning Management System (LMS): a systematic literature review. International Journal of Computer Applications, 175(13). — P. 16-26. DOI: https://doi.org/10.5120/ijca2020920611 (in English)

Alizadeh M. (2024) Exploring engagement and perceived learning outcomes in an immersive flipped learning context. arXiv preprint arXiv:2409.12674. DOI: https://doi.org/10.32664/ijitgeb.v6i2.155 (in English)

Alyasin A., Nasser R., El Hajj M., Harb H. (2023). Assessing Learning Outcomes in Higher Education: From Practice to Systematization. TEM Journal, 12(3). DOI: https://doi.org/10.18421/TEM123-41 (in English)

Brusilovsky P., Millán E. (2007) User models for adaptive hypermedia and adaptive educational systems. In The adaptive web: methods and strategies of web personalization. — P. 3-53. Berlin,

Heidelberg: Springer Berlin Heidelberg. DOI: https://doi.org/10.1007/978-3-540-72079-9_1 (in English)

Capuano N., Gaeta M., Salerno S., Mangione G.R. (2011, November) An ontology-based approach for context-aware e-learning. In 2011 Third International Conference on Intelligent Networking and Collaborative Systems. — P. 789-794. IEEE, DOI: https://doi.org/10.1109/INCoS.2011.53 (in English)

Doo M.Y., Kim J. (2024) The relationship between learning engagement and learning outcomes in online learning in higher education: A meta-analysis study. Distance Education, 45(1). — P. 60-82. DOI: https://doi.org/10.1080/01587919.2024.2303484 (in English)

Harden R.M. (2002) Learning outcomes and instructional objectives: is there a difference?. Medical teacher, 24(2). — P. 151-155. DOI: https://doi.org/10.1080/0142159022020687 (in English)

Goss H. (2022). Student learning outcomes assessment in higher education and in academic libraries: A review of the literature. The Journal of Academic Librarianship, 48(2). — 102485 p. DOI: https://doi.org/10.1016/j.acalib.2021.102485 (in English)

Ifenthaler D., Mah D.K., Yau J.Y.K. (2019) Utilising learning analytics for study success: Reflections on current empirical findings. In Utilizing learning analytics to support study success. — P. 27-36). Cham: Springer International Publishing. DOI: https://doi.org/10.1007/978-3-319-64792-0_2 (in English)

López-Nuñez J.A., Alonso-García S., Berral-Ortiz, B., Victoria-Maldonado, J.J. (2024) A systematic review of digital competence evaluation in higher education. Education Sciences, 14(11), 1181. DOI: https://doi.org/10.3390/educsci14111181 (in English)

Mulder M., Weigel T., Collins K. (2007) The concept of competence in the development of vocational education and training in selected EU member states: a critical analysis. Journal of Vocational Education and Training, 59(1). — P. 67-88. DOI: https://doi.org/10.1080/13636820601145630 (in English)

Nazyrova A., Milosz M., Bekmanova G., Omarbekova A., Mukanova A., Aimicheva G. (2023) Analysis of the consistency of prerequisites and learning outcomes of educational programme courses by using the ontological approach. Applied Sciences, 13(4), 2661. DOI: https://doi.org/10.3390/app13042661 (in English)

Romero C., Ventura S. (2020) Educational data mining and learning analytics: An updated survey. Wiley interdisciplinary reviews: Data mining and knowledge discovery, 10(3), e1355. DOI: https://doi.org/10.1002/widm.1355 (in English)

Zawacki-Richter O., Marín, V.I., Bond M., Gouverneur F. (2019) Systematic review of research on artificial intelligence applications in higher education–where are the educators?. International journal of educational technology in higher education, 16(1). — P. 1-27. DOI: https://doi.org/10.1186/s41239-019-0171-0 (in English)

ACADEMIC SCIENTIFIC JOURNAL OF COMPUTER SCIENCE
ISSN 1991-346X
Volume 4. Namber 356 (2025). 184–202

https://doi.org/10.32014/2025.2518-1726.391

UDC 336.77:338.43(477)

© N. Karymsakova[1], A. Boltaboyeva [2,3,4 *], D. Turmakhanbet [2,3,4], M. Maulenbekov[2,3,5], T. Abdirova[6], 2025.

[1]ALT University named after M. Tynyshpayev, Almaty, Kazakhstan;
[2]Joldasbekov Institute of Mechanics and Engineering, Almaty, Kazakhstan;
[3]LLP "Kazakhstan R&D Solutions", Almaty, Kazakhstan;
[4]Kazakh National University named after Al-Farabi, Almaty, Kazakhstan;
[5]Kazakh National Research Technical University named after K.I. Satpayev, Almaty, Kazakhstan;
[6]State Communal Enterprise on the Right of Economic Management "Regional Cardiology Center", Almaty, Kazakhstan.
E-mail: boltaboyeva_assiya3@live.kaznu.kz

UNSUPERVISED LEARNING FOR THE IDENTIFICATION OF CRITICAL CONDITIONS IN RENEWABLE ENERGY PRODUCTION

Nurgul Karymsakova — PhD, Associate Professor, Researcher, ALT University named after M. Tynyshpaev, Almaty, Kazakhstan, E-mail: n.karymsakova@alt.edu.kz, https://orcid.org/0000-0001-8187-2369;
Assiya Boltaboyeva — Researcher, Joldasbekov Institute of Mechanics and Engineering; Research Associate at LLP "Kazakhstan R&D Solutions"; 3rd-year PhD student at the Faculty of Information Technologies and Artificial Intelligence, Al-Farabi Kazakh National University, Almaty, Kazakhstan, E-mail: boltaboyeva_assiya3@live.kaznu.kz, https://orcid.org/0000-0002-7279-9910;
Dinara Turmakhanbet — IoT Engineer at LLP "Kazakhstan R&D Solutions"; 4th-year student, Faculty of Information Technologies and Artificial Intelligence, Al-Farabi Kazakh National University; Joldasbekov Institute of Mechanics and Engineering, Almaty, Kazakhstan,
E-mail: turmahanbetdinara@gmail.com, https://orcid.org/0009-0004-8388-4979;
Marlen Maulenbekov — IoT Engineer at LLP "Kazakhstan R&D Solutions"; 4th-year student, Faculty of Electronics and Engineering, Satbayev University (Kazakh National Research Technical University named after K. Satpayev); Joldasbekov Institute of Mechanics and Engineering, Almaty, Kazakhstan,
E-mail: marlen.maulenbek@gmail.com, https://orcid.org/0009-0002-0703-7634;
Aliya Baydauletova — Neurologist and Somnologist, Candidate of Medical Sciences, Medical Center "Neuroclinika", Almaty, Kazakhstan,
E-mail: baidaulet123@gmail.com, https://orcid.org/0009-0000-5510-3590.

Abstract. Anaerobic digestion is a multifaceted biochemical process that frequently encounters instability arising from fluctuations in feedstock characteristics, shifts in operating conditions, and changes within the microbial community. While mechanistic approaches such as ADM1 provide an established

theoretical framework, they typically demand substantial parameter tuning and often fail to fully capture the rapidly changing behavior of industrial-scale reactors. To overcome these constraints, this study employs unsupervised learning—particularly clustering analyses—to distinguish discrete, recurring operational regimes in full-scale anaerobic digesters. Using long-term datasets collected from commercial facilities, the analysis incorporates multiple indicators, including concentrations of volatile fatty acids, biogas composition, and other key process variables. The clustering results consistently revealed three dominant regimes: a stable mode associated with high methane productivity; transitional conditions marked by moderate imbalances; and inhibited states characterized by elevated volatile fatty acids and lower methane content. Importantly, these data-driven regimes corresponded closely with real operational events such as abrupt increases in organic loading, underscoring the method's ability to detect emerging disturbances before they escalate. The study demonstrates that unsupervised learning can serve as a valuable complement to traditional modeling by offering a practical means of interpreting complex process dynamics. Its capacity to highlight regime shifts and signal early deviations suggests clear potential for improving monitoring strategies, reinforcing system stability, and enabling more proactive, data-informed management of industrial anaerobic digestion operations.

Keywords: anaerobic digestion, unsupervised learning, clustering, process monitoring, biogas production, operational regimes, machine learning

Funding: *This research was funded by the Science Committee of the Ministry of Science and Higher Education of the Republic of Kazakhstan (Grant No. AP26103739) and LLP Kazakhstan R&D Solutions.*

Н. Карымсакова[1], А. Болтабоева[2,3,4] *, Д. Тұрмаханбет[2,3,4], М. Мауленбеков[2,3,5], Т. Абдирова[6] 2025.

[1]М. Тынышпаев атындағы ALT университеті, Алматы, Қазақстан;
[2]У.А. Жолдасбеков атындағы Механика және Инженерия Институты, Алматы, Қазақстан;
[3]"Kazakhstan R&D Solutions" ЖШС, Алматы, Қазақстан;
[4]Әл-Фараби атындағы Қазақ ұлттық университеті, Алматы, Қазақстан;
[5]Қ. Сәтбаев атындағы Қазақ ұлттық техникалық зерттеу университеті, Алматы, Қазақстан;
[6]Шаруашылық жүргізу құқығындағы мемлекеттік коммуналдық кәсіпорын «Облыстық кардиологиялық орталық, Алматы, Қазақстан.
E-mail: boltaboyeva_assiya3@live.kaznu.kz

ЖАҢАРТЫЛАТЫН ЭНЕРГИЯ ӨНДІРІСІНДЕГІ КРИТИКАЛЫҚ РЕЖИМДЕРДІ АНЫҚТАУҒА АРНАЛҒАН МҰҒАЛІМСІЗ ОҚЫТУ

Карымсакова Нургуль — PhD, М. Тынышпаев атындағы ALT университеті, Алматы, Қазақстан,

E-mail: n.karymsakova@alt.edu.kz, ORCID ID: https://orcid.org/0000-0001-8187-2369;
Болтабоева Асия — зерттеушы, У.А. Жолдасбеков атындағы Механика және Инженерия институты, ТОО «Kazakhstan R&D Solutions» ғылыми қызметкері, Әл-Фараби атындағы Қазақ Ұлттық Университетінің Ақпараттық технологиялар және жасанды интеллект факультетінің 3-курс докторанты, Алматы, Қазақстан,
E-mail: boltaboyeva_assiya3@live.kaznu.kz; https://orcid.org/0000-0002-7279-9910;
Турмахан Динара — IoT инженері, ТОО «Kazakhstan R&D Solutions», 4-курс студенті, Әл-Фараби атындағы Қазақ Ұлттық университетінің Ақпараттық технологиялар және жасанды интеллект факультеті; У.А. Жолдасбеков атындағы Механика және Инженерия институты, Алматы, Қазақстан,
E-mail: turmahanbetdinara@gmail.com, https://orcid.org/0009-0004-8388-4979;
Мауленбеков Марлен — IoT инженері, ТОО «Kazakhstan R&D Solutions», 4-курс студенті, К. Сәтбаев атындағы Қазақ ұлттық зерттеу техникалық университетінің Электроника және инженерия факультеті; У.А. Жолдасбеков атындағы Механика және Инженерия институты, Алматы, Қазақстан,
E-mail: marlen.maulenbek@gmail.com, https://orcid.org/0009-0002-0703-7634;
Абдирова Тамара — невролог және сомнолог, медицина ғылымдарының кандидаты, «Нейроклиника» медициналық орталығы, Алматы, Қазақстан,
E-mail: baidaulet123@gmail.com, https://orcid.org/0009-0000-5510-3590.

Аннотация: Анаэробты ашыту – құрамындағы шикізаттың өзгеруі, технологиялық параметрлердің ауытқуы және микробтық қауымдастық динамикасы салдарынан жиі тұрақсыздыққа ұшырайтын көпқырлы биохимиялық процесс. ADM1 сияқты механистік модельдер теориялық тұрғыдан маңызды негіз ұсынғанымен, оларды қолдану айтарлықтай параметрлік баптауды талап етеді және өнеркәсіптік ауқымдағы реакторлардың тез өзгеретін мінез-құлқын толық көрсете бермейді. Осы шектеулерді жеңілдету үшін бұл зерттеуде бақылаусыз оқыту әдістері, атап айтқанда кластерлік талдау, толық ауқымды анаэробты биореакторларда қайталанатын операциялық режимдерді анықтау үшін қолданылды. Коммерциялық нысандардан ұзақ мерзімді жиналған деректер ұшпа май қышқылдарының концентрациясын, биогаз құрамын және басқа да негізгі технологиялық көрсеткіштерді қамтитын көпөлшемді жиынтықтарға негізделді. Кластерлік талдау үш басым режимді анықтады: жоғары метан өнімділігімен сипатталатын тұрақты жұмыс күйі; орташа теңгерімсіздіктер байқалатын өтпелі кезеңдер; және ұшпа май қышқылдарының жиналуы мен метан мөлшерінің төмендеуімен ерекшеленетін тежелген күй. Бұл режимдер органикалық жүктеменің күрт артуы сияқты нақты технологиялық оқиғалармен жақсы сәйкес келді, әдістің ерте ескертетін құрал ретінде құндылығын дәлелдеді. Зерттеу бақылаусыз оқытудың анаэробты ашыту жүйелерін бақылауды жақсартуға, тұрақтылығын арттыруға және индустриялық процестерді неғұрлым проактивті басқаруға мүмкіндік беретінін көрсетеді.

Түйін сөздер: анаэробты ашу, бақылаусыз оқыту, кластерлеу, процесс мониторингі, биогаз өндіру, операциялық режимдер, машиналық оқыту

© Н. Карымсакова[1], А. Болтабоева[2,3,4,*], Д. Тұрмаханбет[2,3,4], М. Мауленбеков[2,3,5], Т. Абдирова[6], 2025

[1]Университет ALT им. М.Тынышпаева", Алматы, Казахстан;

[2]Институт Механики и Инженерии им. У. А. Джолдасбекова, Алматы, Казахстан;

[3]ТОО "Kazakhstan R&D Solutions", Алматы, Казахстан;

[4]Казахский Национальный Университет им. аль-Фараби, Алматы, Казахстан;

[5]Казахский национальный исследовательский технический университет им. К. Сатпаева, Алматы, Казахстан;

[6]ГКП на ПХВ «Областной кардиологический центр, Алматы, Казахстан.
E-mail: boltaboyeva_assiya3@live.kaznu.kz

ОБУЧЕНИЕ БЕЗ УЧИТЕЛЯ ДЛЯ ВЫЯВЛЕНИЯ КРИТИЧЕСКИХ РЕЖИМОВ В ПРОИЗВОДСТВЕ ВОЗОБНОВЛЯЕМОЙ ЭНЕРГИИ

Карымсакова Нургуль — PhD, научный сотрудник, Университет ALT им. М.Тынышпаева", Алматы, Казахстан,
E-mail:n.karymsakova@alt.edu.kz, ORCID ID: https://orcid.org/0000-0001-8187-2369;
Болтабоева Асия — исследователь, Институт Механики и Инженерии им. У.А. Джолдасбекова, научный сотрудник в ТОО «Kazakhstan R&D Solutions», Докторант 3-го года в Факультете Информационных Технологий и Искусственного Интеллекта, Казахский Национальный Университет им. аль-Фараби, Алматы, Казахстан,
E-mail: boltaboyeva_assiya3@live.kaznu.kz, https://orcid.org/0000-0002-7279-9910;
Тұрмаханбет Динара — инженер IoT в ТОО «Kazakhstan R&D Solutions», студентка 4 курс, Факультет Информационных Технологий и Искусственного Интеллекта, Казахский Национальный Университет им. аль-Фараби; Институт Механики и Инженерии им. У.А. Джолдасбекова, Алматы, Казахстан,
E-mail: turmahanbetdinara@gmail.com, https://orcid.org/0009-0004-8388-4979;
Мауленбеков Марлен — инженер IoT в ТОО «Kazakhstan R&D Solutions», студент 4 курс, Факультет электроники и инженерии, Казахский национальный исследовательский технический университет им. К. Сатпаева; Институт Механики и Инженерии им. У.А. Джолдасбекова, Алматы, Казахстан,
E-mail: marlen.maulenbek@gmail.com , https://orcid.org/0009-0002-0703-7634;
Байдаулетова Алия — невролог и сомнолог, кандидат медицинских наук, Медицинский центр «Нейроклиника», Алматы, Казахстан,
E-mail: baidaulet123@gmail.com, https://orcid.org/0009-0000-5510-3590.

Аннотация: Анаэробное сбраживание представляет собой многокомпонентный биохимический процесс, который нередко становится нестабильным из-за изменений в составе сырья, колебаний технологических параметров и динамики микробного сообщества. Хотя механистические модели, такие как ADM1, дают важную теоретическую основу, их применение требует значительной калибровки параметров и зачастую не отражает быстро меняющееся поведение промышленных реакторов. Чтобы преодолеть эти ограничения, в данном исследовании использованы методы обучения без учителя, в частности кластерный анализ, направленный на выявление повторяющихся эксплуатационных режимов в полноразмерных анаэробных

дигестерах. Длительные исторические данные с промышленных объектов включали многомерные показатели, такие как концентрации летучих жирных кислот, состав биогаза и ключевые параметры процесса. Анализ кластеров выявил три основных режима: стабильный режим с высокой метановой продуктивностью; переходные состояния с умеренными отклонениями; и ингибированные режимы, характеризующиеся накоплением летучих жирных кислот и снижением содержания метана. Важно, что обнаруженные режимы хорошо совпадали с реальными производственными ситуациями, включая резкие повышения органической нагрузки, что подтвердило ценность метода как инструмента раннего предупреждения. Исследование показывает, что методы обучения без учителя могут существенно повысить качество мониторинга, укрепить стабильность процесса и способствовать более проактивному управлению промышленными системами анаэробного сбраживания.

Ключевые слова: анаэробное сбраживание, обучение без учителя, кластеризация; мониторинг процессов, производство биогаза, эксплуатационные режимы, машинное обучение, летучие жирные кислоты

Финансирование: *Это исследование было профинансировано Комитетом науки Министерства науки и высшего образования Республики Казахстан (Грант № AP26103739) и ТОО «Kazakhstan R&D Solutions».*

Введение. Анаэробное сбраживание признано одной из ключевых биотехнологий для утилизации органических отходов и производства возобновляемой энергии, преимущественно в форме биогаза (Zhu et al., 2025; Meola et al., 2025; Rutland et al., 2025; Zhu et al., 2025). Этот многостадийный биохимический процесс, протекающий при отсутствии кислорода с участием разнообразного микробного сообщества, обеспечивает преобразование различных типов органического сырья — включая пищевые отходы, осадки сточных вод и сельскохозяйственные остатки — в метан (CH_4) и диоксид углерода (CO_2) (Zhu et al., 2025; Meola et al., 2025; Obileke et al., 2024; Ali et al., 2025).

Несмотря на значительный потенциал, процессы анаэробного сбраживания обладают высокой сложностью и выраженной нелинейностью, так как зависят от множества взаимосвязанных факторов: состава субстрата, органической нагрузки (OLR), гидравлического времени удерживания (HRT), температуры, уровня pH и динамики микробного сообщества (Rutland et al., 2025; Tasmurzayev et al., 2022; Simeonov et al., 2025; Sidi Habib et al., 2024; Long et al., 2025). Нестабильность реактора, возникающая из-за колебаний характеристик сырья или накопления промежуточных продуктов, таких как летучие жирные кислоты (ЛЖК), может приводить к снижению выработки метана, ухудшению качества биогаза и даже к полному нарушению процесса

(Zhu et al., 2025; Zhu et al., 2025; Ali et al., 2025; Jia et al., 2025; Beschkov et al., 2025).

Высокие скорости разложения пищевых отходов часто вызывают накопление ЛЖК, что снижает pH ниже оптимальных значений для метаногенов и, как следствие, ингибирует метаногенез (Tasmurzayev et al., 2022; Lee et al., 2023; Amangeldy et al., 2023; Tongco et al., 2025). Такое нарушение баланса может привести к технологическим сбоям, поэтому эффективный мониторинг и управление процессом являются критически важными для поддержания стабильности, оптимизации производительности и обеспечения экономической эффективности промышленных систем анаэробного сбраживания (Zhu et al., 2025; Rutland et al., 2025; Tasmurzayev et al., 2022; Simeonov et al., 2025; Sidi et al., 2024).

Традиционные методы контроля процессов анаэробного сбраживания, как правило, основаны на периодическом отборе проб и трудоёмких лабораторных анализах (Rutland et al., 2025; Almansa et al., 2023; Sappl et al., 2023). Эти методы не обеспечивают получение данных в реальном времени, что затрудняет оперативное реагирование на внезапные изменения или развитие нестабильности (Rutland et al., 2025; Jia et al., 2023).

Механистические модели, такие как общепринятая Anaerobic Digestion Model No. 1 (ADM1) (Meola et al., 2025; Sappl et al., 2023; Dekhici et al., 2025: Amangeldy et al., 2024), позволяют глубоко понять биохимические и физико-химические процессы (Moradvandi et al., 2025), однако требуют значительной калибровки множества параметров и больших вычислительных ресурсов (Meola et al., 2025; Sappl et al., 2023). Их применение в промышленных масштабах ограничено необходимостью дорогостоящих офлайн-измерений и высокой чувствительностью к ошибкам сенсоров (Meola et al., 2025; Rutland et al., 2025; Sappl et al., 2023). Упрощённые механистические модели, в свою очередь, часто не способны адекватно описывать сложную нелинейную динамику, наблюдаемую в реальных системах анаэробного сбраживания (Tasmurzayev et al., 2022; Sappl et al., 2023; Dekhici et al., 2025: Amangeldy et al., 2024).

Модели, разработанные на основе лабораторных или пилотных данных, нередко плохо масштабируются на промышленные условия из-за большей сложности микробных сообществ и изменчивости эксплуатационных параметров (Dekhici et al., 2025). Это подчёркивает наличие разрыва в способности точно предсказывать и интерпретировать поведение промышленных анаэробных реакторов (Long et al., 2025; Amangeldy et al., 2024).

В последние годы методы, основанные на данных, особенно машинное обучение (ML) и искусственный интеллект (AI), стали мощным инструментом для решения задач мониторинга и оптимизации процессов анаэробного сбраживания (Sappl et al., 2023; Dekhici et al., 2025: Amangeldy et al., 2024).

В отличие от механистических моделей, алгоритмы машинного обучения способны выявлять сложные нелинейные зависимости между входными и выходными параметрами без необходимости глубокого понимания внутренних биохимических механизмов. Они позволяют предсказывать показатели, которые ранее требовали дорогостоящих анализов, тем самым повышая эффективность и стабильность работы систем (Tasmurzayev et al., 2022).

Различные алгоритмы машинного обучения — такие как искусственные нейронные сети (ANNs), регрессия на опорных векторах (SVR), случайные леса и градиентный бустинг — успешно применялись для прогнозирования выработки биогаза и метана при различных условиях [9]. Машинное обучение также обеспечивает эффективные инструменты для анализа больших массивов эксплуатационных данных (Sappl et al., 2023), выявления аномалий и изучения динамики поведения систем (Sappl et al., 2023; Dekhici et al., 2025: Amangeldy et al., 2024). Эти возможности особенно важны для установок анаэробного сбраживания, работающих с переменным сырьём и подверженных внешним воздействиям, таким как температурные колебания (Moradvandi et al., 2025).

Несмотря на растущее применение методов машинного обучения в анаэробном сбраживании, большинство исследований сосредоточено на прогнозировании производительности или выявлении отказов (Sidi et al., 2024; Sappl et al., 2023; Dekhici et al., 2025: Amangeldy et al., 2024). Потенциал методов обучения без учителя, в частности кластерного анализа, для систематического выявления дискретных и повторяющихся эксплуатационных режимов в промышленных анаэробных реакторах остаётся недостаточно изученным (Sappl et al., 2023).

Настоящее исследование направлено на восполнение данного пробела. Основная цель работы — определить, могут ли методы кластеризации использоваться для идентификации различных эксплуатационных режимов в промышленных анаэробных реакторах на основе многомерных данных, включающих концентрации летучих жирных кислот, состав биогаза (CH_4, CO_2), а также другие ключевые технологические параметры — pH, органическую нагрузку и гидравлическое время удерживания (Long et al., 2025; Tongco et al., 2025; Sappl et al., 2023; Amangeldy et al., 2024; Moradvandi et al., 2025).

Выявление таких режимов позволит операторам глубже понимать базовое поведение реактора при изменяющихся условиях, принимать более обоснованные технологические решения и осуществлять проактивное управление. Научная новизна исследования заключается в применении кластерного анализа для характеристики и классификации сложных, высокоразмерных промышленных данных анаэробного сбраживания, что обеспечивает новые возможности для анализа динамики систем и способствует разработке более надёжных стратегий управления (Sappl et al., 2023).

Схема выполнения исследования представлена на рисунке 1.

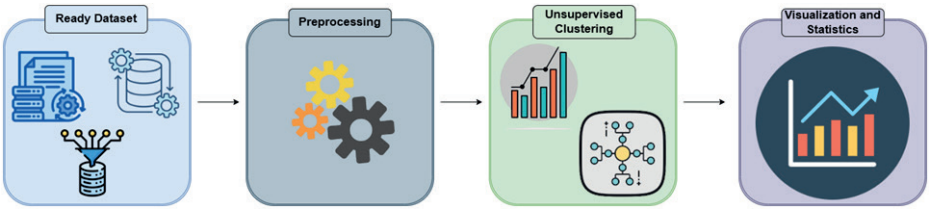


Рисунок 1 - Схема методологии исследования.

Материалы и методы. *2.1. Архитектура системы и сбор данных*

Основой данного исследования послужил набор данных, полученный в результате мониторинга работы восьми биореакторов на протяжении 127 дней. Эти данные были структурированы таким образом, чтобы отразить ключевые биохимические и эксплуатационные аспекты процесса анаэробного сбраживания. В качестве основных входных переменных для анализа были выбраны суточные концентрации летучих жирных кислот (ЛЖК) — чувствительных маркеров микробного дисбаланса, а также состав биогаза (CH_4 и CO_2), напрямую отражающий эффективность протекания процесса. Для получения более полной картины к этим показателям были добавлены эксплуатационные параметры, такие как pH, температура и органическая нагрузка, в результате чего была сформирована итоговая матрица, включающая 741 наблюдение и 15 информативных переменных. Переменные, включённые в набор данных, представлены в таблице 1.

Таблица 1. Переменные, используемые в биореакторах

VFAg	Уксусная, пропионовая, изомасляная, н-масляная, изовалериановая, валериановая, гексановая, гептановая кислоты, а также суммарные ЛЖК (для каждого реактора D1–D8)	г·Л$^{-1}$	127 дней × 8 реакторов	Промежуточные метаболиты, маркеры нестабильности
Названиие параметра	Включенные переменные	Единицы измерения	Колиечсивр записей	Роль в анализе
Gas comp	CH_4, CO_2 (объёмные доли), общий CH_4, общий CO_2, стандартное отклонение (CH_4, CO_2)	объем %	Агрегировано по эксплуатационным интервалам (~127 дней)	Выходная эффективность, состав и стабильность биогаза
Process parameters (merged)	pH, температура, органическая нагрузка (OLR), гидравлическое время удерживания (HRT)	–	127 дней × 8 реакторов	Эксплуатационный контекст и управляющие переменные

Первым шагом в нашей работе стала подготовка данных для анализа. Мы выполнили ежедневное выравнивание временных рядов и применили

метод k-ближайших соседей (KNN-imputation) для обработки пропущенных значений. Затем была проведена z-стандартизация всех переменных, чтобы привести их к общей шкале и обеспечить корректную работу алгоритмов машинного обучения. Далее мы перешли к созданию новых признаков (feature engineering), сформировав составные индексы, которые более точно характеризуют состояние процесса. В частности, были рассчитаны ключевые показатели, такие как соотношение пропионовой и уксусной кислот, соотношение CO_2 к CH_4 и логарифм общей концентрации летучих жирных кислот. Эти индексы позволили классифицировать состояние реактора на три режима: стабильный, начальный дисбаланс и ацидогенный стресс.

Для автоматического обнаружения отклонений от нормальной работы была реализована модель обучения без учителя. Мы использовали алгоритмы кластеризации, такие как K-Means и DBSCAN, для выявления аномалий, не соответствующих стабильным эксплуатационным режимам. В результате система автоматически формирует оповещения об аномалиях для оператора.

На заключительном этапе внимание было уделено предоставлению оператору инструментов для принятия решений. Система генерирует визуализации на основе анализа главных компонент (PCA), статистические сводки и прямые уведомления. Основываясь на этой информации, оператор может своевременно предпринимать корректирующие действия, например регулировать подачу сырья или щёлочность, чтобы предотвратить сбой процесса и стабилизировать работу биореактора.

Предложенная архитектура, представленная на рисунке 1, была разработана для бесшовной интеграции в промышленные установки анаэробного сбраживания и согласована с иерархическими структурами управления, обычно применяемыми в процессных отраслях. На уровне 0 (L0) информация о процессе поступает от сертифицированных датчиков и исполнительных механизмов, которые изначально интегрированы в инфраструктуру распределённой системы управления (DCS) или системы диспетчерского контроля и сбора данных (SCADA).

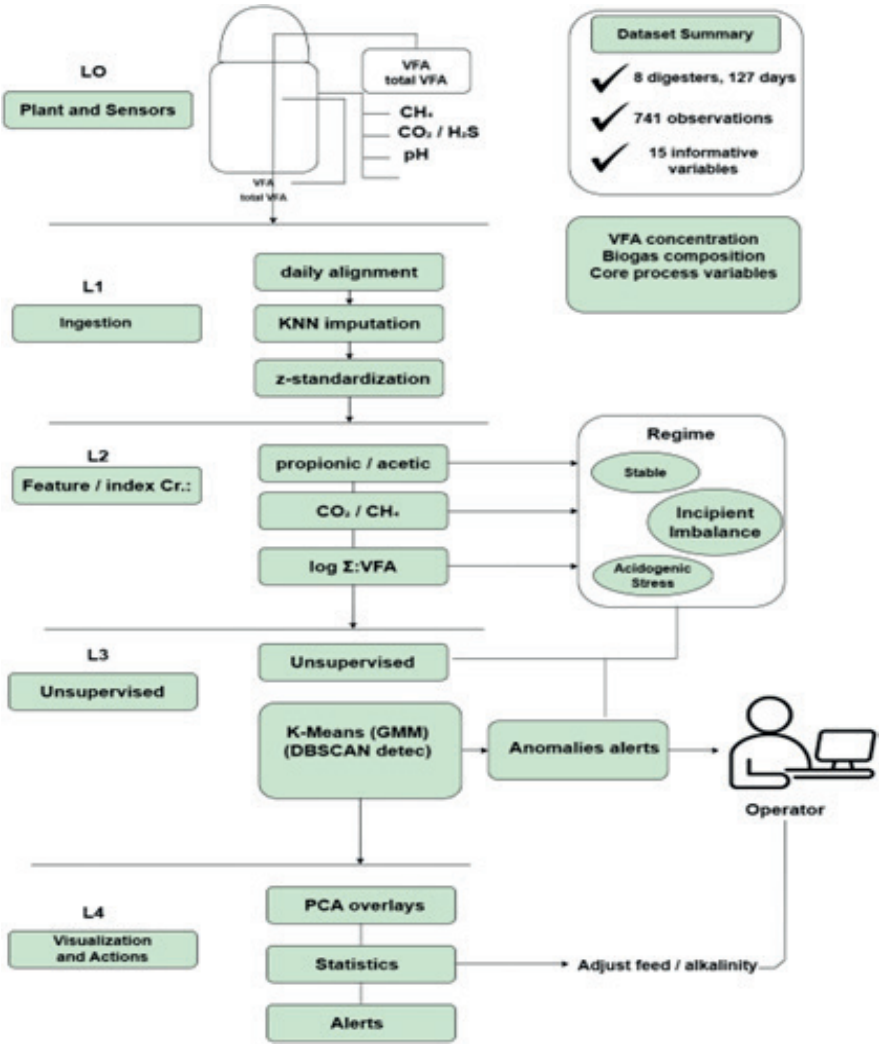


Рисунок 2 - Иерархическая архитектура мониторинга процесса биогазопроизводства и обнаружения аномалий.

Непрерывные измерения летучих жирных кислот (VFA), pH и состава биогаза (CH_4, CO_2, H_2S) обеспечивают базовые показатели работы реактора; при этом протоколы калибровки датчиков строго регистрируются, а процедуры контроля качества автоматически помечают периоды калибровки и ненадёжные значения. На уровне 1 (L1) уровень ввода данных реализован в среде хранения данных предприятия (plant historian) и промежуточном программном обеспечении для интеграции данных. Здесь автоматизированные конвейеры согласовывают данные с дневным разрешением, применяют реконструкцию пропусков методом k-ближайших соседей (KNN) и выполняют z-стандартизацию для всех сигналов.

Интерфейсы с лабораторными информационными системами (LIMS) дополнительно обеспечивают автоматическую загрузку офлайн-результатов анализов, что гарантирует сопоставимость лабораторных данных и данных онлайн-датчиков, а также преемственность между экспериментальной валидацией и промышленной эксплуатацией. На уровне 2 (L2) формирование признаков выполняется на периферийных серверах или в виртуализированных средах, напрямую подключённых к сетевой инфраструктуре предприятия. Производные показатели, включая соотношение пропионовой и уксусной кислот, соотношение CO_2/CH_4 и логарифм суммарной концентрации ЛЖК, вычисляются почти в реальном времени. Эти вычисленные индексы повышают чувствительность к метаболическому дисбалансу по сравнению с исходными измерениями, а контейнеризированные конвейеры обеспечивают переносимость между разнородным оборудованием и масштабируемость -от пилотных установок с одним реактором до многореакторных промышленных комплексов.

На уровне 3 (L3) модели обучения без учителя внедряются в контролируемой среде машинных операций (MLOps). Алгоритмы кластеризации, такие как K-Means и гауссовые смесевые модели (GMM), классифицируют эксплуатационные состояния как стабильное, начальный дисбаланс или ацидогенный стресс, в то время как DBSCAN используется для выявления аномалий и выбросов. Механизмы обнаружения дрейфа данных непрерывно отслеживают изменения распределений и инициируют повторное обучение моделей, чтобы сохранять точность прогнозирования при изменении качества сырья или условий работы. Важно отметить, что все модели функционируют в консультативном режиме, обеспечивая совместимость с системами безопасности и сохраняя полномочия оператора. На уровне 4 (L4) результаты моделирования интегрируются в интерфейс оператора через SCADA и человеко-машинные интерфейсы (HMI). Визуальная аналитика в форме наложений анализа главных компонент (PCA), статистических панелей и автоматических уведомлений об аномалиях обеспечивает наглядную интерпретацию результатов моделей. Оповещения распространяются как в операторской, так и по защищённым мобильным каналам, что обеспечивает своевременное информирование, а рекомендации - например, корректировка подачи сырья или дозирования щёлочи - передаются в консультативном виде, поддерживая проактивные действия без нарушения автономности системы.

С точки зрения эксплуатации, архитектура соответствует промышленным стандартам кибербезопасности (IEC 62443) и использует строгие механизмы контроля доступа для предотвращения несанкционированных изменений. Процедуры обслуживания включают плановую перекалибровку датчиков, автоматическую проверку данных и постоянный мониторинг состояния системы. Ключевые показатели эффективности (KPI) установлены на двух уровнях: эксплуатационные KPI - точность обнаружения, уровень

ложных срабатываний и время раннего предупреждения о дисбалансе; и аналитические KPI - разделимость кластеров, объяснённая дисперсия в PCA и чувствительность обнаружения аномалий. В совокупности эти показатели обеспечивают количественную основу для оценки эффективности и непрерывного совершенствования системы. В целом, данный путь промышленного внедрения демонстрирует, что архитектура является не только теоретически обоснованной, но и практически реализуемой. Она поддерживает воспроизводимую интеграцию в существующие производственные инфраструктуры, гарантирует надёжность и безопасность в реальных условиях и обеспечивает масштабируемость для нескольких реакторов и производственных площадок. Связывая экспериментальные наборы данных с системами оперативного управления, архитектура формирует чёткий путь к промышленному применению методов обучения без учителя для выявления режимов в процессах анаэробного сбраживания.

2.2. Предобработка данных

Для обеспечения сопоставимости переменных, имеющих разные шкалы и физические единицы измерения, все непрерывные признаки были подвергнуты z-стандартизации перед кластеризацией. Этот шаг необходим, поскольку алгоритмы, такие как K-Means, основаны на евклидовых расстояниях, которые сильно зависят от масштаба переменных. Без нормализации признаки с большими числовыми диапазонами (например, общие ЛЖК в г·л$^{-1}$) будут доминировать над признаками с малыми диапазонами (например, pH), что приведёт к искажённым кластерам. Стандартизация преобразует каждый признак к нулевому среднему и единичной дисперсии, уравнивая их влияние и обеспечивая, чтобы кластеризация отражала внутреннюю структуру данных, а не масштаб измерений (Jain, 2010). Преобразование определяется выражением:

$$z = \frac{x-p}{\sigma},$$

где: x — наблюдаемое значение,

μ — среднее значение,

σ — стандартное отклонение признака.

После этой процедуры все признаки становятся безразмерными и центрируются вокруг нуля, что позволяет корректно сравнивать разнородные биохимические и эксплуатационные переменные.

Признаки, имевшие более 50 % пропущенных значений, были исключены, а оставшиеся пробелы заполнены методом множественной импутации k-ближайших соседей (k = 3). Все непрерывные переменные подвергались z-стандартизации для устранения масштабных искажений. Были дополнительно рассчитаны диагностические индексы: соотношение пропионовой и уксусной кислот как показатель ингибирования метаногенеза; соотношение CO_2/CH_4 как индикатор глубины сбраживания; натуральный

логарифм общей концентрации ЛЖК для коррекции правостороннего смещения распределений; и составной индекс нестабильности, определённый как нормализованная сумма z-оценок pH, ΣVFA и CH_4.

Для снижения шумов и уменьшения размерности применялся анализ главных компонент (PCA); сохранялись компоненты, объясняющие не менее 95 % общей дисперсии (обычно четыре–пять). Итоговая матрица признаков включала 741 наблюдение и 15 информативных переменных, представляющих ключевые промежуточные продукты, конечные продукты и эксплуатационные параметры анаэробного процесса.

После выделения четырёх режимов каждый кластер был профилирован по медианам, межквартильным диапазонам и коэффициентам вариации для всех исходных и производных переменных. Критерий Крускала–Уоллиса ($p < 0.01$) подтвердил значимые различия между кластерами для общей концентрации ЛЖК, соотношения пропионовой и уксусной кислот, содержания CH_4 и составного индекса нестабильности. Пост-хок тест Данна с поправкой Хольма выявил наиболее контрастные пары, а обобщённые значения η^2 варьировали от 0.12 (для pH) до 0.54 (для индекса нестабильности), что указывает на средние и большие эффекты. Проекция центроидов кластеров на первые две главные компоненты показала чёткое разделение режима III, характеризующегося ΣVFA выше 1 $г \cdot л^{-1}$ и CH_4 ниже 55 %, что соответствует ацидогенной фазе. Временное сопоставление меток кластеров показало, что переходы от режима I к режиму III совпадали с резкими скачками OLR, подчёркивая технологическую интерпретируемость кластеров и демонстрируя применимость обучения без учителя для мониторинга и управления промышленными анаэробными реакторами в реальном времени.

2.3. Неконтролируемая кластеризация

Для выявления скрытых эксплуатационных режимов в процессе анаэробного сбраживания были применены три метода неконтролируемой кластеризации: K-Means, DBSCAN и гауссовые смесевые модели (GMM). Кластеризация K-Means была сначала оценена при варьировании количества кластеров и анализе кривой внутрикластерной инерции (рисунок 3). Критерий «локтя» показал выраженный излом при $k = 3$, что также подтвердилось максимальным средним коэффициентом силуэта = 0,43. Такое разбиение дало три компактные и хорошо разделённые группы, соответствующие различным состояниям процесса.

Дополнительно применялся DBSCAN (Density-Based Spatial Clustering of Applications with Noise) для исследования плотностной структуры набора данных. Анализ чувствительности параметров показал, что порог $\varepsilon \approx 0{,}6$ является оптимальным, исходя из графика k-расстояний (рисунок 5). При этом значении DBSCAN объединил большинство наблюдений в один основной кластер и выделил подмножество точек, преимущественно из ацидогенного режима, как шум. Этот результат показывает, что DBSCAN более подходит для обнаружения аномалий, чем для полной классификации режимов (рисунок 2).

Наконец, были протестированы гауссовые смесевые модели (GMM), чтобы учесть возможное перекрытие распределений в пространстве данных. GMM сошлась к тому же трёхмодовому решению, что и K-Means, подтверждая устойчивость трёхкластерной структуры. В совокупности комбинация разбиений по признакам, плотностных и вероятностных подходов подтвердила существование трёх интерпретируемых режимов: стабильная работа, начальный дисбаланс и ацидогенный стресс.

Результаты и обсуждение. Очищенный набор данных содержал 741 наблюдение, выровненное по дням, с пятнадцатью z-стандартизированными переменными, представляющими концентрации летучих жирных кислот, состав биогаза и эксплуатационные параметры. Анализ главных компонент показал, что первые пять компонентов объясняют 95,6 % дисперсии, что подтверждает сохранение основной структуры данных в сокращённом пространстве и обеспечивает интерпретируемую визуализацию результатов кластеризации.

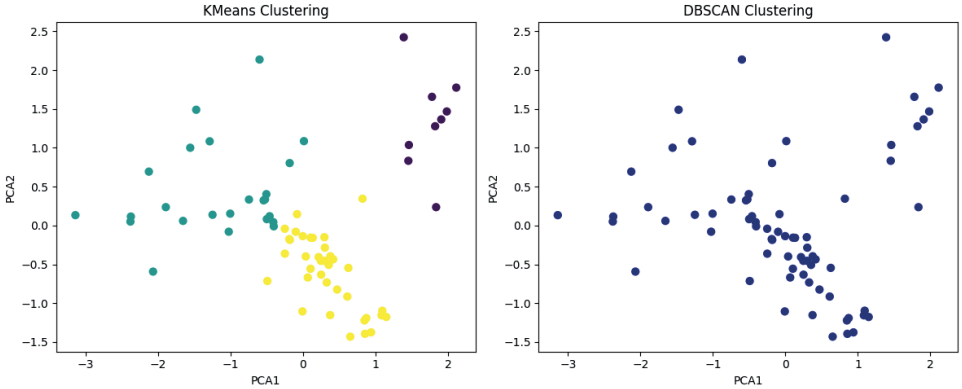


Рисунок 3 - Сравнение результатов кластеризации на основе PCA: метод K-Means выделяет три компактных режима, тогда как DBSCAN объединяет большинство точек в один кластер, помечая некоторые аномалии как шум.

На рисунке 3 представлено сравнение методов кластеризации K-Means и DBSCAN, применённых к данным, преобразованным методом анализа главных компонент (PCA). Алгоритм K-Means сформировал три компактных и хорошо разделённых кластера, соответствующих различным эксплуатационным состояниям. В противоположность этому, алгоритм DBSCAN при выбранном параметре ε объединил большинство точек в один плотный кластер и пометил лишь часть выборок как шум. Такое поведение показывает, что кластеризация, основанная на центроидах (K-Means), более подходит для классификации режимов, тогда как плотностная кластеризация (DBSCAN) эффективнее работает как инструмент обнаружения аномалий.

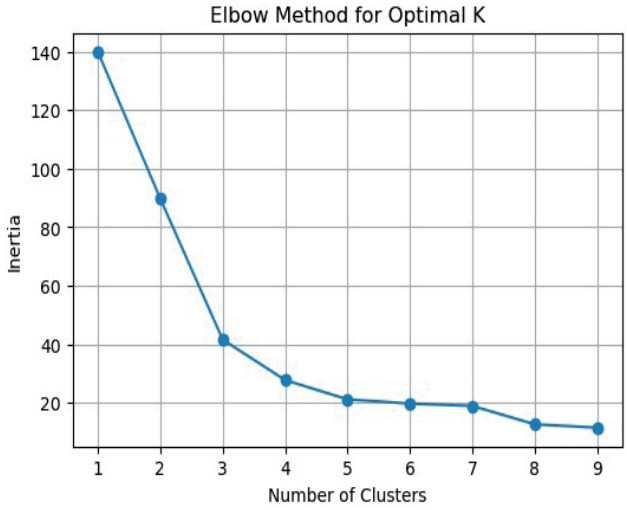


Рисунок 4 - Кривая метода «локтя» для определения оптимального числа кластеров (k = 3 выбрано как наилучший баланс).

Надёжность решения, полученного методом K-Means, подтверждается результатами метода «локтя», представленными на рисунке 4. Наблюдалось резкое снижение инерции до значения k = 3, после чего кривая выравнивалась. Этот «локоть» совпадает с максимальным коэффициентом силуэта 0,43, что подтверждает оптимальность трёхкластерного решения. Гауссовы смесевые модели независимо сошлись к трёхмодальному распределению, что дополнительно укрепляет данный вывод.

Анализ исходных технологических переменных позволил интерпретировать три выделенных режима. Первый кластер отражал стабильный метаногенез, при котором концентрация метана составляла около 66 об.% и суммарное содержание ЛЖК было ниже 0,5 г/л. Второй кластер соответствовал начальному дисбалансу, характеризующемуся концентрацией метана около 60 об.%, уровнем ЛЖК в пределах 0,5–1,0 г/л и соотношением пропионовой к уксусной кислоте выше 0,2. Третий кластер представлял ацидогенный стресс, для которого характерны ΣЛЖК свыше 1 г/л и содержание метана ниже 55 об.%.

Непараметрический критерий Крускала–Уоллиса подтвердил наличие статистически значимых различий между кластерами по показателям метана, диоксида углерода и ЛЖК ($p < 0,01$). Обобщённые значения η^2 варьировали от 0,34 до 0,54, что соответствует эффектам средней и высокой величины.

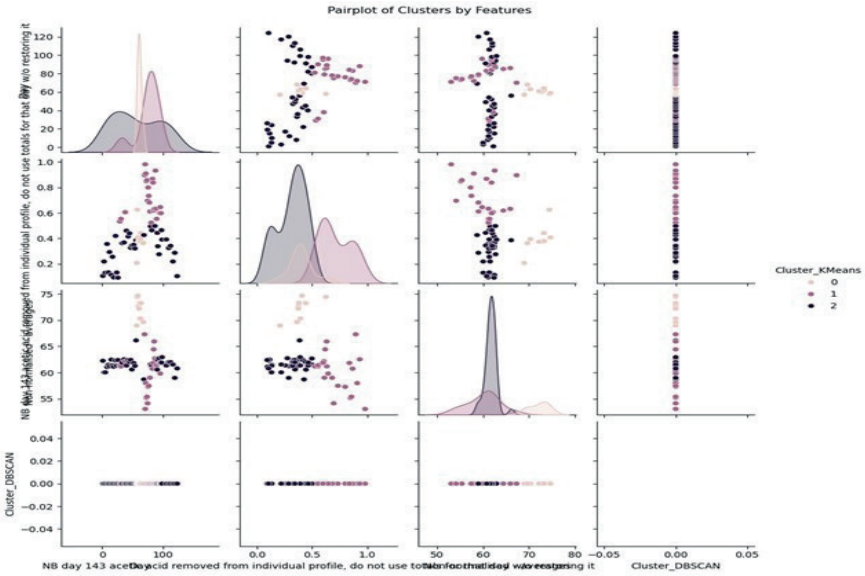


Рисунок 5 - Парный график распределений признаков по кластерам, показывающий различия и разделение между исходными и производными технологическими переменными для кластеров, полученных методом K-Means.

Рисунок 5 (парный график признаков) иллюстрирует разделение выделенных режимов в пространстве исходных переменных. Он показывает, что концентрация метана, суммарное содержание ЛЖК и составной индекс нестабильности являются наиболее информативными признаками, определяющими различие между режимами. Несмотря на высокую наглядность, большой размер данного графика предполагает его размещение в дополнительном материале, тогда как в основном тексте можно сослаться лишь на его ключевые результаты.

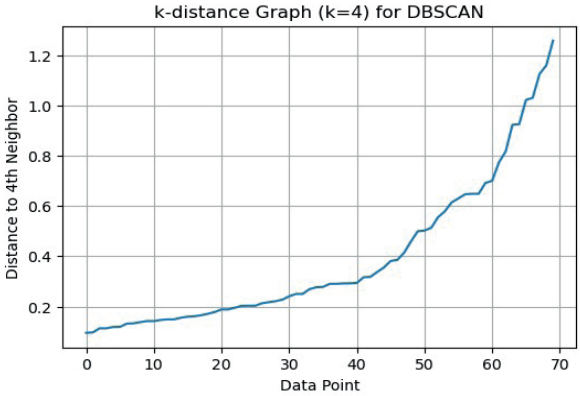


Рисунок 6 - График k-расстояний для DBSCAN (k = 4), указывающий ε ≈ 0,6 как оптимальный порог для обнаружения аномалий.

Поведение алгоритма DBSCAN было дополнительно исследовано с использованием графика k-расстояний, представленного на рисунке 6, который показал оптимальное значение порога ε около 0,6. При этом значении DBSCAN сгруппировал большинство наблюдений в один кластер, в основном соответствующий стабильному режиму, при этом выявив восемнадцать выбросов, преимущественно связанных с состоянием ацидогенного стресса. Этот результат подтверждает полезность модели для обнаружения аномалий, а не для систематической классификации режимов.

Наконец, сопоставление меток кластеров с временной шкалой эксплуатации показало, что переходы от стабильного к ацидогенному режиму совпадали с резкими увеличениями органической нагрузки, зарегистрированными операторами установки. Такое совпадение подтверждает интерпретируемость результатов кластеризации: модель обучения без учителя не только выявила закономерности, ранее распознаваемые экспертами, но и показала потенциал в качестве автоматизированной системы раннего предупреждения. На практике присвоение новых наблюдений ацидогенному кластеру в режиме реального времени позволит операторам заблаговременно принимать корректирующие меры — например, регулировать скорость подачи субстрата или щёлочность — прежде чем произойдёт снижение качества биогаза.

Несмотря на обнадёживающие результаты, исследование имеет определённые ограничения. Оно охватывало период 127 дней и восемь реакторов, что привело к относительно невысоким значениям коэффициента силуэта (< 0,5) и частичному перекрытию кластеров. Включение дополнительных показателей, таких как щелочность, аммиак или окислительно-восстановительный потенциал, а также расширение периода наблюдения и тестирование альтернативных методов кластеризации (например, fuzzy C-means или скрытых марковских моделей) могли бы повысить чёткость границ между режимами и точнее отразить переходные состояния.

Тем не менее, полученные результаты убедительно демонстрируют, что методы обучения без учителя, особенно кластеризация, основанная на центроидах, способны выявлять значимые эксплуатационные режимы в промышленных анаэробных реакторах и могут служить основой для прогностического мониторинга и поддержки принятия решений.

Заключение. Данное исследование показывает, что методы обучения без учителя, в частности кластерный анализ, могут эффективно выявлять повторяющиеся эксплуатационные режимы в полноразмерных анаэробных реакторах на основе многомерных технологических данных. Анализ концентраций летучих жирных кислот, состава биогаза и ключевых эксплуатационных параметров позволил выделить три различных режима — стабильный, переходный и ингибированный, — которые тесно коррелируют с эксплуатационными событиями, такими как резкие повышения органической

нагрузки. Это подтверждает пригодность методов кластеризации в качестве инструмента раннего предупреждения и поддержки принятия решений для операторов установок.

С практической точки зрения полученные результаты демонстрируют, что идентификация режимов на основе данных помогает операторам предвидеть развитие нестабильности до снижения выхода метана, что позволяет своевременно предпринимать корректирующие действия, повышающие стабильность процесса, качество биогаза и общую эффективность работы установки. Внедрение таких инструментов в системы мониторинга может снизить эксплуатационные риски, уменьшить экономические потери и способствовать более надёжному производству возобновляемой энергии.

Тем не менее, анализ имел определённые ограничения, связанные с относительно коротким периодом наблюдений (127 дней), ограниченным числом реакторов и отсутствием дополнительных физико-химических параметров, таких как щелочность, аммиак или окислительно-восстановительный потенциал. Эти факторы могли повлиять на чёткость границ кластеров и ограничить возможность обобщения результатов на более широкий диапазон условий эксплуатации.

В дальнейшем исследования должны быть направлены на расширение периода наблюдений, включение более разнообразных установок и интеграцию более широких наборов сенсорных данных, а также на изучение гибридных подходов, сочетающих механистические знания с адаптивностью методов, основанных на данных. Развитие кластеризации в реальном времени и создание интуитивно понятных систем поддержки операторских решений станут ключевыми шагами для внедрения этих методов в промышленную практику и обеспечения стабильных, эффективных и устойчивых систем производства биогаза.

References

Zhu X., Blanco E., Bhatti M., Borrion A. (2025) Improving anaerobic digestion process against acetate accumulation: insights into organic loading rates and nano magnetite additions. Chemical Engineering Journal. — Vol. 512. — P. 162641. DOI: 10.1016/j.cej.2025.162641 (In English).

Meola A., Wolf K., Weinrich S. (2025) Meta-tuning and fast optimization of machine learning models for dynamic methane prediction in anaerobic digestion. Bioresource Technology. — Vol. 432. — P. 132654. DOI: 10.1016/j.biortech.2025.132654 (In English).

Rutland H., You J., Liu H., Bowman K. (2025) Application of machine learning for FOS/TAC soft sensing in bio-electrochemical anaerobic digestion. Molecules. — Vol. 30. — P. 1092. DOI: 10.3390/molecules30051092 (In English).

Zhu X., Blanco E., Bhatti M., Borrion A. (2025) Promoting methanogenesis and stability in anaerobic digestion with nano magnetite under VFA-induced stress. Biofuel Research Journal. — Vol. 12. — P. 2432–2450. DOI: 10.18331/BRJ2025.12.2.5 (In English).

Obileke K., Makaka G., Tangwe S., Mukumba P. (2024) Improvement of biogas yields in an anaerobic digestion process via optimization technique. Environmental Development and Sustainability. — Vol. 27. — P. 15025–15051. DOI: 10.1007/s10668-024-04540-6 (In English).

Ali M., Pena C.B., De Vrieze J. (2025) Anaerobic co-digestion of sewage sludge and food waste: staging and carriers enhance system performance and process stability (In English).

Tasmurzayev N., Amangeldy B., Baigarayeva Z., Mansurova M., Resnik B., Amirkhanova G. (2022) Improvement of HVAC system using the intelligent control system, 2022 IEEE 7th International Energy Conference (ENERGYCON). DOI: 10.1109/ENERGYCON53164.2022.9830375 (In English).

Simeonov I., Chorukova E., Kabaivanova L. (2025) Two-stage anaerobic digestion for green energy production: a review. Processes. — Vol. 13. — P. 294. DOI: 10.3390/pr13020294 (In English).

Sidi Habib S., Torri S., Mol S.K. (2024) New methodologies for the optimization of operational parameters of bio gas power plants: a review. Journal of Renewable Energy and Environment. — Vol. 11. DOI: 10.30501/jree.2024.444163.1840 (In English).

Long F., Xu M., Liao W., Liu H. (2025) Machine learning for predicting and optimizing the performance of a commercial-scale anaerobic digester with diverse feedstocks and operating conditions. Bioresource Technology. — Vol. 435. — P. 132940. DOI: 10.1016/j.biortech.2025.132940 (In English).

Jia R., Song Y.-C., An Z., Kim K., Lee C.-Y., Bae B.-U. (2023) A new comprehensive indicator for monitoring anaerobic digestion: a principal component analysis approach. Processes. — Vol. 12. — P. 59. DOI: 10.3390/pr12010059 (In English).

Beschkov V.N., Angelov I.K. (2025) Volatile fatty acid production vs. methane and hydrogen in anaerobic digestion. Fermentation. — Vol. 11. — P. 172. DOI: 10.3390/fermentation11040172 (In English).

Lee J.-H., Lee J.-H., Kim S.-Y., Yoon Y.-M. (2023) Effect of addition of zero-valent iron (Fe) and magnetite (Fe_3O_4) on methane yield and microbial consortium in anaerobic digestion of food wastewater. Processes. — Vol. 11. — P. 759. DOI: 10.3390/pr11030759 (In English).

Amangeldy B., Tasmurzayev N., Mansurova M., Imanbek B., Sarsembayeva T. (2023) Design and development of IoT based medical cleanroom. Communications in Computer and Information Science. — Vol. 1864. DOI: 10.1007/978-3-031-41774-0_36 (In English).

Tongco J.V., Jannat M.A.H., Kim S., Park S., Hwang S. (2025) Process performance and biogas output: impact of fluctuating acetate concentrations on methanogenesis in horizontal anaerobic reactors. Energies. — Vol. 18. — P. 3120. DOI: 10.3390/en18123120 (In English).

Almansa X.F., Starostka R., Raskin L., Zeeman G., De Los Reyes F., Waechter J., Yeh D., Radu T. (2023) Anaerobic digestion as a core technology in addressing the global sanitation crisis: challenges and opportunities. Environmental Science & Technology. — Vol. 57. — P. 19078–19087. DOI: 10.1021/acs.est.3c05291 (In English).

Sappl J., Harders M., Rauch W. (2023) Machine learning for quantile regression of biogas production rates in anaerobic digesters. Science of the Total Environment. — Vol. 872. — P. 161923. DOI: 10.1016/j.scitotenv.2023.161923 (In English).

Dekhici B., Short M. (2025) Data-driven modelling of biogas production using multi-task Gaussian processes. Systems and Control Transactions. — P. 26–32 (In English).

Amangeldy B., Tasmurzayev N., Shinassylov S., Mukhanbet A., Nurakhov Y. (2024) Integrating machine learning with intelligent control systems for flow rate forecasting in oil well operations. Automation. — Vol. 5. DOI: 10.3390/automation5030021 (In English).

Moradvandi A., Heegstra S., Ceron-Chafla P., Schutter B.D., Abraham E., Lindeboom R.E.F. (2025) Model predictive control of feed rate for stabilizing and enhancing biogas production in anaerobic digestion under meteorological fluctuations. Journal of Process Control. — Vol. 147. — P. 103375. DOI: 10.1016/j.jprocont.2025.103375 (In English).

ACADEMIC SCIENTIFIC JOURNAL OF COMPUTER SCIENCE
ISSN 1991-346X
Volume 4. Namber 356 (2025). 203–215

https://doi.org/10.32014/2025.2518-1726.392

IRSTI 28.23.01
UDC 004.032.26

© A. Kulakayeva[1], E. Daineko[1], B. Medetov[2], A. Nurlankyzy[1,3*], 2025.
[1] International Information Technology University, Almaty, Kazakhstan;
[2] L.N. Gumilyov Eurasian National University, Astana, Kazakhstan;
[3] Satpayev University, Almaty, Kazakhstan.
E-mail: nurlankyzyaigulya@gmail.com

EVALUATION OF THE EFFECTIVENESS OF MODERN NEURAL NETWORK ARCHITECTURES FOR VAD UNDER LOW SNR RATIO CONDITIONS

Aigul Kulakayeva — PhD, Associate Professor of International Information Technology University, Almaty, Kazakhstan,
E-mail: a.kulakayeva@iitu.edu.kz, https://orcid.org/0000-0002-0143-085X;
Yevgeniya Daineko — PhD, Associate Professor, Vice Rector for Global Partnerships and Continuous Education of International Information Technology University, Almaty, Kazakhstan,
E-mail: y.daineko@iitu.edu.kz, https://orcid.org/0000-0001-6581-2622;
Bekbolat Medetov — PhD, Associate Professor of L.N. Gumilyov Eurasian National University, Astana, Kazakhstan,
E-mail: bm02@mail.ru, https://orcid.org/0000-0002-5594-8435;
Aigul Nurlankyzy — PhD, Senior lecturer, Satbaev University; Almaty University of Power Engineering and Telecommunications named after Gumarbek Daukeyev, Almaty, Kazakhstan,
E-mail: nurlankyzyaigulya@gmail.com; https://orcid.org/0000-0002-0791-8573.

Abstract. This study addresses the problem of enhancing the robustness of Voice Activity Detection (VAD) systems against acoustic noise, which is particularly relevant for practical applications in mobile devices and telecommunication systems. An experimental comparison of five hybrid neural network architectures (CNN+BiGRU, CNN+GRU, CNN+LSTM, CNN+BiLSTM, and CNN+TDNN) was conducted using the large-scale Kazakh Speech Corpus (KSC2), augmented with both synthetic and real noise from the ESC-50 dataset across a signal-to-noise ratio (SNR) range of –20 to +30 dB. Mel-Frequency Cepstral Coefficients (MFCC) were used to construct the feature space, providing informative representations of audio signals for subsequent machine learning. The experimental results showed that all tested architectures achieved high values of Accuracy and F1-score (above 99.3%) across the entire range of acoustic conditions. The CNN+BiGRU architecture demonstrated the best balance between precision, recall, and noise robustness, whereas CNN+TDNN stood out for its minimal computational cost

with comparable performance metrics. Error matrix analysis confirmed the models' ability to reliably distinguish between speech and noise, even under low signal-to-noise ratio (SNR) conditions. These findings indicate the high suitability of hybrid neural network architectures for addressing the VAD task in real-world noisy environments. CNN+BiGRU and CNN+TDNN are the most promising for practical deployment, enabling their use in mobile, embedded, and cloud-based speech systems.

Key words: voice activity detectors, signal-to-noise ratio, recurrent neural networks, convolutional neural networks, machine learning

© А. Кулакаева[1], Е. Дайнеко[1], Б. Медетов[2], А. Нурланкызы[1,3 *], 2025.
[1] Халықаралық ақпараттық технологиялар университеті, Алматы, Қазақстан;
[2] Л.Н. Гумилев атындағы Еуразия ұлттық университеті, Астана, Қазақстан;
[3] Сәтбаев Университеті, Алматы, Қазақстан.
E-mail: nurlankyzyaigulya@gmail.com

СИГНАЛ/ШУЫЛ ҚАТЫНАСЫ ТӨМЕН ЖАҒДАЙЛАРДА ЗАМАНАУИ НЕЙРОНДЫҚ ЖЕЛІЛІК VAD АРХИТЕКТУРАЛАРЫНЫҢ ТИІМДІЛІГІН БАҒАЛАУ

Айгуль Кулакаева — PhD, Халықаралық ақпараттық технологиялар университетінің қауымдастырылған профессоры, Алматы, Қазақстан,
E-mail: a.kulakayeva@iitu.edu.kz, https://orcid.org/0000-0002-0143-085X;
Евгения Дайнеко — PhD, Қауымдастырылған-профессор, Халықаралық ақпараттық технологиялар университетінің Жаһандық серіктестік және қосымша білім беру жөніндегі проректоры, Алматы, Қазақстан,
E-mail: y.daineko@iitu.edu.kz, https://orcid.org/0000-0001-6581-2622;
Бекболат Медетов — PhD, Л.Н. Гумилев Атындағы Еуразия ұлттық университетінің қауымдастырылған профессоры, Астана, Қазақстан,
E-mail: bm02@mail.ru, https://orcid.org/0000-0002-5594-8435;
Айгуль Нурланкызы — PhD, аға оқытушы, Сәтбаев университеті; Ғұмарбек Дәукеев атындағы Алматы Энергетика және байланыс университеті, Алматы, Қазақстан,
E-mail: nurlankyzyaigulya@gmail.com, https://orcid.org/0000-0002-0791-8573.

Аннотация. Бұл жұмыста дауыс белсенділігін анықтау жүйелерінің (VAD, Voice Activity Detection) акустикалық кедергілерге төзімділігін арттыру мәселесі қарастырылады. Бұл мәселе әсіресе мобильді құрылғылар мен телекоммуникациялық жүйелерде практикалық қолдану үшін өзекті болып табылады. Эксперименттік түрде бес гибридті нейрондық желілік архитектуралардың (CNN+BiGRU, CNN+GRU, CNN+LSTM, CNN+BiLSTM және CNN+TDNN) салыстырмалы талдауы жүргізілді. Зерттеу Kazakh Speech Corpus (KSC2) атты кең ауқымды қазақша сөйлеу корпусына негізделіп, оған ESC-50 мәліметтер базасынан алынған синтетикалық және нақты шу деңгейлері –20 дБ-ден +30 дБ-ге дейін қосылды. Белгілік кеңістікті қалыптастыру үшін мел-жиілік кепстральды коэффициенттері (MFCC) қолданылды, бұл

аудиосигналдарды одан әрі машиналық оқыту үшін ақпараттық тұрғыдан тиімді ұсынуға мүмкіндік берді. Эксперименттік нәтижелер барлық сынақтан өткен архитектуралардың акустикалық жағдайлардың барлық ауқымында жоғары дәлдік (Accuracy) және F1-score көрсеткіштеріне (99,3%-дан жоғары) қол жеткізетінін көрсетті. CNN+BiGRU архитектурасы дәлдік, толықтық және шуға төзімділік арасындағы ең жақсы тепе-теңдікті көрсетті, ал CNN+TDNN архитектурасы сапа көрсеткіштері салыстырмалы болғанымен, есептеу ресурстарын аз қажет етуімен ерекшеленеді. Қате матрицаларын талдау тіпті төмен сигнал/шум қатынасында да модельдердің сөйлеу мен шуды сенімді түрде ажырата алатынын растады. Алынған нәтижелер гибридті нейрожелі архитектураларының нақты, шуды көп ортада Voice Activity Detection (сөйлеу белсенділігін анықтау) мәселесін шешуге жоғары бейімділігін дәлелдейді. CNN+BiGRU және CNN+TDNN модельдері практикалық қолдану үшін ең перспективалы болып табылады және мұндай модельдерді мобильді, кірістірілген және бұлтты сөйлеу жүйелерінде пайдалануға мүмкіндік береді.

Түйін сөздер: сөйлеу белсенділігін анықтауыштар, сигнал/шуыл қатынасы, рекуррентті нейрондық желілер, конволюциялық нейрондық желілер, машиналық оқыту

© А. Кулакаева[1], Е. Дайнеко[1], Б. Медетов[2], А. Нурланкызы[1,3], 2025.

[1] Международный Университет Информационных Технологий, Алматы, Казахстан;

[2] Евразийский Национальный Университет им. Л.Н. Гумилева, Астана, Казахстан;

[3] Сатпаев Университет, Алматы, Казахстан.

E-mail: nurlankyzyaigulya@gmail.com

ОЦЕНКА ЭФФЕКТИВНОСТИ СОВРЕМЕННЫХ НЕЙРОСЕТЕВЫХ АРХИТЕКТУР VAD ПРИ НИЗКОМ ОТНОШЕНИИ СИГНАЛ/ШУМ

Айгуль Кулакаева — PhD, Ассоциированный профессор Международного университета информационных технологий, Алматы, Казахстан,
E-mail: a.kulakayeva@iitu.edu.kz, https://orcid.org/0000-0002-0143-085X;
Евгения Дайнеко — PhD, Ассоциированный-профессор, Проректор по глобальному партнерству и дополнительному образованию Алматы, Казахстан,
E-mail: y.daineko@iitu.edu.kz, https://orcid.org/0000-0001-6581-2622;
Бекболат Медетов — PhD, Ассоциированный профессор Евразийского национального университета им. Л.Н. Гумилева, Астана, Казахстан,
E-mail: bm02@mail.ru, https://orcid.org/0000-0002-5594-8435;
Айгуль Нурланкызы — PhD, старший преподаватель, Сатпаев университет; Алматинский университет энергетики и связи им. Гумарбека Даукеева, Алматы, Казахстан,
E-mail: nurlankyzyaigulya@gmail.com, https://orcid.org/0000-0002-0791-8573.

Аннотация. В настоящей работе рассмотрена задача повышения устойчивости систем обнаружения голосовой активности VAD (Voice Activity

Detection) к воздействию акустических помех, что особенно актуально для практического применения в мобильных устройствах и телекоммуникационных системах. Проведено экспериментальное сравнение пяти гибридных нейросетевых архитектур, таких как CNN+BiGRU, CNN+GRU, CNN+LSTM, CNN+BiLSTM и CNN+TDNN на большом корпусе казахской речи Kazakh Speech Corpus (KSC2) Институт умных систем и искусственного интеллекта Назарбаев Университета, дополненном синтетическим и реальным шумом из базы ESC-50 (уличные шумы, звуки животных, бытовые шумы) при уровне отношение сигнал/шум от –20 до +30 дБ. Для формирования признакового пространства использовались мел-частотные кепстральные коэффициенты (MFCC), что обеспечило информативное представление аудиосигналов для последующего машинного обучения. Экспериментальные результаты показали, что все протестированные архитектуры достигают высоких значений Accuracy и F1-score (более 99,3%) во всем диапазоне акустических условий. Архитектура CNN+BiGRU продемонстрировала наилучший баланс между точностью, полнотой и устойчивостью к шуму, тогда как CNN+TDNN выделяется минимальными вычислительными затратами при сопоставимых показателях качества. Анализ матриц ошибок подтвердил способность моделей надежно различать речь и шум даже при низких уровнях отношения сигнал/шум. Полученные результаты свидетельствуют о высокой пригодности гибридных нейросетевых архитектур для решения задачи VAD в реальных, зашумленных условиях. Наиболее перспективными для практического внедрения являются CNN+BiGRU и CNN+TDNN, что открывает возможности для использования таких моделей в мобильных, встраиваемых и облачных речевых системах.

Ключевые слова: детекторы голосовой активности, отношение сигнал/шум, рекуррентные нейронные сети, сверточные нейронные сети, машинное обучение

Благодарность. *Данное исследование финансируется Комитетом науки Министерства науки и высшего образования Республики Казахстан (Грант №AP22684173 «Разработка высокоэффективного нейросетевого метода обнаружения голосовой активности при низком уровне отношения сигнал/шум»)*

Введение. В современных системах цифровой обработки речи обнаружение голосовой активности (VAD – Voice Activity Detection) играет важную роль и используется в таких областях, как распознавание речи, телекоммуникации и умные голосовые помощники. Практически все современные аудиотехнологии требуют надежной и точной работы системы VAD. Это особенно важно при использовании таких технологий в мобильных устройствах, умных гаджетах для интернета вещей (IoT) и различных системах безопасности. Качество

работы VAD напрямую влияет на эффективность голосовых помощников, распознавание команд и работу автоматизированных сервисов. Если система VAD работает нестабильно, это может привести к ошибкам в распознавании речи, задержкам в ответах устройств или даже сбоям в работе важных сервисов, что особенно критично для безопасности и повседневного использования современных технологий.

Во многих жизненных ситуациях – например, в автобусе, на заводе или в шумном торговом центре – голос трудно распознать из-за сильного фонового шума. Поэтому задача надежного обнаружения речи при низком отношении сигнал/шум (ОСШ) остается очень важной. В подобных условиях обычные системы часто ошибаются, ведь речь почти сливается с окружающими звуками, и выделить ее становится сложно. Традиционные алгоритмы VAD быстро теряют эффективность при ОСШ ниже 0 дБ, что ограничивает возможности автоматических речевых систем в практических приложениях.

Современные тенденции развития искусственного интеллекта позволяют создавать нейросетевые VAD-системы, способные адаптироваться к шуму и эффективно работать даже в сложных акустических условиях. Однако задачи обеспечения устойчивой работы VAD при экстремально низком ОСШ, а также адаптации таких систем для языков с ограниченным объемом обучающих данных и внедрения в устройства с ограниченными вычислительными ресурсами по-прежнему остаются нерешенными. Высокая востребованность в таких исследованиях связана с необходимостью повышения надежности, точности и универсальности голосовых интерфейсов для телекоммуникаций, транспорта, индустрии безопасности, робототехники и цифровых сервисов. Поэтому научные исследования, направленные на разработку устойчивых к шуму нейросетевых систем VAD, сохраняют свою актуальность и практическую значимость для дальнейшего развития интеллектуальных аудиотехнологий.

Классические алгоритмы VAD базируются на простых статистических признаках – таких как энергия, нулевая частота пересечений (Zero Crossing Rate), спектральная энтропия, линейное предсказание (LPC) и спектральное вычитание. Они характеризуются низкой вычислительной сложностью и высокой скоростью, что удобно для работы в реальном времени. Однако при снижении ОСШ ниже 5 дБ их эффективность резко падает (Lu, 2010; Hu, 2007). Более совершенные системы, реализованные в кодеках G.729B, AMR и WebRTC, используют спектральную энергию и эвристики, оставаясь надежными при ОСШ > 10 дБ, но становятся неустойчивыми при ОСШ < 0 дБ – растет количество ложных срабатываний (FP – False Positive) и пропусков речи (FN – False Negative).

Для повышения устойчивости VAD применяются методы шумоподавления – минимизация спектральной энтропии (SMPR), спектральное вычитание (SS), вейвлет-преобразования (DWT) и их комбинации с алгоритмами

машинного обучения. Гибридные архитектуры, такие как SS+DWT+SVM, демонстрируют высокие показатели качества (PESQ, STOI) даже при –10 дБ (Iqbal, 2025; Nagaraja, 2024), однако требуют ручной настройки признаков и недостаточно эффективны при динамических шумах. Предложенный метод динамической фильтрации D-FBSS с отдельными DNN-моделями для разных типов шума улучшает точность (Aliouat, 2025), но требует больших обучающих данных и ресурсов.

С распространением глубокого обучения интерес к нейросетевым архитектурам VAD резко возрос. Они показывают лучшие результаты по точности и обобщающей способности в условиях шума (Hughes, 2013; Soni, 2021). Сверточные сети (CNN) успешно извлекают устойчивые временно-частотные признаки (Tan, 2024), рекуррентные LSTM и GRU учитывают временные зависимости (Han, 2024), а двунаправленные BiLSTM, BiGRU повышают полноту, хотя менее подходят для реального времени (Wilkinson, 2021; Suvorov, 2018). TDNN (Time Delay Neural Network), адаптированные из ASR, обеспечивают эффективную обработку временных зависимостей с низкой задержкой и высокой устойчивостью к шуму (Snyder, 2017).

Современные гибридные решения (например, DWT-CNN-MCSE) обеспечивают прирост точности распознавания при ОСШ –10 дБ (Cherukuru, 2024), извлекая пространственно-временные признаки. Однако большинство исследований ограничено фиксированными уровнями шума (–5, 0, 5 дБ), редко охватывает ОСШ ниже –10 дБ, фокусируется на англоязычных корпусах (TIMIT, LibriSpeech) и редко оценивает вычислительную сложность моделей.

Несмотря на современные достижения, задача точного обнаружения речи при очень низком ОСШ (до –20 дБ) остается нерешенной, особенно для языков с ограниченным объемом фонетических данных, например казахского. Актуальным становится использование реалистично зашумленных данных (KSC2 (Mussakhojayeva, 2022), ESC-50 (Piczak, 2015), исследование разных нейросетевых архитектур (CNN+BiGRU, CNN+GRU, CNN+LSTM, CNN+BiLSTM, CNN+TDNN), анализ метрик качества (Accuracy, Recall, Precision, F1), влияния ОСШ (–20…+30 дБ) и баланса между качеством и вычислительной сложностью моделей. Все это подтверждает актуальность разработки устойчивых и эффективных VAD-систем для работы в реальных шумных условиях.

Материалы и методы. В качестве данных использовался Kazakh Speech Corpus (KSC2), содержащий записи 30 дикторов, к которым добавлялись различные типы шумов (белый, бытовой, природный, городской, животные) из набора ESC-50. Все записи были размножены с уровнями ОСШ от –20 до +30 дБ, что обеспечило широкий спектр акустических сценариев. На этапе обработки аудиофайлы были автоматически сегментированы на отдельные слова (словоформы), из которых затем извлекались мел-частотные кепстральные коэффициенты (MFCC), нормализованные в диапазоне [0, 1].

Модели обучались и тестировались на сбалансированных по классам выборках, разделенных по дикторам в пропорции 80:20. Были реализованы пять гибридных нейросетевых архитектур (CNN+BiGRU, CNN+GRU, CNN+LSTM, CNN+BiLSTM, CNN+TDNN), различающихся типом рекуррентных и сверточных слоев. Обучение проводилось с использованием Adam-оптимизатора и функции потерь categorical crossentropy. Эффективность моделей оценивалась по метрикам Accuracy, Precision, Recall, F1-score и анализу матриц ошибок на разных уровнях ОСШ. Эксперименты выполнялись на вычислительной платформе с GPU RTX 4090 и современными инструментами Python (Librosa, TensorFlow, Scikit-learn).

Результаты. Для оценки качества детектирования речевой активности в зашумленных условиях были протестированы пять гибридных нейросетевых архитектур: CNN+BiGRU, CNN+BiLSTM, CNN+GRU, CNN+LSTM и CNN+TDNN. Все модели обучались и тестировались на едином датасете с варьируемым ОСШ от –20 до +30 дБ, что обеспечило объективное сравнение их производительности. Оценка велась по стандартным метрикам бинарной классификации: F1-score, Accuracy, Precision и Recall, что позволяет комплексно оценить работу моделей в разных акустических условиях.

На рисунке 1 представлена динамика изменения точности классификации (Accuracy) в зависимости от уровня ОСШ для всех протестированных моделей. По мере увеличения значения ОСШ наблюдается стабильный рост Accuracy для каждой архитектуры. Следовательно, уже при значениях ОСШ выше 0 дБ точность всех моделей превышает 95%. Это подтверждает способность архитектур эффективно работать в относительно чистых акустических условиях.

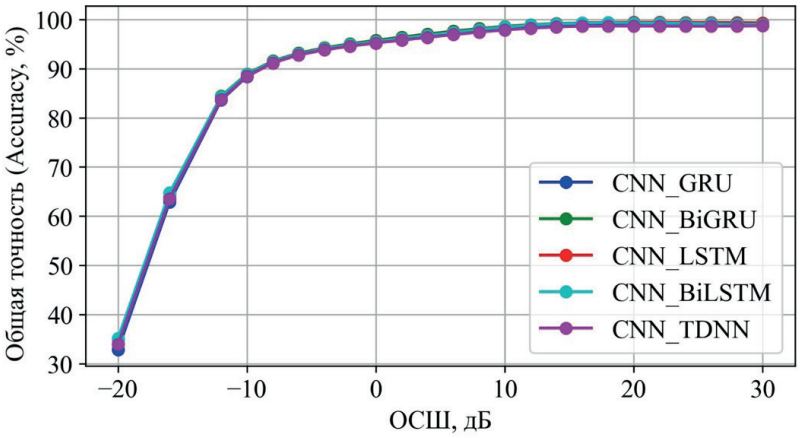


Рисунок 1 – Зависимость общей точности от уровня ОСШ для всех протестированных моделей

Однако в области отрицательных значений ОСШ (от –20 дБ до 0 дБ) проявляются характерные различия между моделями. Архитектуры CNN+BiGRU и CNN+LSTM показывают более высокие значения Accuracy на всем шумовом диапазоне, сохраняя лидерство даже при экстремальном уровне зашумления. Напротив, для CNN+TDNN и CNN+GRU характерно существенное снижение точности при ОСШ ниже –10 дБ, что ограничивает их применение в особо сложных акустических условиях.

Таким образом, сравнительный анализ на всем диапазоне ОСШ демонстрирует разную степень устойчивости протестированных моделей к шумовым искажениям, что необходимо учитывать при выборе архитектуры для практических задач в реальных условиях эксплуатации.

В таблице 1 представлены усредненные по всем значениям ОСШ результаты для каждой архитектуры. Наивысшие значения F1-score (99,6%) показали модели CNN+BiGRU, CNN+LSTM и CNN+GRU, что отражает их высокую согласованность в классификации речевых и неречевых сегментов.

Таблица 1. Сравнение метрик классификации для различных нейросетевых архитектур при варьирующемся уровне отношения С/Ш

Модель	F1, %	Precision, %	Accuracy, %	Recall, %
CNN+BiGRU	99,6	99,4	99,6	99,8
CNN+BiLSTM	99,5	99,2	99,6	99,8
CNN+LSTM	99,6	99,5	99,7	99,8
CNN+TDNN	99,3	98,7	99,3	99,9
CNN+GRU	99,6	99,3	99,6	99,9

Максимальная Accuracy (99,7%) отмечена у CNN+LSTM, демонстрируя ее стабильность по всем условиям тестирования. Модель CNN+TDNN, хотя и показала минимальное значение Accuracy (99,3%), обеспечила максимальный Recall (99,9%), что свидетельствует о высокой чувствительности к речевым событиям даже при низком ОСШ. Таким образом, различия между архитектурами проявляются в балансе между точностью и полнотой, что важно при выборе модели для практического внедрения.

Параметры обучения приведены в таблице 2. Для всех моделей использовались одинаковые условия: 10 эпох, размер батча 1024. Accuracy на обучающей и тестовой выборках стабильно достигала 96% для всех архитектур, значения Loss – в диапазоне 9,8–10,5%.

Таблица 2. Результаты обучения моделей: точность, потери и количество параметров

Модель	Количество эпох	Точность (train/test)	Потери (train/test)	Количества параметров
CNN+BiGRU	10	96%/96%	9,8%/9,9%	11 106
CNN+BiLSTM	10	96%/96%	9,8%/9,8%	13 538
CNN+GRU	10	96%/96%	10,2%/10,1%	6 050

CNN+LSTM	10	96%/96%	10,1%/10,3%	7 010
CNN+TDNN	10	96%/96%	10,4%/10,5%	5 650

Кроме того, на рисунке 2 представлены детализированные графики, иллюстрирующие процесс обучения нейросетевой модели CNN-BiGRU. На левом графике отображается динамика изменения значения функции потерь (loss – для обучающей выборки, val_loss – для валидационной), а на правом – изменение точности классификации (acc – для обучающей выборки, val_acc – для валидационной) на протяжении всех эпох обучения.

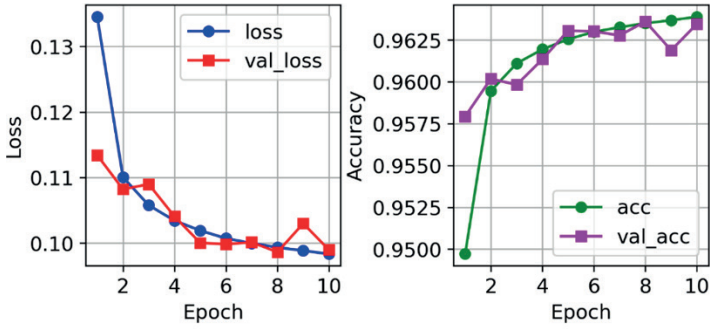


Рисунок 2 – Обучение модели CNN-BiGRU

Как видно из графиков, на первых этапах обучения наблюдается быстрое снижение значения функции потерь, что свидетельствует о быстрой адаптации модели к данным. Параллельно с этим фиксируется устойчивый рост точности классификации на обеих выборках. Уже к 6–7-й эпохе модель достигает высокой точности – более 96%, а значения функции потерь стабилизируются, что указывает на завершение фазы активного обучения и переход к режиму устойчивой работы без признаков переобучения. Отсутствие существенного расхождения между кривыми для обучающей и валидационной выборок подтверждает, что модель обладает хорошей способностью к обобщению и не склонна к переобучению на обучающих данных.

Также графики динамики обучения свидетельствуют о стабильной сходимости и отсутствии переобучения. Число параметров у моделей варьируется от 5 650 (CNN+TDNN) до 13 538 (CNN+BiLSTM), что отражает различия в архитектурной сложности и потенциальном влиянии на вычислительные затраты в реальном применении.

На рисунке 3 показано, как F1-score модели CNN+BiGRU меняется в зависимости от ОСШ. При ОСШ –20 дБ F1-score составляет 46%, что свидетельствует о резком снижении качества классификации в условиях экстремального шума.

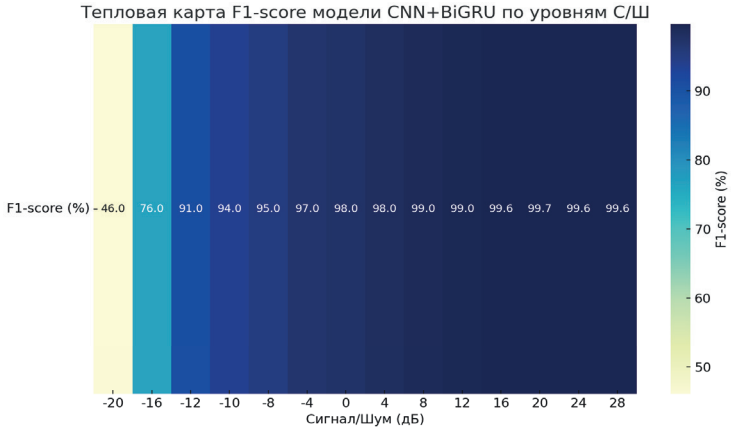


Рисунок 3 – Зависимость F1-меры от значения отношения С/Ш для модели CNN+BiGRU

Уже при –12 дБ F1-score возрастает до 91%, демонстрируя восстановление эффективности даже при отрицательных значениях ОСШ. При ОСШ выше –4 дБ F1-score стабильно превышает 98% вплоть до +18 дБ, отражая высокую устойчивость модели. В диапазоне от –18 дБ до –4 дБ наблюдается наибольший прирост F1-score, что указывает на чувствительность модели к уровню шума в наиболее сложных условиях. Такая зависимость позволяет определить диапазон рабочих режимов модели и границы, при которых обеспечивается требуемое качество классификации.

Для детального анализа характера ошибок классификации рассмотрена модель CNN+BiGRU на уровне ОСШ –6 дБ – этот уровень выбран как показательный для сравнения, поскольку он позволяет выявить особенности работы архитектуры в сложных, но не экстремальных, шумовых условиях. На рисунке 4 приведена нормализованная матрица ошибок (confusion matrix) для CNN+BiGRU, отражающая доли правильно и ошибочно классифицированных сегментов по классам «речь» и «шум».

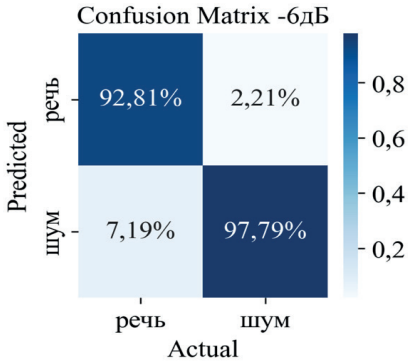


Рисунок 4 – Матрицы ошибок *CNN+BiGRU* при низком уровне ОСШ

Для данной архитектуры точность определения речевых сегментов составляет 92,81%, а шумовых – 97,79%. Доля ложноположительных результатов (False Positives) составляет 2,21%, а ложноотрицательных (False Negatives) – 7,19%. Такой анализ позволяет наглядно продемонстрировать, что даже при значительном уровне акустических помех модель сохраняет приемлемый баланс между точностью и полнотой, корректно различая большинство речевых и неречевых фрагментов.

Выбор именно CNN+BiGRU обусловлен тем, что эта архитектура сочетает высокие показатели качества с устойчивостью к шуму, и может служить наглядным примером типичного распределения ошибок для современных нейросетевых VAD-систем при сложных акустических сценариях.

Обсуждение. Полученные результаты показывают, что все протестированные нейросетевые архитектуры обеспечивают высокое качество обнаружения речевой активности при широком диапазоне уровней ОСШ. Архитектуры CNN+BiGRU, CNN+LSTM и CNN+GRU достигли максимального F1-score (99,6%), что свидетельствует о надежной классификации речи даже в сложных условиях. Модель CNN+LSTM также показала наивысшую среднюю Accuracy (99,7%), а CNN+GRU и CNN+TDNN отличились максимальным Recall (99,9%), что важно для задач, где критична полнота обнаружения.

CNN+TDNN при минимальном числе параметров (5 650) демонстрирует сравнимую точность с более сложными моделями, что делает ее оптимальной для встроенных систем с ограниченными вычислительными ресурсами.

Рост F1-score и Accuracy моделей с увеличением ОСШ подтверждает их устойчивость к шуму. Так, у CNN+BiGRU при ОСШ –12 дБ F1-score превышает 90%, а при ОСШ ≥ 0 дБ стабильно превышает 98%. Матрица ошибок для CNN+BiGRU при ОСШ –6 дБ подтверждает сбалансированность показателей точности и полноты даже при сильных помехах.

Сравнение архитектур также показывает, что более сложные модели требуют больше времени и ресурсов на обучение, но не всегда значительно превосходят по качеству более компактные решения.

Таким образом, CNN+BiGRU и CNN+TDNN можно рассматривать как наиболее сбалансированные по качеству и вычислительной эффективности варианты для практического применения в шумных акустических условиях. В дальнейшем требуется исследование работы моделей на реальных аудиопотоках и с динамическими шумами.

Заключение. В данной работе выполнено экспериментальное сравнение пяти гибридных нейросетевых архитектур, таких как CNN+BiGRU, CNN+GRU, CNN+LSTM, CNN+BiLSTM и CNN+TDNN для детектирования речевой активности VAD в условиях широкого диапазона шумовых помех. Для анализа использовался Kazakh Speech Corpus (KSC2), дополненный реалистичными шумами из ESC-50 при ОСШ от –20 до +30 дБ, а в качестве признаков применялись MFCC.

Полученные результаты показали, что все протестированные модели обеспечивают высокое качество классификации на всем диапазоне условий (Accuracy и F1-score свыше 99,3%). Лучший баланс между точностью, полнотой и устойчивостью продемонстрировала архитектура CNN+BiGRU, а CNN+TDNN показала высокие результаты при минимальной вычислительной сложности, что особенно важно для встраиваемых решений.

Анализ матриц ошибок подтвердил, что предложенные подходы позволяют эффективно разделять речь и шум даже при низких значениях ОСШ. Наиболее перспективными для практического внедрения являются архитектуры CNN+BiGRU и CNN+TDNN, пригодные для использования в голосовых ассистентах, телекоммуникационных и мониторинговых системах в сложных акустических условиях.

В перспективе предполагается интеграция разработанных VAD-моделей с другими системами обработки аудиосигналов, такими как автоматическое распознавание речи и системы интеллектуального управления, для создания комплексных решений.

References

Aliouat M., Djendi D. (2025) An Efficient Speech Enhancement Algorithm Using Dynamic Filter Banks Based on Noise Type, International Journal of Speech Technology. — Vol. 28. — P. 65–79. https://doi.org/10.1007/s10772-025-09999-2 (in English)

Cherukuru S., Mustafa G. (2024) A DWT-CNN-MCSE-Based Speech Enhancement System for Improved Speech Recognition in Noisy Environments, Journal of Ambient Intelligence and Humanized Computing. — Vol. 15(4). — P. 1235–1249. https://doi.org/10.1007/s12652-023-04513-8 (in English)

Han I., Om C.-N., Kim U.-I. (2024) A Gated Recurrent Unit Based Robust Voice Activity Detector, Multimedia Tools and Applications. — Vol. 83(14). — P. 41939–41949. https://doi.org/10.1007/s11042-023-17123-w (in English)

Hu Y., Loizou P.C. (2007) A Comparative Intelligibility Study of Single-Microphone Noise Reduction Algorithms, The Journal of the Acoustical Society of America. — Vol. 122(3). — P. 1777–1786. https://doi.org/10.1121/1.2757668 (in English)

Hughes T., Mierle K. (2013) Recurrent Neural Networks for Voice Activity Detection, Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). — P. 7378–7382. https://doi.org/10.1109/ICASSP.2013.6639132 (in English)

Iqbal M.M., Arain S.S., Talpur A.A., Shaikh A.A. (2025) An Enhanced Speech Enhancement Model Based on SS and DWT for Improved Noise Suppression, Applied Acoustics. — Vol. 214. — Art. 109430. https://doi.org/10.1016/j.apacoust.2023.109430 (in English)

Lu Y., Loizou P.C. (2010) A Geometric Approach to Spectral Subtraction, Speech Communication. — Vol. 52(7–8). — P. 614–627. https://doi.org/10.1016/j.specom.2010.01.005 (in English)

Mussakhojayeva S., Khassanov Y., Varol H.A. (2022) KSC2: An Industrial-Scale Open-Source Kazakh Speech Corpus, Proceedings of the 23rd INTERSPEECH Conference. — P. 1367–1371. (in English)

Nagaraja N., Kumuda R., Bhat M. (2024) Voice Activity Detection Using MMSE-SPZC in Real-World Noisy Environments, Computer Speech & Language. — Vol. 80. — Art. 101540. https://doi.org/10.1016/j.csl.2023.101540 (in English)

Piczak K.J. (2015) ESC: Dataset for Environmental Sound Classification, Proceedings of the 23rd ACM International Conference on Multimedia. — P. 1015–1018. https://doi.org/10.1145/2733373.2806390 (in English)

Snyder D., Garcia-Romero D., Povey D., Khudanpur S. (2018) Deep Neural Network Embeddings for Text-Independent Speaker Verification, Interspeech 2017. — P. 999–1003. https://doi.org/10.21437/Interspeech.2017-620 (in English)

Soni M.H., Shah R.R., Patil H.A. (2021) Time–Frequency Attention for Noise-Robust Speaker Verification Using CNN and TDNN, IEEE/ACM Transactions on Audio, Speech, and Language Processing. — Vol. 29. — P. 879–892. https://doi.org/10.1109/TASLP.2021.3061045 (in English)

Suvorov D., Zhukov R., Tsetserukov D., Zenkevich S. (2018) Audiovisual Voice Activity Detector Based on Deep Convolutional Neural Network and Generalized Cross-Correlation, Mekhatronika Avtomatizatsiya Upravlenie. — P. 53–57. https://doi.org/10.17587/mau.19.53-57 (in English)

Tan Y.W., Ding X.F. (2024) Heterogeneous Convolutional Recurrent Neural Network with Attention Mechanism and Feature Aggregation for Voice Activity Detection, APSIPA Transactions on Signal and Information Processing. — Vol. 13(1). — Art. e6. https://doi.org/10.1561/116.00000158 (in English)

Wilkinson N., Niesler T. (2021) A Hybrid CNN–BiLSTM Voice Activity Detector, ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing – Proceedings. — 2021-June. — P. 6803–6807. https://doi.org/10.1109/ICASSP39728.2021.9415081 (in English)

ACADEMIC SCIENTIFIC JOURNAL OF COMPUTER SCIENCE
ISSN 1991-346X
Volume 4. Namber 356 (2025). 216–233

https://doi.org/10.32014/2025.2518-1726.393

IRSTI 52.35.01
UDC 665.61.7: 519.873

© B. Orazbayev[1], A. Zhumadillayeva[1*], K. Orazbayeva[2], R. Yessirkessinov[3], Zh. Tuleuov[4], 2025.

[1] L.N. Gumilyov Eurasian National University, Astana, Kazakhstan;
[2] Esil university, Astana, Kazakhstan;
[3] Kazakh-British Technical University, Almaty, Kazakhstan;
[4]Atyrau refinery, Atyrau, Kazakhstan.
E-mail: zhumadillayeva_ak@enu.kz

DEVELOPMENT OF MODELS OF SULFUR PRODUCTION PROCESSES BASED ON ARTIFICIAL NEURAL NETWORKS AND SIMULATION

Orazbayev Batyr — doctor of technical sciences, professor, L.N. Gumilyov Eurasian National University, Astana, Kazakhstan,
E-mail: batyr_o@mail.ru, https://orcid.org/0000-0003-2109-6999;
Zhumadillayeva Ainur — candidate of Technical Sciences, Associate Professor, L.N. Gumilyov Eurasian National University, Astana, Kazakhstan,
E-mail: zhumadillayeva_ak@enu.kz, https://orcid.org/0000-0003-1042-0415;
Orazbayeva Kulman — doctor of technical sciences, professor, Esil University, Astana, Kazakhstan,
E-mail: kulman_o@mail.ru, https://orcid.org/0000-0002-1741-7553;
Yessirkessinov Ramazan. **—** student of the Kazakh-British Technical University, Almaty, Kazakhstan,
E-mail: r_yessirkessinov@kbtu.kz, https://orcid.org/0009-0003-1912-1013;
Tuleuov Zhasulan — Senior Operations Coordinator, Coke and Sulfur Production, Atyrau Oil Refinery, Atyrau, Kazakhstan,
E-mail: zh.tuleuov@anpz.kz, https://orcid.org/0000-0003-2353-2416

Abstract. Today, in Kazakhstan, as in other oil-producing countries, the share of high-sulfur crude oil production is increasing. In this regard, enhancing the efficiency of sulfur production processes from harmful sulfur-containing gases released during oil refining, which produce a useful and necessary product for many industries, is a highly relevant scientific and practical task. This study aims to improve the efficiency of sulfur production processes by modeling them using one of the artificial intelligence methods -artificial neural networks. To create an adequate mathematical model of the sulfur production process at the Atyrau Oil Refinery (AOR), it is proposed to use neural network technology. In this work, a hierarchical multilayer neural network model was developed for effectively modeling the sulfur production plant under conditions of uncertainty and limited initial data. The Python programming language and its necessary tools were chosen

and justified for implementing the neural model. An iterative gradient algorithm based on the backpropagation method was proposed as a learning method for the constructed multilayer neural network, and its main stages were described. During the study, the proposed iterative gradient algorithm was implemented in Python using data obtained through fuzzy set theory methods. Based on the created program, graphs of output values and error values depending on the number of epochs were constructed on a computer, and the results were interpreted. It was found that the results obtained using artificial neural networks and fuzzy set methods are consistent with each other. The training results of the neural network model showed minimal errors, and the amount of produced sulfur was as close as possible to the target value.

Keywords: artificial neural network, activation function, Python programming language, model, Claus process, backpropagation algorithm, iterative gradient algorithm

Financing. *This research was supported the Science Committee of the Ministry of Science and Higher Education of the Republic of Kazakhstan (Grant No. fund this research AP23490206 Development of a decision support system for managing the sulfur production process).*

© Б. Оразбаев[1], А. Жумадиллаева[1*], К. Оразбаева[2], Р. Есиркесинов[3], Ж. Тулеуов[4], 2025.

[1]Л.Н. Гумилев атындағы Еуразия ұлттық университеті, Астана, Қазақстан;
[2]Esil university, Астана, Қазақстан;
[3]Қазақстан-Британ техникалық университеті, Алматы, Қазақстан;
[4]Атырау мұнай өңдеу зауыты Атырау, Қазақстан.
E-mail: zhumadillayeva_ak@enu.kz

КҮКІРТ ӨНДІРУ ПРОЦЕССТЕРІНІҢ МОДЕЛЬДЕРІН ЖАСАНДЫ НЕЙРОНДЫҚ ЖЕЛІЛЕР НЕГІЗІНДЕ ӘЗІРЛЕУ ЖӘНЕ МОДЕЛЬДЕУ

Оразбаев Батыр — техника ғылымдарының докторы, профессор, Л.Н. Гумилев атындағы Еуразия ұлттық университеті, Астана, Қазақстан,
E-mail: batyr_o@mail.ru, https://orcid.org/0000-0003-2109-6999;
Жумадиллаева Айнур — техника ғылымдарының кандидаты, қауымдастырылған профессор, Л.Н. Гумилев атындағы Еуразия ұлттық университеті, Астана, Қазақстан,
E-mail: zhumadillayeva_ak@enu.kz, https://orcid.org/0000-0003-1042-0415;
Оразбаева Кулман — техника ғылымдарының докторы, профессор, Esil university, Астана, Қазақстан,
E-mail: kulman_o@mail.ru, https://orcid.org/0000-0002-1741-7553;
Есиркесинов Рамазан — Қазақ-Британ техникалық университеті студенті, Алматы, Қазақстан,
E-mail: r_yessirkessinov@kbtu.kz, https://orcid.org/0009-0003-1912-1013;
Тулеуов Жасулан — Кокс және күкірт өндірісі аға операциондық үйлестірушісі, Атырау мұнай өңдеу зауыты, Атырау, Қазақстан,
E-mail: zh.tuleuov@anpz.kz, https://orcid.org/0000-0003-2353-2416.

Аннотация. Бүгінгі таңда Қазақстанда, басқа да мұнайлы елдердегі сияқты, өндірілетін күкіртті мұнайдың үлесі өсу тенденциясына ие. Осыған байланысты көптеген салалар үшін пайдалы әрі қажетті өнім болып табылатын мұнай өңдеу процестері кезінде бөлінетін зиянды күкірті газдардан күкірт өндіру процестерінің тиімділігін арттыру өте өзекті ғылыми-практикалық мәселе болып табылады. Сол себептен бұл жұмыс күкірт өндіру процесінің тиімділігін жасанды интеллект тәсілдерінің бірі -нейрондық желі арқылы модельдеу арқылы арттыруға бағытталған. Зерттеуде Атырау мұнай өңдеу зауытының күкірт өндіру қондырғысының мақсатты өнімі -күкірт өндіру процесін ақпарат тапшылығы мен анықсыздығы жағдайында математикалық адекватты моделі оқытылып, әзірленген. Зерттеуде күкірт өндіру қондырғысынын ақпарат тапшылығы мен белгісіздігі жағдайында тиімді модельдеу үшін иерархиялық көпқабатты нейрондық желі құрылымы құрылған және нейрондық желіні программалық жүзеге асыру үшін қолданылатын программалау тілі ретінде Python және оның қажетті құралдары негізделіп, алынған. Құрылған көпқабатты нейрондық желіні оқыту тәсілі ретінде қателердің кері таралуы алгоритмі негізінде итеративті градиенттік алгоритмі ұсынылып, оның негізгі кезеңдері сипатталған. Зерттеу барысында айқын емес жиындар тәсілдерін қолдану арқылы алынған мәліметтерді пайдаланып, Python тілінде жұмыста әзірленген итеративті градиент алгоритмі программалық жүзеге асырылды. Құрылған программа негізінде компьютерде дәуір санына байланысты шығыс мәндерінің және қателер мәндерінің графиктері тұрғызылып, талқыланды. Жасанды нейрондық желілер мен айқын емес жиындар тәсілдері қолдану арқылы арқылы алынған нәтижелері бір біріне сәйкес келетіндігі негізделді. Нейрондық технология негізінде құрылған модельді оқыту нәтижесінде қателіктер ең аз және өндірілген күкірт көлемі оның мақсатты мәніне мүмкіндігінше жақын болатыны анықталды.

Түйін сөздер: жасанды нейронды желі, белсендіру функциясы, Python программалау тілі, модель, Клаус процесі, қателердің кері таралу алгоритмі, итеративті градиент алгоритмі

© Б. Оразбаев[1], А. Жумадиллаева[1*], К. Оразбаева[2], Р. Есиркесинов[3], Ж. Тулеуов[4], 2025.

[1]Л.Н. Гумилев атындағы Еуразия ұлттық университеті, Астана, Қазақстан;
[2]Esil university, Астана, Қазақстан;
[3]Қазақстан-Британ техникалық университеті, Алматы, Қазақстан;
[4]Атырау мұнай өңдеу зауыты Атырау, Қазақстан.
E-mail: zhumadillayeva_ak@enu.kz

РАЗРАБОТКА МОДЕЛЕЙ ПРОЦЕССОВ ПРОИЗВОДСТВА СЕРЫ НА ОСНОВЕ ИСКУССТВЕННЫХ НЕЙРОННЫХ СЕТЕЙ И МОДЕЛИРОВАНИЕ

Оразбаев Батыр — доктор технических наук, профессор, Евразийский национальный университет имени Л.Н. Гумилева, Астана, Казахстан,
E-mail: batyr_o@mail.ru, https://orcid.org/0000-0003-2109-6999;
Жумадиллаева Айнур — кандидат технических наук, ассоциированный профессор, Евразийский национальный университет имени Л.Н. Гумилева, Астана, Казахстан,
E-mail: zhumadillayeva_ak@enu.kz, https://orcid.org/0000-0003-1042-0415;
Оразбаева К.Н. — доктор технических наук, профессор, Университет «Есиль», Астана, Казахстан,
E-mail: kulman_o@mail.ru, https://orcid.org/0000-0002-1741-7553;
Есиркесинов Рамазан — студент Казахстанско-Британского технического университета, Алматы, Казахстан,
E-mail: r_yessirkessinov@kbtu.kz, https://orcid.org/0009-0003-1912-1013;
Тулеуов Жасулан — старший координатор производства кокса и серы, Атырауский нефтеперерабатывающий завод, Атырау, Казахстан,
E-mail: zh.tuleuov@anpz.kz, https://orcid.org/0000-0003-2353-2416

Аннотация. Сегодня в Казахстане, как и в других нефтедобывающих странах, доля добываемой сернистой нефти имеет тенденцию к увеличению. В связи с этим повышение эффективности процессов производства серы из вредных серосодержащих газов, выделяющихся при нефтепереработке, становится актуальной научно-практической задачей. Сера является полезным и необходимым продуктом для многих отраслей промышленности. Данная работа направлена на повышение эффективности процесса производства серы посредством его моделирования с использованием одного из методов искусственного интеллекта — искусственных нейронных сетей. Для создания адекватной математической модели процесса производства серы на Атырауском нефтеперерабатывающем заводе (НПЗ) предлагается применять нейросетевые технологии. В работе создана иерархическая многослойная нейросетевая модель для эффективного моделирования установки по производству серы в условиях неопределенности и ограниченности исходной информации. Для программной реализации нейронной модели обоснованно выбран язык программирования Python и его инструменты. В качестве метода обучения построенной многослойной нейронной сети

предложен итерационный градиентный алгоритм, основанный на методе обратного распространения ошибки, с описанием его основных этапов. В ходе исследования разработанный итерационный градиентный алгоритм был реализован на языке Python с использованием данных, полученных с помощью методов теории нечетких множеств. На основе созданной программы были построены графики зависимости выходных значений и значений ошибок от количества эпох, а также дана интерпретация полученных результатов. Было установлено, что результаты, полученные с использованием методов искусственных нейронных сетей и нечетких множеств, согласуются друг с другом. Результаты обучения модели на основе нейронной технологии показали минимальные ошибки, а объем производимой серы оказался максимально близким к целевому значению.

Ключевые слова: искусственная нейронная сеть, функция активации, язык программирования Python, модель, процесс Клауса, алгоритм обратного распространения ошибки, алгоритм итерационного градиента

Кіріспе. Қазіргі уақытта барлық мұнайлы елдердегідей Қазақстанда да мұнай өңдеу тереңдігі мен мұнай өңдеу өнімдері ассортименттерін арттыру мұнай өңдеудің басым бағыты болып саналады. Ал мұнай өндіруде күкіртті мұнай үлесінің артуы, мұнай өңдеу өнімдеріне экологиялық талаптардың қатаңдау тенденциясы мұнай өңдеу процесстерінде бөлінетін күкіртті сутек газдарын өңдеп, әлемдік нарықта жоғары сұранысқа ие күкірт алу аса өзекті ғылым-практикалық мәселеге жатады (Kuznetsov et al., 2021; Edmonson et al., 2020).

Мұнай өңдеу зауыттарында күкірт сутектен тауарлық күкіртті алудың ең көп тараған технологиялық процесі - Клаус процесі жүретін күкірт өндіру қондырғысында өндіреді (КӨҚ) (Golubeva et al., 2020; Gazimzyanov et al., 2020). Аталған қондырғыда күкірт өндіру процесін зерттеп, Клаус процесін оптималды жүргізіп, тиімді басқару үшін КӨҚ-да күкірт алу процесстерінің адекватты математикалық модельдерін құру қажет. Алайда КӨҚ-ның термореакторында, Клаус және салқын абсорбциялау (СА) реакторларында катализатор қатысуымен өтетін күрделі процесстердің адекватты математикалық модельдерін дәстүрлі тәсілдермен құруға қажетті деректер мен ақпараттар өндірістік жағдайда жетіспейді. Ал оларды жинау теориялық тұрғыдан мүмкін болғанымен, экономикалық тұрғыдан практикада тиімсіз, өте көп шығындарды талап етеді. Сондықтан күкірт өндіру процесстерінің адекватты, тиімді математикалық модельдерін құру үшін арнаулы, жасанды интеллект тәсілдерін, мысалы жасанды нейрондық желілер мен айқын емес емес жиындар тәсілдерін қолдану қажеттігі туындайды.

Күкірт өндіру технологиясы мен процесстерін зерттеу, математикалық модельдеу және оптимизациялау мәселелеріне арналған негізгі зерттеулерді

шолу нәтижелерін келтірейік. Каталитикалық реакциялар мен процестерді зерттегенде, зерттелетін жүйе күйлерінің өзгерістерін сипаттауға мүмкіндік беретін математикалық модельдеу тәсілдері жиі қолданылады (Hashemi et al., 2019). Дегенмен, кейбір процестер үшін химиялық технологиялық процестің барлық ерекшеліктерін ескеретін толық математикалық модельдерін дәстүрлі тәсілдермен құру өте көп еңбекті қажет ететін жұмыс болып табылады. Бұл жағдайда модельдерді құру кезінде бірқатар жеңілдетулер мен шектеулер енгізіледі, бұл өз кезегінде сандық есептеулерге қателер енгізеді және алынған нәтижелерді айтарлықтай бұрмалауы мүмкін.

(Naumova et al., 2023; Kadyrov, 2023) жұмыстарында күкірт өндіру технологиялары талданып, олардың ерекшеліктері зерттелген. Карпунин, Труфанов, Nagamalleswara және басқалары өздерінің (Karpunin, 2023; Garmroodi et al., 2019) жұмыстарында күкірт өндіру процесстерін модельдеу және оптимизациялау мәселелерін зерттеген. Бұл зерттеулерде күкірт өндіру нысандары мен процесстерін модельдеу және оптимизациялау үшін дәстүрлі тәсілдер қолданылған. Алайда практикада басқа да өндірістік нысандар сияқты күкірт өндіру қондырғылары да дәстүрлі тәсілдермен модедьдеу және оптимизациялау қажетті ақпараттардың тапшылығы, олардың айқынсыздығымен жиі сипатталады. Яғни бұл жағдайда күкірт өндіру нысандары мен процесстерін модельдеу практикалық тұрғыдан өте тиімсіз немесе мүмкін емес болады. Сондықтан мұндай ақпарат тапшылығы, айқынсыздығымен сипатталатын өндірістік нысандардың модельдерін құрып, оптимизациялау үшін анықсыздық жағдайында да тиімді болып табылатын жасанды интеллект тәсілдерін қолдану қажет болады.

Айқын емес жиындар тәсілдері негізінде жұмыстың зерттеу нысаны - Атырау мұнай өңдеу зауытының (МӨЗ) КӨҚ-ның күкірт алу процесстері жүретін негізгі агрегаттарының, ақпарат тапшылығы мен айқынсыздығы жағдайында, тиімді модельдерін құру тәсілі авторлардың жарияланған (Orazbayev et al., 2024; Orazbaev et al., 2019) жұмыстарында сипатталған.

Жасанды нейрондық желілер (ЖНЖ) моделі келесі өңдеу үшін қолданылатын өзара байланысты және жинақталатын эксперименттік деректер мен білімдер элементтерінің (жасанды нейрондар) желісі болып табылады (Xiaolong, 2020). Белгілі бір модель туралы білімді жинақтау үшін нейрондар арасындағы синаптикалық салмақтар деп аталатын байланыстар пайдаланылады. Желі кіріс ақпаратты өңдейді және уақыт өте келе оның күйін өзгерту процесінде шығыс сигналдарының жиынтығын құрайды (Aksenov et al., 2022).

Нейрондық желілерді оқытудың белгілі алгоритмі бақылаудағы оқыту әдістеріне жататын қателердің кері таралу алгоритмі (Tijanić et al., 2020), ол машиналық оқытудың ең танымал алгоритмдерінің бірі. Бұл алгоритм әдісі оның жасырын қабаттарындағы нейрондардың салмақтары үшін түзету мәндерін есептеу үшін нейрондық желінің шығыс қателігін пайдалануға

негізделген. Қателердің кері таралу алгоритмі қалыпты жұмыс кезінде сигналдардың тікелей таралуына қарама-қарсы бағытта, желі шығыстарынан оның кірістеріне қателік сигналдарының таралуын қарастырады. Бұл алгоритмнің артықшылықтарына енгізудің қарапайымдылығы және ауытқулар мен деректердің ауытқуларына төзімділігі жатады. Кері таралу алгоритмін пайдалану кезінде нейрондық желіні оқытудың мақсаты нейрондық желінің салмақтарын белгілі бір кірістер жинағын қолданған кезде нейрондық шығыстардың қажетті жиынтығын (шығыс нейрондары) алуға болатындай етіп реттеу.

Химиялық технологиялық процестерді модельдеудегі ЖНЖ артықшылығы күрделі математикалық тәуелділіктермен сипатталатын процестердің икемді модельдерін құру мүмкіндігімен анықталады. Сондай-ақ нейрондық желілер болжау айнымалылары (кірістері) мен болжанатын айнымалылар (шығыстар) арасында байланыс бар жағдайларда, тіпті бұл байланыс өте күрделі сипатта болса да, болжау және топтар арасындағы корреляция немесе айырмашылықтарды шартты түрде көрсету қиын болғанда да, басқару мәселелерін шешудің ең тиімді құралы болып табылатынын атап өту қажет. Жасанды нейрондық желі моделі негізінде күрделі сипатталатын өндірістік нысандарды оптимизациялау, тиімді басқару есебін шешу алгоритмдерін құру өзекті ғылыми-техникалық мәселе болып табылады. Сол себептен бұл жұмыс кері оқыту алгоритмін осы жұмыстың зерттеу объектісі КӨҚ негізгі агрегаттары модельдерін құрып, оқытуға қолдану мүмкіндігін қарастырылады.

Бұл ұсынылған жұмыс нейрондық желілер негізінде зерттеу нысаны реакторлары модельдерін құрып, олардың авторлардың (Orazbaev et al., 2019) зерттеуінде айқын емес жиындар теориясы негізінде құрылған КӨҚ негізгі агрегаттарының модельдерімен салыстыруға арналған.

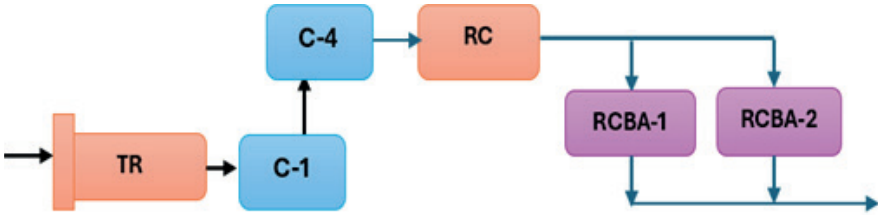
Зерттеу мақсаты, материалдары мен тәсілдері. Зерттеу мақсаты - күкірт өндіру қондырғысында жүретін күкірт өндіру процесстерін жасанды нейрондық желілер негізінде тиімді модельдеу тәсілін әзірлеу және зерттеу нысанының негізгі агрегаттары жұмыс режимдерін компьютерлік модельдеу.

Ұсынылған жұмыста зерттеу материалдары ретінде КӨҚ-ның күкірт өндіру секциясының өзара байланысқан келесі негізгі агрегаттары:

- термореактор (TR);
-конденсаторлар (C-1, C-2);
- Клаус реакторы (RC);
-салқын абсорбциялау (RCBA-1, RCBA-2) реакторлары

жұмыс режимдерін, күйлерін сипаттайтын эксперименталдық-статистикалық деректер мен өндірілген күкірттің сапасын сипаттайтын лабораториялық, эксперттік ақпараттар қолданылады.

Зерттеу нысанының, яғни КӨҚ-ның күкірт өндіру секциясының аталған негізгі агрегаттары келтірілген схемасы 1-суретте келтірілген.



TR-термореактор; С-1, С-2 -конденсаторлар; RC-Клаус реакторы;
RCBA-1, RCBA-2 -салқын абсорбциялау реакторлары
Сурет 1 – КӨҚ-ның күкірт өндіру секциясының өзара байланысқан негізгі агрегаттары

Келтірілген 1-суретте көрсетілген негізгі КӨҚ сұйық күкірт өндіру процестерін жүзеге асады. Күкірт өндіру секциясыны қышқыл газ термиялық түрлендіру процесінен TR термореакторында аралық өнім алынады. Термореактор шығысынан аралық өнім С-1, С-2 конденсаторлары арқылы каталитикалық түрлендіру процесі жүретін RC Клаус реакторына түседі. Клаус реакторы шығысынан күкірт параллель қосылған RCBA-1 және RCBA-2 реакторларына беріледі. Суық қабаттағы адсорбция процесі өтетін RCBA-1, RCBA-2 реакторларында өндірілген сапалы сұйық күкірт шұңқырына және одан күкіртті кристалдандыру және түйіршікті күкіртті алу үшін күкіртті кристалдау қондырғысына жеткізіледі.

Зерттеу материалдары ретінде бұл жұмыста сондай-ақ белгілі пассивті және белсенді эксперименттер негізінде жинақталып, математикалық статистиканың (Corazza et al., 2022; Shao et al., 2020) тәсілдерімен өңделген Атырау мұнай өңдеу зауытының (МӨЗ) КӨҚ күкірт өндіру секциясының сипатталған негізгі агрегаттарының жұмыс режимдерін параметрлері жайлы мәндері қолданылады. Зерттеуде пассивті эксперименттер ретінде КӨҚ жұмысын басқаратын тәжірибелі операторлары толтыратын режим парақтарындағы негізгі параметрлердің мәндері және мамандар қатысумен зертханада анықталған ақпараттар пайдаланылады. Сонымен қатар зерттеу материалдары ретінде Атырау МӨЗ-ның КӨҚ технологиялық регламентінің алынған ақпараттар (Tuleuov, 2023) және жасанды нейронды желілер тәсілдері мен оларды оқыту алгоритмдері (Xiaolong, 2020; Aksenov et al., 2022; Tijanić et al., 2020; Nilsson, 2019) қолданылады.

Күкірт өндіру секциясының келтірілген негізгі агрегаттарының статистикалық модельдерін құруда эксперименттік деректерді статистикалық өңдеу тәсілдері MathCad программалар пакеті негізінде пайдаланылды. Зерттеу материалдары ретінде қолданылатын күкірт өндіру секциясының негізгі агрегаттарының өлшенетін кіріс, режимдік және шығыс параметрлері мәндері өлшеу аспаптары көмегімен алынады. Ал айқынсыздықпен сипатталатын күкірттің сапа көрсеткіштері (күкірт құрамындағы күкірт, су, күл, органикалық заттар үлесі) қолданыстағы стандарт талаптарындағы айқын емес нұсқаулар мен маман-эксперттердің табиғи тілінде сипатталады.

Зерттеу нәтижелері. Зерттеу нысанының анықсыздыққа байланысты дәстүрлі тәсілдер мен құрылуы күрделі немесе мүмкін болмағанда тиімді болатын нейрондық желі моделін жобалау немесе таңдау үшін

- нейрондардың қабаттары арасындағы байланыстардың құрылымы;

– нейрондардың түрі (белсендіру функциясы тұрғысынан);

-нейрондық желіні оқыту кезінде салмақ коэффициенттерін анықтау ережелері анықталды.

Оқыту алгоритмін пайдалану үшін деректерді нормалдау (қалыпқа келтіруді) жүзеге асырылды, яғни оқыту, тексеру деректерін бір сандық шкалаға сәйкестендірілді. Бұл жұмыста құрылатын нейрондық модельді оқыту үшін жинақталған деректер мәндерін олардың ең үлкен мәніне бөлу арқылы нормалданған. Нәтижесінде барлық деректер 0-ден 1-ге дейінгі мәндерге түрлендірілді.

Қолданылатын программалау ортасы таңдалды. Python программалау тілінде программа кодын енгізу үшін Anaconda 3 программалық ортасы алынды және келесі көмекші құралдар қолданылды:

1) Matplotlib – Python-да ғылыми графиктерді құруға арналған негізгі кітапхана: *https://matplotlib.org/*

2) NumPy – Python тіліндегі ғылыми есептеулерге арналған негізгі пакеттердің бірі. Ол көп өлшемді массивтермен жұмыс істеу функционалдығы, жоғары деңгейлі математикалық функциялардан турады: *https://numpy.org/*

3) Pandas – деректерді өңдеу мен талдауға арналған Python кітапханасы: *https://pandas.pydata.org/*

4) Scikit-learn – машиналық оқытуға арналған ең танымал Python кітапханасы: https://scikit-learn.org/

5) SciPy – Python тіліндегі ғылыми есептеулерге арналған функциялар жиынтығы: https://www.scipy.org/

6) TensorFlow – нейрондық желіні құру және оқыту мәселелерін шешу үшін машиналық оқытуға арналған ашық программалық кітапхана: https://www.tensorflow.org/

7) Keras – Python-дағы ашық нейрондық желі кітапханасы TensorFlow қосымшасы болып табылады: https://keras.io

Осылардың негізінде күкірт өндіру процесінде өндірілетін күкірт көлеміне кіріс және жұмыс параметрлерінің әсерін зерттеу кезінде бұл жұмыста жасанды нейрондық желілерге негізделген модель құрылып, пайдаланылды (Shaimukhametov et al., 2019). ЖНЖ оқыту үшін қателерді кері тарату алгоритмі және итеративті градиентті оқыту алгоритмі қолданылады, ол тізбекті қосылымдары бар көпқабатты нейрондық желілердің қажетті шығыстарынан ағымдағы орташа квадраттық ауытқуын минимизациялау үшін қолданылады (Protalinsky, et al. 2021). Бұл көпқабатты перцептронның жұмыс қатесін азайту және қажетті нәтижені алу үшін қолданылатын итеративті градиент алгоритмі болып табылады.

Қателердің кері таралуы желілерінде кірістер $\sigma_1, \sigma_2, \dots, \sigma_n$ сырттан немесе алдыңғы қабаттан келеді. Әр $\sigma_1, \sigma_2, \dots, \sigma_n$ кіріс өздерінің салмағына көбейтіліп, олардың нәтижелері қосылады:

$$NET = \sigma_1 w_1 + \sigma_2 w_2 + \dots + \sigma_n w_n, \quad (1)$$

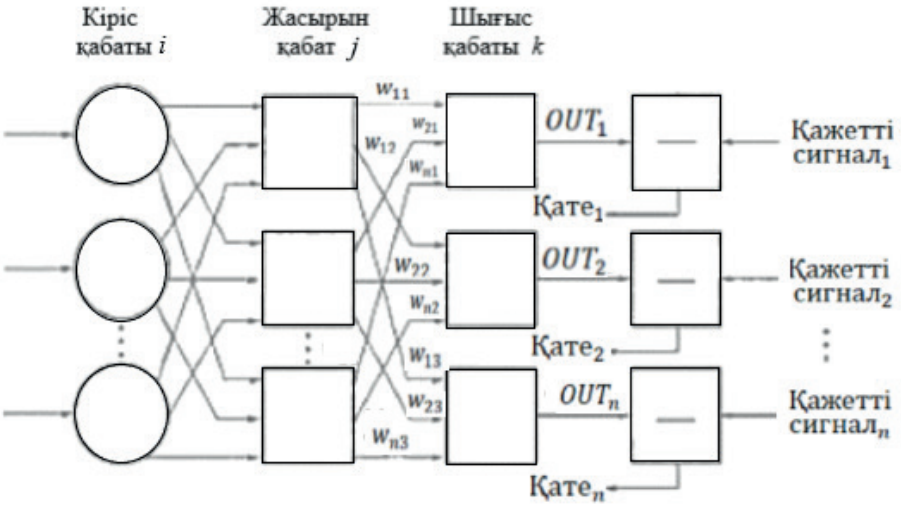
мұнда NET – желідегі әрбір нейрон үшін есептелетін сома; $\sigma_1, \sigma_2, \dots, \sigma_n$ – желідегі нейрондардың кірістері; $w_1, w_2, \dots, w_n$ – желідегі әр нейронның өзгермелі салмақ коэффициенті; F – белсендіру функциясы; OUT – нейрондық желінің шығыс сигналы.

Кері таралу алгоритмдері үшін келесі белсендіру функциясы қолданылған:

$$OUT = \frac{1}{1 + e^{-NET}}. \quad (2)$$

(2) формула сигмоидальды белсендіру функциясын сипаттайды. Ол OUT мәні нөл мен бір арасында болатындай ауқымды азайтады. Бұл функция барлық жерде дифференциалдануы керек болғандықтан жұмыста қолданылған қателерді кері тарату алгоритмі үшін сигмоидальды белсендіру функциясы ең қолайлы болып саналады.

2-суреттте келтірілген бір-бірімен байланысқан нейрондар бірнеше қабаттарға біріктірілген иерархиялық нейрондық желі, көпқабатты персептрон құрылымын қарастырайық.



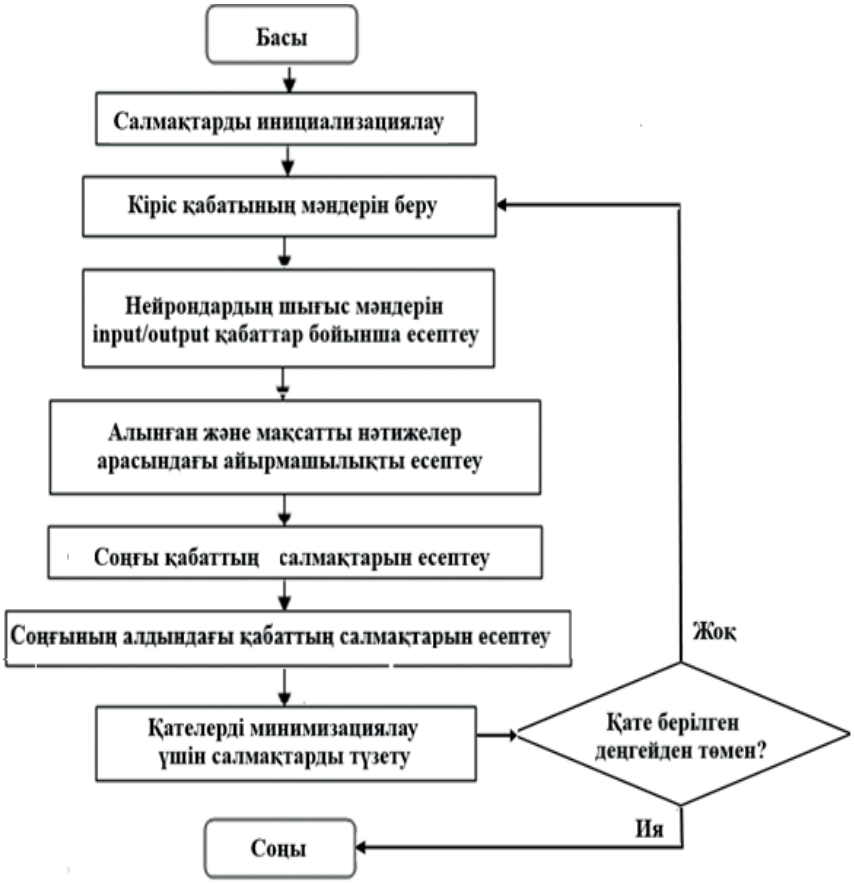
Сурет 2 – Иерархиялық көпқабатты нейрондық желі құрылымы

Көпқабатты персептрон нейрондардың бірнеше тізбектей қосылған қабаттарынан тұратын желі болып табылады. Көпқабатты персептронды

оқытудың мақсаты оның салмақтарын белгілі бір кірістер жиынын қолдану қажетті нәтижелер жиынтығына әкелетіндей реттеу болып табылады.

Бұл жұмыста қателерді кері тарату алгоритмін қолдана отырып, желінің барлық деңгейлері үшін көпдеңгейлі нейрондық желіні оқытудың екі тәсілдемесі: қателерді тура және артқа тарату ұсынылып, қолданылған. Тура өтуде кіріс векторы нейрондық желінің кіріс деңгейіне беріледі, содан кейін ол желі арқылы қабаттан қабатқа бөлінеді. Процесс желінің кіріс образына нақты реакциясы болып табылатын шығыс сигналдарының жиынтығын шығарады. Тура өту кезінде желінің барлық синаптикалық салмақтары бекітіледі, ал кері өту кезінде қатені түзету ережесі бойынша барлық синаптикалық салмақтар реттеледі. Желінің нақты шығысы қажетті мәнінен ауытқиды, нәтижесінде қателік сигналы пайда болады. Содан кейін бұл сигнал желі арқылы синаптикалық байланыстардың бағытына қарама-қарсы бағытта таралады.

Математикалық модельдер жүйесін құрудың жоғарыда келтірілген нәтижелері негізінде, Python тілінде жасанды нейрондық желіні оқытамыз. Төменде 3-суретте нейрондық желіні оқыту алгоритмінің схемасы берілген.



Сурет 3 – Жасанды нейрондық желінің қолданылған оқыту алгоритмінің схемасы

Айқын емес жиындар тәсілдерін қолдану нәтижесінде алынған мәліметтерді (Orazbayev et al., 2019) пайдалана отырып, Python тілінде қателердің кері таралуы алгоритмі программалық жүзеге асырылды. Нейрондық желі кірістерін алдын ала өңдеудің кең таралған түрлерінің бірі деректерді қалыпқа келтіру болып табылады. Деректерді нормалдау баған элементтерін деректердің максималды мәніне бөлу арқылы жүзеге асырылды.

Төменде дайындалдған Python тілінде жазылған программа коды келтірілген.

```
Import numpy as np
from numpy import *
import matplotlib.pyplot as plt
layer_0 = np.array([[0.9464],
        [0.9554],
         [0.714286],
  [0.928571]])
layer_t_0 = layer_0.transpose()
i1 = 0.9464
i2 = 0.9554
i3 = 0.714286
weights_0_1 = np.array([[0.2, 0.2, 0.3, 0.4],
[0.5, 0.2, 0.3, 0.2],
        [0.3, 0.8, 0.3, 0.3],
         [0.2, 0.2, 0.4, 0.2],
         [0.5, 0.3, 0.2, 0.3]])
weights_1_2 = np.array([[0.3, 0.3, 0.2, 0.2, 0.5],
        [0.2, 0.3, 0.6, 0.3, 0.4],
          [0.4, 0.4, 0.2, 0.3, 0.3],
          [0.4, 0.6, 0.6, 0.7, 0.7],
          [0.2, 0.2, 0.3, 0.3, 0.5]])
weights_2_3 = np.array([0.2, 0.2, 0.2, 0.3, 0.3])
alpha = 0.2
goal = np.array([0.821875])
x = 0
for x in range(0,200):
layer_1 = weights_0_1.dot(layer_0)
   i = 0
   for element in layer_1:
   layer_1[i] = 1 / (1 + np.exp(-element))
      i += 1
   sigmoid_layer_1 = layer_1
   layer_t_1 = sigmoid_layer_1.transpose()
   deriv_sigmoid_layer_1 = sigmoid_layer_1 * (1 - sigmoid_layer_1)
```

```
    layer_2 = weights_1_2.dot(sigmoid_layer_1)
      j = 0
       for element in layer_2:
          layer_2[j] = 1 / (1 + np.exp(-element))
          j += 1
       sigmoid_layer_2 = layer_2
       layer_t_2 = sigmoid_layer_2.transpose()
       deriv_sigmoid_layer_2 = sigmoid_layer_2 * (1 - sigmoid_layer_2
       sum_output = weights_2_3.dot(sigmoid_layer_2)
       sigmoid_layer_3 = 1 / (1 + np.exp(-sum_output))
       output = sigmoid_layer_3
       print(output)
    deriv_sigmoid_layer_3 = sigmoid_layer_3 * (1 - sigmoid_layer_3)
    error = (output - goal) ** 2
    error_layer_2 = error * weights_2_3
    error_layer_t_2 = error_layer_2.transpose()
    error_layer_1 = error_layer_2.dot(weights_1_2)
       error_layer_t_1 = error_layer_1.transpose()
    w1 = weights_0_1[0] + error_layer_t_1[0] * deriv_sigmoid_layer_1[0] *
layer_t_0 * alpha
    w2 = weights_0_1[1] + error_layer_t_1[1] * deriv_sigmoid_layer_1[1] *
layer_t_0 * alpha
    w3 = weights_0_1[2] + error_layer_t_1[2] * deriv_sigmoid_layer_1[2] *
layer_t_0 * alpha
    w4 = weights_0_1[3] + error_layer_t_1[3] * deriv_sigmoid_layer_1[3] *
layer_t_0 * alpha
    w5 = weights_0_1[4] + error_layer_t_1[4] * deriv_sigmoid_layer_1[4] *
layer_t_0 * alpha
    weights_0_1 = vstack([w1, w2, w3, w4, w5])
    m1 = weights_1_2[0] + error_layer_t_2[0] * deriv_sigmoid_layer_2[0] *
layer_t_1 * alpha
    m2 = weights_1_2[1] + error_layer_t_2[1] * deriv_sigmoid_layer_2[1] *
layer_t_1 * alpha
    m3 = weights_1_2[2] + error_layer_t_2[2] * deriv_sigmoid_layer_2[2] *
layer_t_1 * alpha
    m4 = weights_1_2[3] + error_layer_t_2[3] * deriv_sigmoid_layer_2[3] *
layer_t_1 * alpha
    m5 = weights_1_2[4] + error_layer_t_2[4] * deriv_sigmoid_layer_2[4] *
layer_t_1 * alpha
    weights_1_2 = vstack([m1, m2, m3, m4, m5])
    weights_2_3 = weights_2_3 + error * deriv_sigmoid_layer_3 * layer_t_2 *
alpha
```

```
import numpy as np
import matplotlib.pyplot as plt
from matplotlib import *
import pandas as pd
fig = plt.figure()
x = np.arange(1, 501)
y = np.array([…] )
ax = plt.gca()
plt.plot(x, y)
ax.set_xlabel("Epochs")
ax.set_ylabel("Output values")
plt.savefig('1.png', transparent=True)
fig = plt.figure()
x = np.arange(1, 201)
y = np.array([…])
ax = plt.gca()
plt.plot(x, y)
ax.set_xlabel("Epochs")
ax.set_ylabel("Errors")
plt.savefig('2.png', transparent=True)
```

Нәтижелерді талқылау. Күкірт өндіру процесінің моделін синтездеуде жасанды нейрондық желінің құру процесінің 3-суретте келтірілген оқыту алгоритмі негізінде келесі негізгі кезеңдерін келтіруге болады:

1. Оқыту жиынынан келесі оқыту жұбын таңдау; кіріс векторын нейрондық желі кірісіне беру;

2. Нейрондық желінің шығысын есептеп, анықтау;

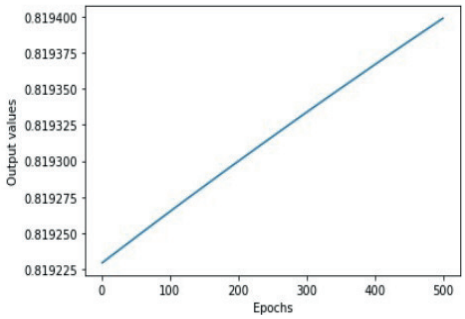
3. Нейрондық желі шығысы мен қажетті шығыс арасындағы айырмашылықты, яғни қатені есептеу;

4. Қатені минимизациялау үшін нейрондық желі салмағын коррекциялау;

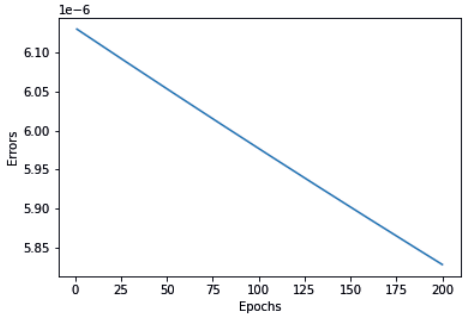
5. Минимизацияланған қате қажетті деңгейден төмен бе? шартын тексеру.

6. Егер минимизацияланған қате қажетті деңгейден төмен болмаса, бүкіл жиынтықтағы қате қажетті минималды деңгейге жеткенше оқыту жиынының әрбір векторы үшін 1–4 қадамдарды қайталау. Ал қажетті деңгейден төмен болса, онда нейрондық желі оқытылған борлып, есептеліп, КӨҚ процесін модельдеуге ұсынылады.

Программалық кодты орындау нәтижелері 4а), 4б) суреттерде келтірілген көрсетілген.



Сурет 4а – Дәуір (эпоха) санына байланысты шығыс мәндерінің графигі



Сурет 4б – Дәуір санына байланысты қателер мәндерінің графигі

Бұл суреттерде келтірілген графиктерді талқылау арқылы желінің шығыс мәндері дәуір санының артуына байланысты артатынын, ал қателер – керісінше дәуір саны артқан сайын төмендейтінін анықталған. Яғни дәуір саны артқан сайын, нәтижелер жақсара түседі. Алайда тым көп дәуір саны есептеу қиындығын және уақытын арттыруы мүмкін. Сондықтан практикада нейрондық желінің қажетті шығысы мен керекті минималды қатеге қол жеткізетін дәуірлер саны анықталып, қабылданады.

Қорытынды. Зерттеуде күкірт өндіру процесстерін жасанды нейрондық желілер негізінде тиімді модельдеу тәсілін әзірлеу және қолдану нәтижелері келтірілген. Көпқабатты жасанды нейрондық желінің қажетті құрылымын мен параметрлерін анықтау және оны қателердің кері таралуы алгоритмі негізінде оқыту нәтижесінде күкірт өндіру процесінің тиімді моделі алынған.

Жұмыстың зерттеу мақсатына қол жеткізілді, яғни күкірт өндіру процесстерін жасанды нейрондық желілер негізінде тиімді модельдеу тәсілі әзірленіп, зерттеу нысаны – Атырау МӨЗ күкірт өндіру қондырғысының негізгі агрегаттары жұмыс режимдерін компьютерлік модельдеу жүзеге асырылған.

Қойылған мақсатты жүзе асыру барысында келесі зерттеу міндеттері орындалып, төмендегі негізгі нәтижелер алынды:

- зерттеу нысанын ақпарат тапшылығы мен белгісіздігі жағдайында тиімді модельдеу үшін иерархиялық көпқабатты нейрондық желі құрылымы негізделіп, алынған;

- алынған нейрондық желіні программалық жүзеге асыру үшін қолданылатын программалау тілі, ортасы және құралдары таңдалған;

-жасанды нейрондық желіні оқыту тәсілі ретінде қателердің кері таралуы алгоритмі модификациясы негізделіп, оқыту алгоритмінің блок-схемасы келтірілген және оның негізгі кезеңдері сипатталған;

-авторлардың айқын емес жиындар тәсілдерін қолдану нәтижесінде алынған мәліметтерді пайдалана отырып, Python тілінде модификацияланған қателердің кері таралуы алгоритмі: итеративті градиент алгоритмі

программалық жүзеге асырылған. Дәуір санына байланысты шығыс мәндерінің және қателер мәндерінің графиктері тұрғызылған.

Тұрғызылған графиктерді талдау нәтижесінде, жасанды нейрондық желілер мен айқын емес жиындар тәсілдері қолдану арқылы арқылы алынған нәтижелері бір біріне сәйкес келеді деген қорытындыға келуге болады. Зерттеулер көрсеткендей, нейрондық желіні оқудағы қателіктер ең аз және шығыс мәні оның мақсатты мәніне мүмкіндігінше жақын. Айқын емес жиындар және жасанды нейрондық желі тәсілдері нәтижелері бойынша екі тәсіл ұқсас нәтиже береді, яғни, шығарылатын нәтижелер, күкірт көлемі сәйкес және олардың өндірісте тәжірибелік жолмен алынған нақты мәндерімен сәйкес келеді деп қорытында жасауға болады.

Әдебиеттер

Аксенов С.В., Новосельцев В.Б. (2022) Организация и использование нейронных сетей (методы и технологии). — Томск: Изд-во ТНУ. — 128 с.

Corazza M., Perna C., Pizzi C., Sibillo M. (2022) Mathematical and Statistical Methods for Actuarial Sciences and Finance. Springer International Publishing/-307 p. https://doi.org/10.1007/978-3-030-99638-3

Edmonson N. (2020) Industrial sulfur demand: Analysis and prospect. Nat Resour Res. 2. — Vol. 3. — P. 205–215 https://doi.org/10.1007/BF02259046

Garmroodi A., Shahsavand A., Mirzaei Sh. (2019) Maximization of sulfur recovery efficiency via coupled modification of GTU and SRU processes.Egyptian Journal of Petroleum. — Vol. 26. — P. 579–592.

Gazimzyanov N.R., Platonov O.I. (2020) Efficiency of a Claus furnace in the coke-oven gas desulphurization circuit of MMK. Coke Chem. — Vol. 60. — P. 199–202. https://doi.org/10.3103/S1068364X17050039

Голубева И.А., Хайруллина Г.Р., Старынин А.Ю., Каратун О.Н. (2020) Анализ производства серы методом Клауса на предприятиях нефтегазового комплекса. Химия нефти и газа. — № 3. — С. 10-18.

Hashemi M., Pourfayaz F., Mehrpooya M. (2019) Energy, exergy, exergoeconomic and sensitivity analyses of modified Claus process in a gas refinery sulfur recovery unit. J. Cleaner Prod. — Vol. 220. No 5. — P. 1071-1085.

Kadyrov S.R. (2023) Review article Investigation and comparison of technologies and methods of sulfur recovery and production processes. Bulletin of the oil and gas industry of Kazakhstan. — Vol. 5. — P. 91–98. https://doi.org/10.54859/kjogi108613

Karpunin I.I. (2023) Modeling and optimization of technological processes applied to packaging production. — Minsk: BNTU. — 124 p.

Кузнецов А.Г., Тулеуов Ж.Н. (2021) Обзор рынка серы и серной кислоты в Казахстане. Исследовательская группа «ИНФОМАЙН». Москва: Нефть и газ. — 108 с.

Naumova V.V., Ermolaeva V.A. Analysis of sulfur production technology using a catalytic reactor. International Journal of Applied Sciences and Technologies “Integral” 2023. — No. 2. — P. 289-299.

Nilsson N. (2019) Introduction to Machine Learning. Unpublished draft. Stanford University. — P.39-68.

Orazbayev B., Zhumadillayeva A., Orazbayeva K., Kurbangalieva N. (2024) Impact of key technological parameters on sulfur production processes. IEEE 13th international conference on Renewable energy research and application — ICRERA 2024. Nagasaki, Japan. 9-13 November. — P. 1-7.

Orazbayev B., Orazbaeva K.N., Utenova B.U., Tuleuov Zh.N. (2019) Models of the sulfur production process based on fuzzy information. Bulletin of KazNITU. — No. 2 (132). — P. 196-202.

Protalinsky O.M., Shcherbatov I.A., Belyaev I.O. (2021) Hybrid method of training neural

networks for classifying the catalytic stage of the Claus process. Theoretical Foundations of Chemical Technology. — Vol. 58, No. 3. — P. 38-43.

Shaimukhametov D.R., Mustafina S.A., Shaimukhametova D.V. (2019) Creation of an artificial neural network for determining the optimal temperature regime for a chemical process. Bulletin of the Bashkir University. — Vol. 24, No. 3. — P. 551-558.

Shao J. (2020) Mathematical Statistics. Springer Texts in Statistics. — 575 p. https://doi.org/10.1007/b97553

Tijanić K., Car-Pušić D., Šperac M. (2020) Cost estimation in road construction using artificial neural network. Neural Comput & Applic. — Vol. 32. — P. https://doi.org/10.1007/s00521-019-04443-y

Тулеуов Ж.Н. (2023) Технологический регламент установки производства серы Атырауского нефтеперерабатывающего завода. Атырау. Издательство Атырауского университета нефти и газа. — 178 с.

Xiaolong Hu. (2020) Global Finite-time Stability for Fractional-order Neural Networks. Opt. Mem. NeuralNetworks. —Vol. 29. — P.77–99. https://doi.org/10.3103/S1060992X20020046

References

Aksenov S.V., Novosel'tsev V.B. (2022) Organizatsiya i ispol'zovaniye neyronnykh setey (metody i tekhnologii) [Organization and use of neural networks (methods and technologies)]. - Tomsk: NTU Publishing House. — 128 p. (In Russian)

Corazza M., Perna C., Pizzi C., Sibillo M. (2022) Mathematical and Statistical Methods for Actuarial Sciences and Finance. Springer International Publishing/-307 p. https://doi.org/10.1007/978-3-030-99638-3 (in English)

Edmonson N. (2020) Industrial sulfur demand: Analysis and prospect. Nat Resour Res. 2. — Vol. 3. — P. 205–215 https://doi.org/10.1007/BF02259046 (in English)

Garmroodi A., Shahsavand A., Mirzaei Sh. (2019) Maximization of sulfur recovery efficiency via coupled modification of GTU and SRU processes.Egyptian Journal of Petroleum. — Vol. 26. — P. 579–592. (in English)

Gazimzyanov N.R., Platonov O.I. (2020) Efficiency of a Claus furnace in the coke-oven gas desulphurization circuit of MMK. Coke Chem. — Vol. 60. — P. 199–202. https://doi.org/10.3103/S1068364X17050039 (in English)

Golubeva I.A., Khayrullina G.R., Starynin A.YU., Karatun O.N. (2020) Analiz proizvodstva sery metodom Klausa na predpriyatiyakh neftegazovogo kompleksa [Analysis of sulfur production by the Claus method at oil and gas enterprises]. Oil and Gas Chemistry. 2020. No. 3. — P 10–18. (In Russian)

Hashemi M., Pourfayaz F., Mehrpooya M. (2019) Energy, exergy, exergoeconomic and sensitivity analyses of modified Claus process in a gas refinery sulfur recovery unit. J. Cleaner Prod. — Vol. 220. No 5. — P. 1071-1085. (in English)

Kadyrov S.R. (2023) Review article Investigation and comparison of technologies and methods of sulfur recovery and production processes. Bulletin of the oil and gas industry of Kazakhstan. — Vol. 5. — P. 91–98. https://doi.org/10.54859/kjogi108613 (in English)

Karpunin I.I. (2023) Modeling and optimization of technological processes applied to packaging production. — Minsk: BNTU. — 124 p. (in English)

Kuznetsov A.G., Tuleuov Zh.N. (2021) Obzor rynka sery i sernoy kisloty v Kazakhstane [Review of the sulfur and sulfuric acid market in Kazakhstan]. INFOMINE Research Group. Moscow, Oil and Gas. — 108 p. (In Russian)

Naumova V.V., Ermolaeva V.A. Analysis of sulfur production technology using a catalytic reactor. International Journal of Applied Sciences and Technologies "Integral" 2023. — No. 2. — P. 289-299. (in English)

Nilsson N. (2019) Introduction to Machine Learning. Unpublished draft. Stanford University. — P. 39-68. (in English)

Orazbayev B., Zhumadillayeva A., Orazbayeva K., Kurbangalieva N. (2024) Impact of key

technological parameters on sulfur production processes. IEEE 13th international conference on Renewable energy research and application — ICRERA 2024. Nagasaki, Japan. 9-13 November. — P. 1-7 (in English)

Orazbayev B., Orazbaeva K.N., Utenova B.U., Tuleuov Zh.N. (2019) Models of the sulfur production process based on fuzzy information. Bulletin of KazNITU. — No. 2 (132). — P. 196-202. (in English)

Protalinsky O.M., Shcherbatov I.A., Belyaev I.O. (2021) Hybrid method of training neural networks for classifying the catalytic stage of the Claus process. Theoretical Foundations of Chemical Technology. — Vol. 58, No. 3. — P. 38-43. (in English)

Shaimukhametov D.R., Mustafina S.A., Shaimukhametova D.V. (2019) Creation of an artificial neural network for determining the optimal temperature regime for a chemical process. Bulletin of the Bashkir University. — Vol. 24, No. 3. — P. 551-558. (in English)

Shao J. (2020) Mathematical Statistics. Springer Texts in Statistics. — 575 p. https://doi.org/10.1007/b97553 (in English)

Tijanić K., Car-Pušić D., Šperac M. (2020) Cost estimation in road construction using artificial neural network. Neural Comput & Applic. — Vol. 32. — P. https://doi.org/10.1007/s00521-019-04443-y (in English)

Tuleuov Zh.N. (2023) Tekhnologicheskiy reglament ustanovki proizvodstva sery Atyrauskogo neftepererabatyvayushchego zavoda. [Technological regulations of the sulfur production unit of the Atyrau oil refinery]. Atyrau. Publishing house of the Atyrau University of Oil and Gas. — 178 p. (In Russian)

Xiaolong Hu. (2020) Global Finite-time Stability for Fractional-order Neural Networks. Opt. Mem. NeuralNetworks. —Vol. 29. — P. 77–99. https://doi.org/10.3103/S1060992X20020046 (in English)

ACADEMIC SCIENTIFIC JOURNAL OF COMPUTER SCIENCE
ISSN 1991-346X
Volume 4. Namber 356 (2025). 234–258

https://doi.org/10.32014/2025.2518-1726.394

УДК 004.855.5

© L. Rzayeva[1], A. Ryzhova[1], M. Zhaparkhanova[1], A. Myrzatay[1], Zh. Kozhakhmet[1], 2025.
[1]Astana IT University, Astana, Kazakhstan;
[2]RSE on PCV "Digital Government Support Center", Astana, Kazakhstan.
E-mail: l.rzayeva@astanait.edu.kz

A NEW LSTM-BASED WEB APPLICATION FOR AUTOMATED PASSWORD STRENGTH EVALUATION

Rzayeva Leila — PhD, Associate Professor, Director of the "CyberTech" Research and innovation Center, Astana IT University, Astana, Kazakhstan,
E-mail: l.rzayeva@astanait.edu.kz, https://orcid.org/0000-0002-3382-4685;
Ryzhova Alissa — Junior Researcher at the "CyberTech" Research Center, Astana IT University, Astana, Kazakhstan,
E-mail: alissaryzh@gmail.com, https://orcid.org/0009-0006-3781-1773;
Zhaparkhanova Merei — Junior Researcher at the "CyberTech" Research and innovation, Kazakhstan,
E-mail: zhaparkhanovamerei.16@gmail.com**,** zhaparkhanovamerei.16@gmail.com, https://orcid.org/0009-0002-7177-7805;
Myrzatay Ali — PhD, Junior Researcher at the "CyberTech" Research and innovation Center, Astana IT University LLP, Astana, Kazakhstan,
E-mail: mirzataitegiali@gmail.com; ORCID ID: https://orcid.org/0000-0002-5339-2437;
Kozhakhmet Zhaksylyk — Junior Researcher at the "CyberTech" Research and innovation Center, Astana IT University LLP, Astana, Kazakhstan,
E-mail: zh.kozhakhmet@astanait.edu.kz; ORCID ID: https://orcid.org/0009-0002-5449-3317.

Abstract. According to the latest Verizon DBIR report, poor handling of credentials, including password reuse and the human factor in their creation remains one of the key attack vectors. A study conducted by the authors showed that most users change passwords only when they are lost, and 35% consider mandatory regular changes inconvenient. This underscores the need for technical solutions capable of clearly demonstrating system vulnerabilities and raising security awareness. Within the study, the human factor in creating usernames and passwords is considered a vulnerability. Identifying the patterns and rules people use significantly reduces the number of combinations an attacker must try to gain access. The proposed method is based on a character-level LSTM model that detects recurring structures and generates generalized masks reflecting characteristic patterns. Training was performed on public datasets containing 31,000 compromised passwords. The model achieved over 90% accuracy on the test set without signs of overfitting.

The approach combines the analysis of individual user habits in generating logins and passwords with automatic keyword extraction from open sources using a keyword extraction algorithm. The method is integrated into a web application that allows local fine-tuning of the model, running it via ONNX, and performing all computations on the device, ensuring data confidentiality and compliance with information security requirements.

Keywords: cybersecurity; password security; machine learning; neural networks; social engineering; digital forensic; cryptography; behavioural analysis

© Л. Рзаева[1], А. Рыжова[1], М. Жапарханова[1], А. Мырзатай[1,2], Ж. Кожахмет[1], 2025.

[1]Astana IT University, Астана, Қазақстан;

[21]Цифрлық үкіметті қолдау орталығы" ШЖҚ РМК, Астана, Қазақстан.

E-mail: l.rzayeva@astanait.edu.kz

ҚҰПИЯСӨЗДІҢ БЕРІКТІГІН АВТОМАТТЫ БАҒАЛАУҒА АРНАЛҒАН LSTM НЕГІЗІНДЕГІ ЖАҢА ВЕБ-ҚОСЫМША

Рзаева Лейла — PhD, қауымдастырылған профессор, «CyberTech» ҒЗО директоры, Astana IT University, Астана, Қазақстан,
E-mail: l.rzayeva@astanait.edu.kz, https://orcid.org/0000-0002- 3382-4685;
Рыжова Алиса — кіші ғылыми қызметкер, «CyberTech» ҒЗО, Astana IT University, Астана, Қазақстан,
E-mail: alissaryzh@gmail.com, https://orcid.org/0009-0006-3781-1773;
Жапарханова Мерей — кіші ғылыми қызметкер, «CyberTech» ҒЗО, Astana IT University, Астана, Қазақстан,
E-mail:zhaparkhanovamerei.16@gmail.comб https://orcid.org/0009-0002-7177-7805;
Мырзатай Али — PhD, аға ғылыми қызметкер, «CyberTech» ҒЗО, Astana IT University, Астана, Қазақстан,
E-mail:mirzataitegiali@gmail.comб https://orcid.org/0000-0002-5339-2437;
Кожахмет Жаксылык — ғылыми қызметкер, «CyberTech» ҒЗО, Astana IT University, Астана, Қазақстан,
E-mail:zh.kozhakhmet@astanait.edu.kzб https://orcid.org/0009-0002-5449-3317.

Аннотация. Бұл мақалада көптілді мәтіндерді өңдеу мен семантикалық қауіптілік талдауына арналған цифрлық криминалистика жүйесінің заманауи архитектурасы ұсынылады. Негізгі мәселе – құрылымдалмаған хабарламалардан маңызды ақпаратты автоматты түрде бөліп алып, оны нақты қауіп категорияларына жатқызу қажеттілігі. Бұл мәселе құқық қорғау органдарының заманауи тергеу процестерінде жиі кездеседі және оны шешу жедел әрекет етуді, масштабталатын жүйелер мен көптілділікке төзімділікті талап етеді. Ұсынылған жүйе Python/FastAPI негізінде құрылған серверлік архитектураны, Qdrant векторлық дерекқорын, алдын ала үйретілген jina-embeddings-v3 үлгісін және Llama-Guard-3-1B қауіп классификациялау моделін пайдаланады. Жүйе мәтінді векторизациялау, семантикалық іздеу, тілдерді автоматты анықтау және көптілді аударма, сонымен қатар

қауіпті контентті санатқа бөлу сынды бірнеше кезеңнен тұрады. Негізгі гипотеза – семантикалық векторлық модельдер мен категориялық гибридті классификацияны біріктіру арқылы қауіпсіздік деңгейін жоғары дәлдікпен анықтауға болады. Эксперименттік нәтижелер жүйенің нақты уақытта өңдеу қабілетін көрсетті (7448 хабарлама ~3 секундта индекстеледі) және Llama-Guard моделінің 75.19% дәлдікпен хабарламаларды қауіптілік деңгейі мен типі бойынша жіктей алатынын растады. Орташа жауап уақыты аудармасыз сұраныстарда ~0.45 сек және аударманы қажет ететін жағдайларда ~5.23 сек құрайды. Жүйенің визуалды модулі пайдаланушыға хабарламаларды топтар бойынша, қолданбалар, алушылар, геолокациялар және уақыттық үлгілер бойынша талдауға мүмкіндік береді. Бұл функционал криминалистикалық сараптаманың тиімділігін арттырып, тергеудің негізгі кезеңдерін автоматтандыруға жағдай жасайды. Интерфейс жүйені ыңғайлы басқаруға, деректерді жеңілдетуге және шешім қабылдауды жылдамдатуға көмектеседі. Ұсынылған шешім құқық қорғау, ұлттық қауіпсіздік, ақпараттық қауіпсіздік және киберқылмысқа қарсы күрес салаларында практикалық қолдануға бейімделген және болашақта аудио/видео деректермен кеңейтілуі мүмкін.

Түйін сөздер: табиғи тілдерді өңдеу (NLP), мәтінді векторизациялау, семантикалық іздеу, көптілді талдау, автоматты машиналық аударма, криминалистика, қатерлерді жіктеу

© Л. Рзаева[1], А. Рыжова[1], М. Жапарханова[1], А. Мырзатай[1,2], Ж. Кожахмет[1], 2025.

[1] НИЦ “CyberTech” Astana IT University, Астана, Казахстан;

[2] РГП на ПХВ Центр поддержки цифрового правительства, Астана, Казахстан.

E-mail: l.rzayeva@astanait.edu.kz

НОВОЕ ВЕБ-ПРИЛОЖЕНИЕ НА ОСНОВЕ LSTM ДЛЯ АВТОМАТИЗИРОВАННОЙ ОЦЕНКИ НАДЕЖНОСТИ ПАРОЛЕЙ

Рзаева Лейла — PhD, Ассоциированный профессор, Директор НИЦ «CyberTech», Astana IT University, Астана, Казахстан,
E-mail: l.rzayeva@astanait.edu.kz, https://orcid.org/0000-0002-3382-4685;
Рыжова Алиса — младший научный сотрудник НИЦ «CyberTech», Astana IT University, Астана, Казахстан,
E-mail: alissaryzh@gmail.com, https://orcid.org/0009-0006-3781-1773;
Жапарханова Мерей — младший научный сотрудник НИЦ «CyberTech», Astana IT University, Астана, Казахстан,
E-mail: zhaparkhanovamerei.16@gmail.com, https://orcid.org/0009-0002-7177-7805;
Мырзатай Али — PhD, старший научный сотрудник НИЦ «CyberTech», Astana IT University, Астана, Казахстан,
E-mail: mirzataitegiali@gmail.com, https://orcid.org/0000-0002-5339-2437;
Кожахмет Жаксылык — научный сотрудник НИЦ «CyberTech», Astana IT University, Астана, Казахстан,
E-mail: zh.kozhakhmet@astanait.edu.kz, https://orcid.org/0009-0002-5449-3317.

Аннотация. Согласно последнему отчёту Verizon DBIR, неграмотное обращение с учётными данными, включая повторное использование паролей и влияние человеческого фактора при их создании, остаётся одним из ключевых векторов атак. Исследование, проведённое авторами, показало, что большинство пользователей меняют пароли лишь при их утрате, а 35 % считают обязательную регулярную смену неудобной. Это показывает, насколько необходимы решения, которые способны наглядно демонстрировать уязвимости и повышать осведомленность в вопросах безопасности паролей. В рамках данного исследования человеческий фактор при создании паролей рассматривается как уязвимость. Определение типичных шаблонов и правил значительно сокращает количество комбинаций и время, которые злоумышленнику необходимо для получения доступа. Предложенный метод основан на LSTM-модели, выявляющей повторяющиеся структуры и маски, отражающие характерные паттерны на уровне отдельных символов. Для обучения использовались датасеты в открытом доступе, включающие 31 000 скомпрометированных паролей. Модель показала более 90% точности на тестовой выборке без переобучения. Подход объединяет анализ индивидуальных привычек и шаблонов пользователя при генерации паролей с автоматическим извлечением ключевых слов из открытых источников с помощью алгоритма keyword extraction. В настоящее время метод интегрирован в веб-приложение, которое позволяет предварительно обучать модель, запускать ее через ONNX и выполнять все вычисления на конечном устройстве пользователя. Это обеспечивает конфиденциальность данных и соответствие требованиям информационной безопасности.

Ключевые слова: кибербезопасность, безопасность паролей, машинное обучение, нейронные сети, социальная инженерия, цифровая криминалистика, криптография, поведенческий анализ

Введение. В эпоху цифровых технологий большинство людей неосознанно раскрывают личную информацию, (имена, даты рождения, клички домашних животных, личные хобби), в социальных сетях, которые часто служат основой для их учетных данных. Эти элементы, особенно в сочетании с распространенными шаблонами или «масками», могут быть легко использованы злоумышленниками с помощью социальной инженерии.

Хотя существует множество инструментов для оценки надежности паролей на основе традиционных политик, таких как длина и разнообразие символов (Bergeron & Dearden, 2024; Atzori et al., 2024; He et al., 2021) они не учитывают индивидуальную логику, лежащую в основе формирования паролей пользователями, поскольку для этого потребовалось бы отслеживать историю использования логинов и паролей, личные данные и привычки формирования пароля.

Несмотря на широкое применение дополнительных методов аутентификации, например таких как биометрия и двухфакторная верификация, односложные буквенные пароли по-прежнему остаются

основным методом доступа. Представленный метод, который оценивает потенциальную уязвимость пароля путем анализа открытых данных и экспорта из менеджеров паролей, показывает, как индивидуальные привычки и присутствие в Интернете влияют на его предсказуемость. В результате они представляют собой серьезную угрозу безопасности: в 2024 году было скомпрометировано более 5,5 миллиарда учетных записей по всему миру, что почти в восемь раз больше, чем в 2023 году (Surfshark, 2020). Данные опроса выявили, что пользователи часто используют привычные методы: 42% смешивают значимые слова и цифры, 34% следуют базовым требованиям платформы, а 32% повторно используют части предыдущих паролей (Urrico, 2024). Существующие решения, как правило, не учитывают индивидуальные правила формирования паролей конкретного пользователя. Некоторые допускают такие пароли, как «имяфамилия1!», или вовсе не анализируют часто повторяющиеся шаблоны в базах данных утечек.

Это исследование, являясь частью большого исследовательского проекта (Rzayeva et al., 2025; Idrissova et al., 2025) фокусируется на широко распространенной тенденции пользователей игнорировать предложения системы по созданию паролей, при этом почти 90% предпочитают создавать собственные комбинации на основе знакомых и предсказуемых шаблонов (Buckman, 2025). Эти шаблоны, как правило, следуют одному из нескольких общих правил: слова из словаря, числовые последовательности, (такие как номера телефонов или даты рождения), или символьные маски — причем символы редко используются отдельно. Описанный подход фокусируется на личных привычках пользователя, учитывает характерные паттерны и предлагает удобный формат в виде расширения для Google Chrome с возможностью анализа в режиме реального времени.

Научный вклад данного исследования содержит: анализ существующих методов взлома паролей и выявление слабых мест в паролях, созданных пользователями; анализ правил формирования паролей на основе данных из утечек; проектирование и обучение LSTM (Long Short-Term Memory) для выявления структурных шаблонов; разработку инструмента, предоставляющего индивидуальную обратную связь с применением методов социальной инженерии.

В отличие от существующих работ, данное исследование сосредоточено на взломах, вызванных человеческим фактором, а не на технических уязвимостях или методах перебора. Если злоумышленник получает доступ к предыдущим комбинациям логина и пароля, (например, через локальные устройства или базы утечек), и сочетает это с открытыми данными из социальных сетей, он может вывести правила генерации паролей пользователя, что значительно ускоряет получение несанкционированного доступа к его учетным записям. Новизна исследования заключается в предложенном методе проверки паролей, который учитывает человеческий фактор и зависимость от социальной инженерии, игнорируемые традиционными средствами проверки. Идентифицируя общие паттерны, мы создаем обобщённые маски паролей,

выделяющие структурные слабые места, потенциально позволяющие в будущем взломать другие учетные записи.

Материалы и методы:

1.2 **Новый метод обнаружения повторяющихся правил при создании логина/пароля**

Анализ поведения пользователей показывает, что при отказе от использования автоматически сгенерированных системой паролей они демонстрируют склонность к повторению одних и тех же логических схем и правил их формирования, внося минимальные изменения. Такая повторяемость может быть связана с намерением упростить процесс запоминания, снижая общую стойкость паролей к подбору.

Новый метод, логическое объяснение которого представлено на рисунке 1, вводится путем анализа прозрачности и предсказуемости логинов и паролей.

Важно отметить, что метод можно использовать только при наличии хотя бы части истории логинов и паролей пользователя. Именно на основе этой истории создается новая научная и прикладная работа. В отличие от существующих методов, этот подход предполагает, что каждый пользователь имеет свои собственные устойчивые правила формирования паролей.

Метод объединяет два типа анализа. Первый - поиск повторяющихся структур и паттернов в последовательности логинов и паролей пользователя. Второй - выделение его личного «словаря» на основе активности в социальных сетях. Используя результаты этих двух шагов, система в режиме реального времени оценивает, насколько новый пароль выбивается из привычных шаблонов, и формирует персональные рекомендации по усилению его надёжности.

Рисунок 1 — Логическая схема предлагаемого метода

Сбор эмпирических данных, их техническая обработка, обучение моделей и проверка пользовательских сценариев — все это были шаги, которые использовались для достижения целей исследования. Исходные данные были созданы с использованием открытых наборов скомпрометированных паролей (RockYou, LinkedIn, GitHub и другие). После удаления дублей, служебных символов и неправильных записей был создан очищенный массив более 31 000 анонимизированных строк, который был готов к дальнейшему анализу.

Обучающая выборка дополнительно пополнялась открытыми датасетами Kaggle (Password dataset 2024) и GitHub (Kkrypt0nn., n.d.-b). На подмножестве этих данных обучалась модель Decision Tree Classifier; перед запуском обучения каждому паролю вручную присваивался тип. В ходе обучения алгоритм последовательно отбирает наиболее информативные признаки и формирует древовидную структуру (рис. 2), в которой внутренние узлы соответствуют логическим условиям, разделяющим выборку на более однородные группы по целевому признаку.

После завершения обучения модель была протестирована на отложенной выборке, что позволило оценить ее способность обобщать закономерности на новые, ранее не встречавшиеся данные. Для анализа результатов были получены прогнозы модели, которые затем сравнивались с фактическими метками классов, что позволило количественно оценить точность классификации. Однако впоследствии это решение было изменено в пользу более гибких и выразительных моделей, способных учитывать последовательный характер паролей и выявлять сложные скрытые зависимости в структуре символов.

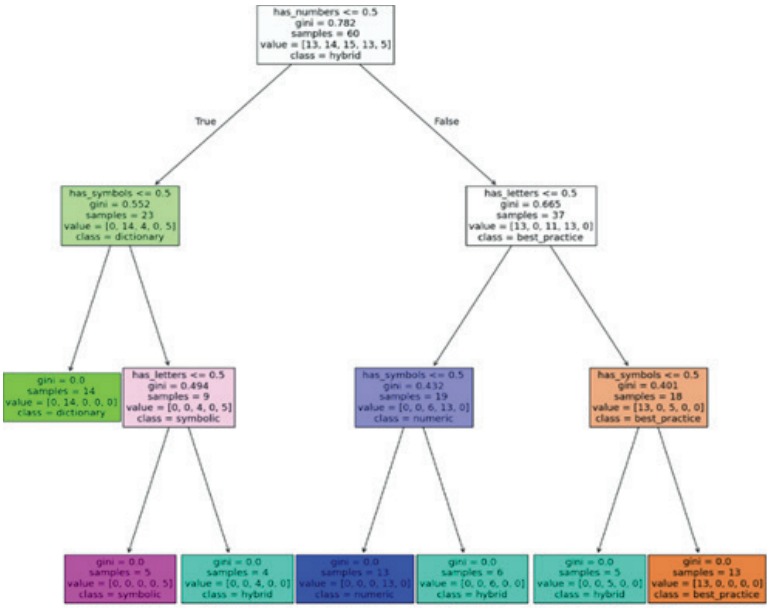


Рисунок 2 – Результаты Decision Tree

Классификатор на основе дерева решений обучается сопоставлению:

$$f(x) \rightarrow \{0, 1, \ldots, K-1\} \quad (1)$$

где x = [x_1, x_2, x_3] бинарные features показывающее присутствие букв (x_1), цифр (x_2) и символов (x_3) в пароле, и K это количество возможных типов пароля.

В каждом узле модель выбирает признак и пороговое значение для разделения данных таким образом, чтобы минимизировать нечистоту Джини:

$$G = 1 - \Sigma\, p_i^2 \quad (2)$$

где p_i – это доля класса i в узле. Дерево рекурсивно разбивает пространство признаков и присваивает наиболее часто встречающуюся метку класса каждому листу. В результате получается интерпретируемая модель классификации типов паролей на основе лексических свойств.

Дерево решений, традиционный класс алгоритмов, используемый в задачах классификации, использовался в качестве одной из основных моделей исследования. На начальном этапе они использовались для разделения паролей по простым признакам, таким как наличие букв, цифр, особых символов и их комбинаций. Это помогло задокументировать стандартные схемы паролей и повторяющихся комбинаций, которые впоследствии использовались как основа для более тщательного анализа и создания обобщенных масок. При этом модель дерева решений служила как стартовым фильтром, так и источником эмпирических гипотез. Это привело к переходу к более гибким последовательностным моделям, таким как LSTM.

В дополнение к этому были проведены исследования метода контекстно-свободной грамматики на основе вероятности (PCFG), который был направлен на определение устойчивых структур в паролях пользователей. В этом случае пароль представляется последовательностью логических фрагментов, состоящих из подпоследовательностей букв (L), цифр (D) и специальных символов (S). Вероятностная грамматика строится по множеству этих разбиений, учитывая тип и длину каждого сегмента.

Для каждого пароля была сгенерирована структура, например, L5-D4 для пароля Merei1611. С точки зрения формальности, PCFG модель разбивает пароль p по частям s_1, s_2, …, s_n, где каждая часть $s_i = (\tau_i, v_i)$, с $\tau_i \in \{L, D, S\}$ представляет тип символов (Буквы, цифры, специальные символы), и v_i относится к части пароля.

Структурная часть пароля представлена так:

$$\sigma(p) = \tau_1|v_1| - \tau_2|v_2| - \ldots - \tau_n|v_n| \quad (3)$$

Например пароль "Merei1611" соответствует шаблону L5–D4.

Вероятность шаблона представлена как:

$$P(\sigma) = C(\sigma) / N$$

где C(σ) это частота шаблона σ в датасете, и N это общее количество возможных шаблонов.

Условная вероятность сегментов:

$$P(v_i \mid \tau_i) = C_i(v_i) / \Sigma C_i \tag{4}$$

где $C_i(v_i)$ подсчет под частей слова v_i среди сегментов типа τ_i, и ΣC_i это общее количество всех сегментов (Фигура 6)

Предполагая независимость между сегментами, общая вероятность пароля составляет:

$$P(p) = P(\sigma(p)) \times \Pi\, P(v_i \mid \tau_i) \tag{5}$$

Затем надёжность пароля рассчитывается как нормализованный отрицательный логарифм по основанию 10 этой вероятности:

$$Score(p) = -(1 / |p|) \times log_{10} P(p) \tag{6}$$

Чем выше балл, тем более редкий и надежный пароль. Модель оценивает уникальность и предсказуемость паролей путем анализа их структурных и лексических паттернов. После обучения на специально подобранном наборе данных, в котором были зафиксированы частоты встречаемости этих структур, а также частоты встречаемости определенных подстрок в рамках категорий. На этапе оценки паролей модель использует логарифмическую метрику, которая отражает вероятность структуры и ее элементов с поправкой на длину.

Модель обучалась итеративно, наблюдая за прогрессом по среднему баллу и точности распознавания структур, которые были известны. Чтобы продемонстрировать это, были созданы два графика. На рис. 4 показано изменение среднего логарифмического балла на обучающих и валидационных выборках по мере увеличения обучающей группы, а на рис. 3 показана точность сопоставления структур на обеих выборках. Это позволяет наблюдать за тем, насколько хорошо модель адаптируется к новым данным, и обнаруживать паттерны, которые часто встречаются в паролях.

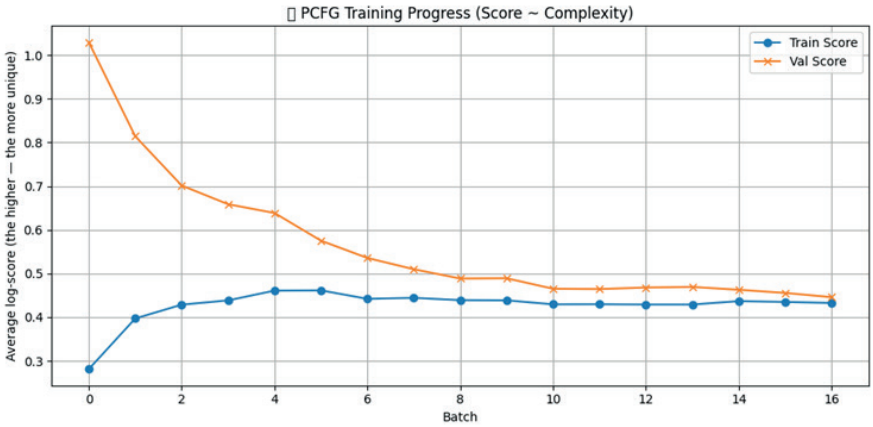


Рисунок 3 – PCFG Training Progress

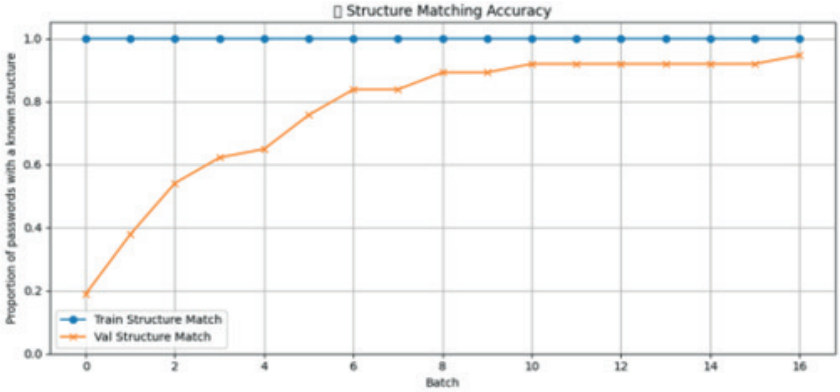


Рисунок 4 –Structure Matching Accuracy

Модели PCFG и деревья решений опираются на заранее определенные правила и характеристики для анализа паролей, например PCFG представляет пароли в виде фрагментов и определяет вероятность появления этих паттернов. Модель была выбрана потому что является негласным стандартом в анализе структуры паролей. Этот метод был применен для изучения распределения структурных паттернов и оценки вероятности использования определенных форматов паролей.

Однако оба подхода, недостаточно гибкие, так как не учитывают влияние человеческого фактора на составление учетных данных. Деревья решений используют жесткие логические разделения, а PCFG плохо справляется с нестандартными шаблонами, индивидуальными привычками или эмоционально заряженными комбинациями.

Эти ограничения привели к переходу на другую архитектуру. Основой анализа паролей, описанного в этой статье, является LSTM. Без ручной настройки правил эта модель способна обучаться на сырых данных и

определять сложные персонализированные шаблоны. В области анализа паролей способность LSTM обрабатывать последовательные данные имеет решающее значение, поскольку сложность пароля может быть изменена даже незначительными изменениями символов. Кроме того, было показано, что модели LSTM эффективны для задач, таких как генерация текста и предсказание следующего элемента.

3.3. Техническое выполнение предложенного метода

Архитектура разработанной модели состоит из трех логических уровней. Первый этап включает входных данных посимвольно изменяя в плотные векторные представления фиксированной размерности, что позволяет проводить эффективное обучение по сравнению с классическим однозначным кодированием, поскольку дает модели возможность улавливать синтаксические и структурные особенности символов, а не только их расположение в алфавите.

На втором этапе применяется двухслойная LSTM модель, которая анализирует контекстуальные отношения во всей строке пароля. Таким образом, модель может распознавать не только комбинации цифр и букв, но и положения особых символов, специфичных для способностей пользователей. Механизм регуляризации помогает предотвратить переобучение на ограниченных наборах данных, а использование двух слоев облегчает обработку долгосрочных и краткосрочных зависимостей между символами.

На последнем этапе используется полностью подключенный слой для преобразования выходных состояний LSTM в вероятностное распределение по всем символам, доступным в словаре. Модель может предсказать не только следующий символ, но и оценить вероятность того, что определенные паттерны будут присутствовать в пароле.

Основные цели моделирования структуры паролей и уровень абстракции отличают модель LSTM, предложенную в этом исследовании, от подхода на основе нейронных сетей (Melicher et al., 2016). Использование рекуррентной нейронной сети на уровне символов позволяет оценить вероятность перехода между отдельными символами. Это позволяет провести детальное моделирование угадываемости паролей на основе зависимостей от места. Их архитектура адаптирована к крупномасштабному моделированию атак, уделив особое внимание вероятности последовательностей символов в паролях.

В отличие от этого, в исследовании уделяется больше внимания поведенческому анализу при создании паролей с целью выявления семантически значимых правил и закономерностей, отражающих когнитивные привычки человека и социальную активность. Для построения обобщенных структурных масок предполагается использование предлагаемой архитектуры LSTM модели для обнаружения повторяющихся подстрок и типичных комбинаций, таких как имена и годы рождения.

Таким образом, наш метод предоставляет дополнительную информацию

о человеческом факторе, который способствует созданию предсказуемых и уязвимых паролей. Это отличается от методов, которые используются для моделирования вероятности очевидности паролей (Melicher et al., 2016).

Алгоритм оптимизации Адам, который обеспечивает стабильную и быструю конвергенцию в глубоких нейронных сетях, использовался для проведения обучения. Кросс-энтропия, которая является стандартной для задач многоклассовой классификации, использовалась в качестве функции потерь. Для обеспечения воспроизводимости и стабильности результатов во всех модулях случайности было установлено фиксированное значение семени. Это позволяет проводить эксперименты снова и снова в тех же условиях и получать те же результаты. Даже при обработке больших объемов строковых данных процесс выполнялся на локальной рабочей станции с графическим ускорителем, что обеспечивало достаточное время обучения. Набор данных для обучения состоял из тысяч паролей, которые были получены из общедоступных утечек, которые были предварительно очищены, чтобы убрать дублики и неинформативные символы. Входные последовательности обычно состояли от 8 до 24 символов, что свидетельствует о том, насколько часто пользователи используют пароли.

В математическом плане LSTM представляет собой набор формул, которые обрабатывают последовательность, такую как текст, шаг за шагом. С помощью специальных «ворот», управляемых значениями от 0 до 1, он определяет, что запомнить, а что забыть. Все это основано на стандартной математике, включая умножение, матричные операции и нелинейные функции, такие как сигмоид и гиперболический тангенс.

В соответствии с документацией PyTorch torch.nn.LSTM, следующие формулы описывают операции LSTM для первого слоя модели PasswordLSTM с embedding_dim = 128, hidden_dim = 256 и num_layers = 2.

Основная часть архитектуры LSTM, входные ворота(it), отвечают за регулирование включения информации в состояние ячейки. Он контролирует степень влияния входных данных на обновление ячейки памяти на текущем временном шаге.

$$i_t = \sigma(W_{ii}x_t + b_{ii} + W_{hi}h_{t-1} + b_{hi}) \quad (7)$$

где

$W_{ii} \in R^{256\times128}, W_{hi} \in R^{256\times256}, b_{ii}, b_{hi} \in R^{256}, x_t \in R^{128}, h_{t-1} \in R^{256}$

$x_t \in R^{128}$ – входной вектор по временному периоду t

$h_{t-1} \in R^{256}$ – скрытое состояние прошлого временного отрезка

$W_{ii} \in R^{256} \times {}^{128}$ – матрица весов для входа

$W_{hi} \in R^{256} \times {}^{256}$ – матрица весов для скрытого состояния

$b_{ii}, b_{hi} \in R^{256}$ – векторы погрешности

$\sigma(\cdot)$ – сигмоидная функция активации

Ворота забывания (f_t) — часть LSTM, решающая, какую информацию удалить из памяти. На вход подаются текущие данные и предыдущее состояние; на выходе — числа от 0 до 1, где 0 означает «полностью забыть», а 1 — «оставить без изменений».

$$f_t = \sigma(W_{if}x_t + b_{if} + W_{hf}h_{t-1} + b_{hf}) \quad (8)$$

где $x_t \in R^{128}$ – входной вектор на шаге t,

$h_{t-1} \in R^{256}$ – скрытое состояние с предыдущего шага

$W_{if} \in R^{256\times128}$ – матрица весов для входа x_t

$W_{hf} \in R^{256\times256}$ – матрица весов для скрытого состояния h_{t-1}

$b_{if,b_{hf}} \in R^{256}$ – векторы смещений,

$\sigma(\cdot)$ – сигмоидальная функция активации.

Клеточные ворота (кандидатное состояние ячейки) ($\tilde{C}_t$) — часть LSTM, формирующая новое содержимое памяти, т.е. предлагающая, что можно добавить. Значения вычисляются на основе входа и предыдущего состояния, но напрямую не добавляются — это делает входной клапан.

$$\tilde{C}_t = tanh(W_{ig}x_t + b_{ig} + W_{hg}h_{t-1} + b_{hg}) \quad (9)$$

где $x_t \in R^{128}$ – входной вектор на шаге t

$h_{t-1} \in R^{256}$ – скрытое состояние с предыдущего шага,

$W_{ig} \in R^{256\times128}$ – матрица весов по входу для кандидата,

$W_{hg} \in R^{256\times256}$ – матрица весов по скрытому состоянию для кандидата,

$b_{ig}, b_{hg} \in R^{256}$ – векторы смещений,

$tanh(\cdot)$ – гиперболический тангенс.

Выходные ворота (o_t) — часть LSTM, определяющая, какую часть текущего состояния памяти показать «наружу», т.е. что станет новым скрытым состоянием ht. Используется текущее состояние памяти Ct и его фильтрация сигмоидой (значения от 0 до 1).

$$o_t = \sigma(W_{io}x_t + b_{io} + W_{ho}h_{t-1} + b_{ho}) \quad (10)$$

где $x_t \in R^{128}$ – входной вектор на шаге t

$h_{t-1} \in R^{256}$ – скрытое состояние с предыдущего шага,

$W_{io} \in R^{256\times128}$ — матрица весов по входу для выходных ворот,
$W_{io} \in R^{256\times256}$ — матрица весов по скрытому состоянию для выходных ворот,
$b_{io}, b_{ho} \in R^{256}$ — векторы смещений,
$\sigma(\cdot)$ — сигмоидальная функция активации.

Состояние ячейки(C_t) процесс обновления внутреннего состояния памяти C_t в LSTM. Оно сочетает старую память C_{t-1}, умноженную на ворота забывания, и новое предложение $\tilde{C}_t$ умноженное на входные ворота.

$$C_t = f_t \odot C_{t-1} + i_t \odot \tilde{C}_t\, b_{ho}) \tag{11}$$

где $C_{t-1} \in R^{256}$ — предыдущее состояние ячейки,
$f_t \in R^{256}$ — выход ворот забывания,
$i_t \in R^{256}$ — выход входных ворот,
$C_t \in R^{256}$ — кандидатное состояние ячейки,
$\odot$ — обозначает покомпонентное умножение.

Скрытое состояние (h_t) — выход LSTM на текущем шаге, то, что «выдаёт» сеть. Вычисляется на основе обновлённой памяти C_t (после $tanh$) и выходных ворот o_t, решающих, какую её часть показать.

$$h_t = o_t \odot tanh(C_t) \tag{12}$$

где $C_t \in R^{256}$ — текущее состояние ячейки,
$o_t \in R^{256}$ — выход выходных ворот,
$h_t \in R^{256}$ — результирующее скрытое состояние,
$tanh(\cdot)$ — гиперболический тангенс,
$\odot$ —обозначает покомпонентное умножение.

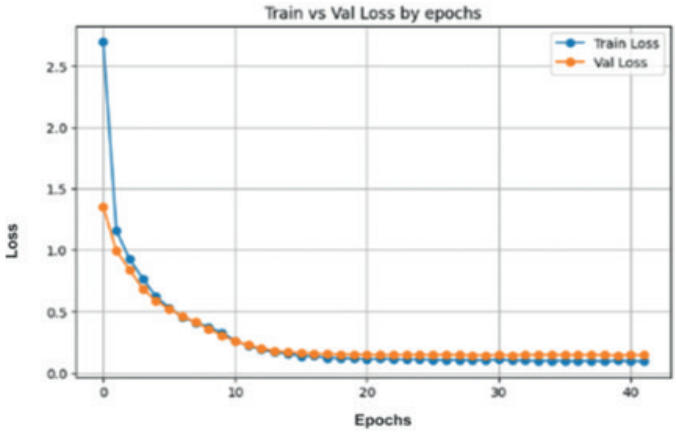


Рисунок 5 –Train vs Validation Loss

Динамика функции потерь при обучении модели Password LSTM на протяжении 43 эпох продемонстрирована на рисунке 8. На начальном этапе кривые обучающей и валидационной потерь демонстрируют резкое снижение, особенно в первые 10 эпох. Модель захватывает базовые структурные зависимости в последовательностях паролей и эффективно уменьшает ошибку предсказания на ранней стадии обучения, что и иллюстрирует быстрый спад.

После ~10-й эпохи скорость снижения потерь замедляется, и примерно к 15-й эпохе обе кривые стабилизируются на значении 0,15. Это плато свидетельствует о достижении сходимости, после чего улучшаются лишь маргинальные величины. Близкое совпадение двух кривых на протяжении всего процесса обучения отражает высокую обобщающую способность и отсутствие переобучения, что особенно важно при работе с посимвольными последовательностями.

В целом динамика обучающей функции потерь подтверждает, что модель была корректно настроена и обучалась на репрезентативном датасете. Согласованное поведение обеих кривых по эпохам демонстрирует устойчивость архитектуры LSTM и этапов препроцессинга, включающих токенизацию символов, нормализацию последовательностей и обогащение признаков. Полученные результаты подтверждают надежность модели в задачах распознавания структур паролей и её готовность к практическому внедрению.

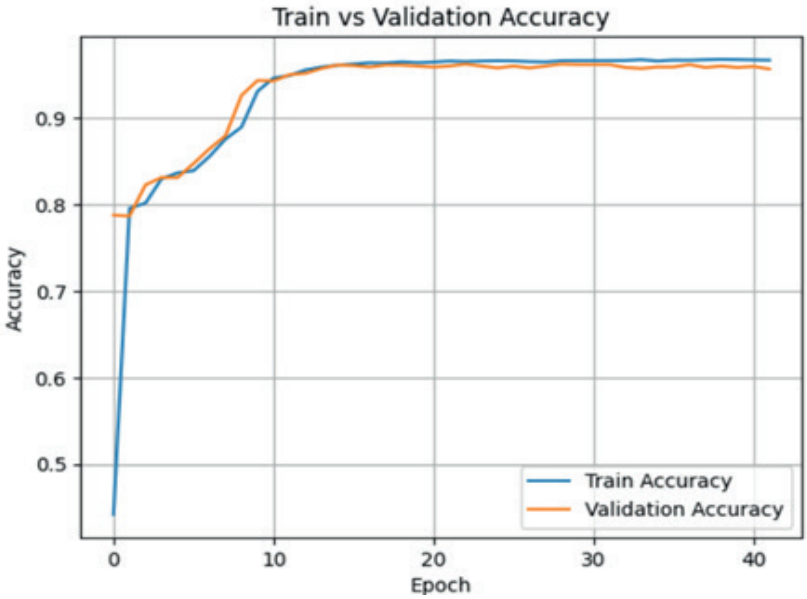


Рисунок 6 –Точность на обучающем и валидационном наборах

Результаты классификации модели Password LSTM как на обучающем, так и на валидационном наборах показаны в рисунке 6. На первых десяти поколениях точность резко увеличивается. Модель переходит от случайного угадывания к стабильно высоким значениям. Несмотря на ограниченный алфавит и сравнительно короткие входы, модель демонстрирует этот рост, быстро извлекая важные паттерны и семантические зависимости из последовательностей паролей.

С пятнадцатого века различия на кривых остаются минимальными, а точность обеих выборок улучшилась до 95%. Это поведение особенно важно для посимвольных моделей последовательностей, потому что оно показывает сохранение качества предсказаний на данных, не использовавшихся при обучении, и отсутствие явных признаков переобучения.

В более поздние времена правильная настройка конвейера предварительной обработки и выбранные механизмы регулирования связаны с высокими значениями точности. Исключение между слоями LSTM и нормализация длины последовательностей обеспечивает устойчивость предсказаний при работе с разнородными паролями. Эти возможности позволяют модели использоваться для задач оценки устойчивости паролей в Интернете и анализа пользовательских шаблонов в средствах безопасности браузеров.

Результаты обучения, показанные на Рисунках 5 и 6, подтверждают надёжность и устойчивость предложенной LSTM-модели. Модель эффективно обнаруживает скрытые закономерности в данных о паролях без переобучения, что подтверждают быстрые сходимости функции потерь и стабильная точность свыше 95% на обучающем и валидационном наборах. Благодаря близкому совпадению кривых модель подходит для задач анализа и прогнозирования структуры паролей. В целом процесс обучения оказался успешным, что привело к созданию надежной и высокопроизводительной архитектуры.

Был проведен двухэтапный анализ, чтобы оценить влияние общедоступных персональных данных на структуру пользовательских паролей. На первом этапе использовался Serper API для сбора текстовых сниппетов с именами, городом, датой рождения и адресом. Ключевые токены и пароли из корпуса были объединены после нормализации текста. Следующим шагом было использование метрик RapidFuzz для оценки точных и частичных совпадений имени/логина, а также семантической близости к OSINT-токенам. Основные характеристики пароля, такие как наличие цифр и символов, были объединены с результатами и представлены в виде корреляционной матрицы.

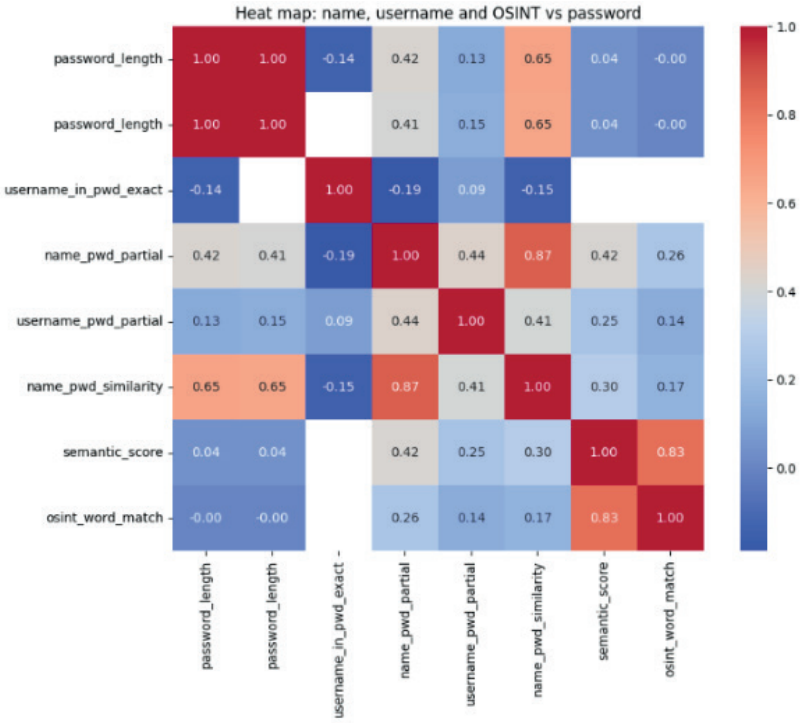


Рисунок 7 –Тепловая карта корреляций между паролями и персональными данными

Проведенный анализ (Рис. 7) показал, что длина пароля положительно коррелирует с его сходством с именем (r = 0,65), тогда как частичные совпадения демонстрируют сильную корреляцию как с полной похожестью (r = 0,87), так и с семантической оценкой (r = 0,42). При этом, если в пароле сохраняются персональные токены, добавление цифр или подстановочных символов лишь незначительно усложняет его предсказуемость. Отсюда следует, что имена и связанные с ними элементы остаются основным источником уязвимости, а усложнение структуры без удаления этих элементов не гарантирует надежную защиту.

После этапа обучения модель была экспортирована в формат ONNX, что позволило интегрировать её в два прикладных инструмента: веб-приложение на Vue 3 и расширение браузера Google Chrome, благодаря этому всё вычисление выполняется локально: без отправки чувствительных данных на удалённые серверы. Исключая серверное хранение или анализ, такая архитектура обеспечивает максимальную конфиденциальность данных: учётные сведения пользователей не собираются, не сохраняются и не обрабатываются на стороне сервера. Интерфейсы работают в режиме реального времени, поэтому пользователь получает мгновенную обратную связь в виде отображения разложения строки на структурные элементы, присвоение уровня безопасности (WEAK, MED, STRONG) и оценку

вероятности на основе выходного распределения модели. Расширение использует Indexed DB для хранения повторяющихся паттернов. Это позволяет отслеживать стандартные шаблоны ввода и адаптировать советы к поведению пользователя. Таким образом, цель реализации модели состоит не только в достижении технической точности, но и в том, насколько она полезна для повседневной кибергигиены.

Опираясь на вероятностное сканирование модели LSTM, система встроена в комплексную платформу анализа паролей. Приложение веб-браузера и расширение Chrome предоставляют обратную связь в режиме реального времени. Веб-приложение загружает экспортированную в ONNX модель и словарь «символ→индекс» при открытии страницы через onnxruntime-web (WASM), что позволяет клиенту выполнять все вычисления инферентности. Асинхронная процедура вызывает функцию предсказания силы для определения двух важных метрик: средней отрицательной лог-вероятности (avg(–log P)) и нормализованного процента «силы» (0–100). Это происходит после того, как пользователь вводит пароль в привязанное поле и нажимает «Проверить». Реактивные переменные Vue.js мгновенно обновляются при помощи этих значений: красная — WEAK (менее 40 %), желтая — MED (40–70 %), зелёная — STRONG (более 70 %), а также выводится среднее значение для продвинутых пользователей (рис. 8).

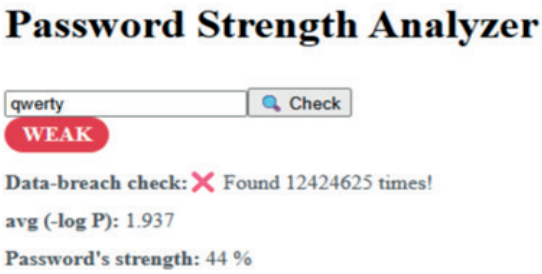


Рисунок 8. Веб-приложение: анализ пароля

После вычисления оценки система сопоставляет пароль с предварительно рассчитанным списком уязвимых подстрок, извлеченных в ходе обучения, помечая совпадения предупреждающими иконками для выделения структурных слабых мест. Но в API Have I Been Pwned отправляется только первые пять символов SHA-1-хеша пароля, таким образом предотвращая возможные атаки повторного использования данных, а возвращенный список суффиксов проходит локальную фильтрацию для отображения количества утечек без вскрытия чувствительных данных. Ниже основной панели динамический компонент Password Chart (Рис. 12) отображает семь наиболее часто повторяющихся паролей, а текстовые списки «Маски, упрощающие угадывание паролей» и «Рекомендуется избегать» (Рис. 13) направляют пользователей вдали от как общих, так и персонализированных паттернов.

Рисунок 9. Веб-приложение: график наиболее повторяемых паролей

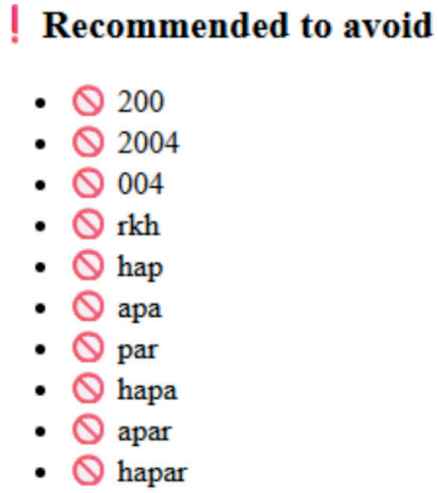


Рисунок 10. Веб-приложение: паттерны, которые рекомендуется избегать

Расширение для Chrome воспроизводит эту логику на каждой странице с полями <input type="password">. Определённое в manifest.json, оно включает всплывающий интерфейс (popup UI), фоновый воркер и контент-скрипт, реализованные на Vue 3, TypeScript и Vite. Во время онбординга пользователь при желании вводит своё имя и город; это инициирует поиск в DuckDuckGo, из которого извлекаются значимые ключевые слова для персонализированных проверок паттернов и сохраняются локально в chrome.storage и IndexedDB. Контент-скрипт добавляет рядом с каждым полем пароля живой бейдж и слушает нажатия клавиш; при каждом вводе выполняются те же ONNX-инференс, сопоставление паттернов и проверка утечек через унифицированную функцию rate (), обновляющие бейдж и встроенную диагностику за миллисекунды. Поддержка «горячей» перезагрузки через сообщения chrome.runtime держит паттерны актуальными без перезагрузки страниц. Все вычисления — исполнение модели, парсинг ключевых слов, проверки паттернов — происходят на устройстве пользователя; из браузера

уходят только неидентифицирующие хеш-префиксы или поисковые запросы, что полностью сохраняет приватность при обеспечении быстрой и информативной обратной связи (Рис. 11).

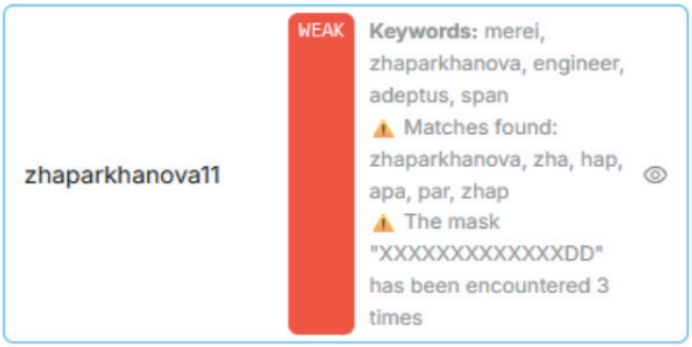


Рисунок 11. Расширение Chrome

Для оценки устойчивости пользовательских паролей к перебору была применена LSTM-модель, обученная на персонализированном датасете. В ходе обучения сеть автоматически выявляла часто повторяющиеся последовательности символов, характерные для конкретного пользователя, и формировала набор «масок» и подстрок, которые статистически повышают предсказуемость паролей данного пользователя на основе его правила формирования паролей и набора наиболее частых слов. Такой персонализированный подход дополнен локальным инференсом (ONNX) и допускает дообучение на конечном устройстве клиента, что исключает вывоз сырых данных и соответствует требованиям корпоративной ИБ.

Сходимость подтверждена динамикой функции потерь и точности: на Рисунке 5 показано, что обучающие и валидационные потери быстро падают в первые 10 эпох и стабилизируются ниже ~0,15, признаков переобучения не наблюдается; Рисунок 6 фиксирует достижение валидационной точности 96,7 % при близких значениях на обучении. Для калибровки «знакомости» структуры использовалась средняя отрицательная лог-вероятность (Avg Log-Score): на валидации получено 0,224 (меньше — значит более предсказуемо). В связке «accuracy + Avg Log-Score» модель не только классифицирует «сильные/слабые» пароли, но и количественно оценивает вероятность структурных и семантических шаблонов.

Согласно исследованию, персональные токены основательно влияют на степень предсказуемости паролей. Например, анализ сформированного моделью набора характерных данных (подстрок) демонстрируют доминацию числовых и именных последовательностей. Это в свою очередь отображает предпочтения пользователей при формировании их паролей, как показано на рисунке 12.

```json
```

```
{
 "merei": 20,
 "erei04": 1,
 "merei04": 1,
 "zha": 49,
 "apa": 96
}```
```

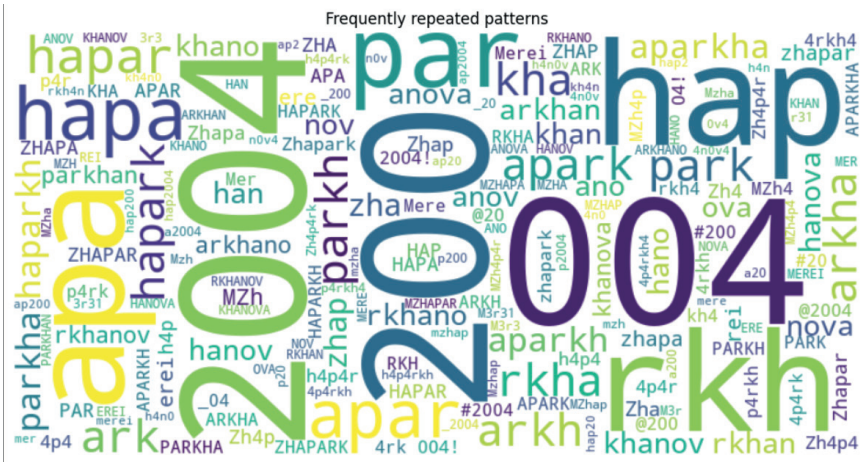


Рисунок 12 – Облако слов по часто повторяющимся паттернам

На рисунке 12 визуализировано какие последовательности преобладают среди числовых последовательностей и токенов, например «2004» и «zhap», «khan», «Merei», что объясняет повышенную предсказуемость даже формально «сложных» паролей с заглавными буквами и спецсимволами. Вместе с этим, проведенный кореляционный анализ OSINT-ключевых слов, извлечённых через Serper API, продемонстрировал выраженную положительную связь между паролями и личными именами ($r \approx 0{,}65$; см. Рисунок 7); где частичные совпадения строк и обнаруженная семантическая близость также увеличивают риск угадывания. Все эти наблюдения поддерживают стратегию персонализированного контроля предсказуемости при формировании паролей, где через контексные подсказки и блокировок определенных персонально клишированных «масок» достигается более высокая непредсказуемость при формирования паролей до их сохранения в системе.

Ниже, в таблице 1 мы провели сравнение полученных результатов с базовыми методами, где модель LSTM показала более высокую точность предсказания лучше оценивала вероятность структуры паролей а также точнее отражает поведенческие особенности испытуемых.

Таблица 1. Сравнение результатов моделей

Модель	Точность	Avg Log-Score	Адаптируемость	Ручное проектирование признаков	Моделирование поведения человека
Decision Tree	81.4%	0.421	Низкая	Требуется	Отсутствует
PCFG	88.2%	0.318	Средняя	Требуется	Частично
LSTM	96.7%	0.224	Высокая	Не требуется	Реализовано

Результаты проведенных опытов подтверждает статистику (All About Cookies, 2023) что большинство пользователей, около 84%, ставят небезопасные пароли опираясь на легко предсказуемые паттерны, такие как: любимые числа (24%), имя питомцев (23%), дата рождения (19%), имена членов семьи или партнеров (16%), любимые произведения (8%), памятные даты (8%) или вообще повторное использование того же пароля в той же системе (17%), если это не запрещено.

Сопоставление с онлайн-проверками также демонстрирует ограниченность эвристик. PasswordMonster (PasswordMonster. (n.d.)) имитирует реальные атаки (словари, подстановки a→@, e→3, последовательности 12345/qwerty) и оценивает время взлома, выступая одновременно диагностическим и образовательным инструментом. Напротив, Kaspersky Password Checker (Kaspersky Lab, 2025) и Security.org (Security.org, 2025) полагаются на простые правила (длина, классы символов), игнорируя структурные/поведенческие паттерны. UIC Password Checker (University of Illinois Chicago. (n.d.)) — правило-ориентированная система с фиксированными критериями. Подобные методы склонны ошибаться и не отражают фактическую стойкость, что подчёркивает потребность в интеллектуальных адаптивных подходах.

Для полноты приведено сравнение классов методов, включая PassGAN в таблице 2(Hitaj, B., et al., 2019), генерирующий «человекоподобные» пароли. Хотя PassGAN силён в массовой генерации кандидатов, ему не хватает интерпретируемости и возможностей обратной связи в реальном времени, необходимых для персонализированных рекомендаций.

Таблица 2. Сравнение различных подходов

Критрии	Decision Tree	PCFG	PassGAN	LSTM
Тип модели	Дискретная логика на деревьях	Статистическая модель на грамматике	Генеративно-состязательная сеть	Рекуррентная нейросеть
Обработка после-довательностей	Не поддерживается	Ограничена заданной грамматикой	Полная поддержка (генеративная выборка)	Полная поддержка (контекстное моделир.)
Требуется ручное проектирование признаков	Да	Да	Нет	Нет

Гибкость для новых паттернов	Низкая	Средняя	Высокая	Высокая
Учёт человеческого фактора	Практически не учитывается	Частичный учёт через вероятности	Косвенно, по распределению данных	Присутствует, моделирует реальное поведение
Устойчивость к нетипичным данным	Низкая	Средняя	Высокая	Высокая
Интерпретируемость	Высокая	Средняя	Очень низкая (чёрный ящик-генератор)	Низкая (чёрный ящик)
Представление пароля	Дерево логических правил	Структура + частоты сегментов	Латентное векторное кодирование	Эмбеддинги и память
Основное ограничение	Требует ручного отбора признаков	Не выявляет вложенные паттерны	Недостаток интерпретируемости и персональной обратной связи	Требует больших данных и времени обучения

Для задач цифровой криминалистики в настоящее время представлена модульная система семантического анализа сообщений, которая комбинирует векторный поиск, классификацию и визуальную аналитику (FastAPI, Qdrant, Llama-Guard-3-1B, React). Платформа сосредоточена на семантической интерпретации и категоризации, отличаясь от традиционных решений. Время построения индекса на корпусе из 7448 сообщений составило 3 с; многоязычные запросы с динамической маршрутизацией на перевод обрабатываются в среднем за 5,23 с, а запросы без перевода — 0,45 с. Классификатор Llama-Guard-3-1B достиг точности 75,19 % при пакетной обработке; типичные ошибки, включая смешение близких категорий и ложные негативы на трех «тонких» кейсах, показывают, что доменная адаптация и балансировка классов являются необходимыми. Платформа облегчает поиск узлов социальных графов, временных аномалий и критичных категориальных паттернов, визуализируя структурные и поведенческие характеристики коммуникаций (по приложениям, получателям, геолокациям, динамике и статусам).

Заключение

Предложенное решение является отечественным альтернативным решением проверки качества паролей пользователей и при интеграции в корпоративную сеть может являться продуктивным инстру

В совокупности результаты показывают: (i) в парольном сценарии LSTM-подход переводит контроль качества пароля от формальных правил к персонализированному on-device анализу, снижая предсказуемость и повышая реальную стойкость; (ii) инфраструктура семантической аналитики

пригодна для экспресс-разведки больших многоязычных массивов текста в криминалистике при приемлемой латентности. Перспективные направления включают расширение персональных словарей (в т.ч. семантическая близость), автоматическую генерацию «анти-масок», интеграцию с менеджерами паролей и корпоративными SSO; мультимодальный OSINT (LinkedIn, X/Twitter, Instagram) с агентами CV/NLP и связкой сущностей для более точного профилирования; а также лонгитюдное исследование, оценивающее, от каких небезопасных привычек (персональные данные, популярные паттерны, минимальные правки старых паролей) пользователи отказываются под воздействием контекстной обратной связи. Дополнительно для криминалистического модуля планируются расширение обучающих корпусов, hard-negative mining, добавление аудио-/видеоаналитики и интеграция с платформами OSINT/Threat Intelligence.

Таким образом, предложенная система сочетает персонализированную оценку стойкости паролей в реальном времени и масштабируемую семантическую аналитику текстов, закрывая разрыв между «безопасностью по правилам» и безопасностью на практике.

Благодарности

Данное исследование проведено при финансовой поддержке Комитета науки Министерства науки и высшего образования Республики Казахстан в рамках договора №388/ПЦФ-24-26 от 01.10.2024 по научному проекту BR24993232 «Разработка инновационных технологий проведения цифровых криминалистических исследований с применением интеллектуальных программно-аппаратных комплексов».

Этическое одобрение

Данное исследование не затрагивает участие людей или животных.

References

Bergeron A., & Dearden T.E. (2024) Secret sharing in online communities: A comparative analysis of offender and non-offender password creation strategies. Journal of Economic Criminology, 6, 100110. *https://doi.org/10.1016/j.jeconc.2024.100110* (in English)

Atzori M., Calò E., Caruccio L., Cirillo S., Polese G., & Solimando G. (2024) Evaluating password strength based on information spread on social networks: A combined approach relying on data reconstruction and generative models. Online Social Networks and Media, 42, 100278. *https://doi.org/10.1016/j.osnem.2024.100278* (in English)

He D., Yu H., Zhou B., Zhu S., Zhang M., Chan S., & Guizani M. (2021) How does social behavior affect your password? IEEE Network, 35(5). — P. 284–289. https://doi.org/10.1109/mnet.101.2000762 (in English)

Buckman B. (n.d.-b) 36 Must-Know Password Statistics for 2025. Huntress. https://www.huntress.com/blog/password-statistics? (in English)

Data breach statistics in 2024 - Surfshark (2020, December 9) Surfshark. *https://surfshark.com/research/study/data-breach-recap-2024* (in English)

Urrico R. (2024, February 13) Reports show rising ransomware attacks and bad password habits threaten financial accounts, among others. Finopotamus. https://www.finopotamus.com/post/reports-show-rising-ransomware-attacks-and-bad-password-habits-threaten-financial-accounts-among-ot (in English)

Password dataset (2024, April 25) Kaggle. https://www.kaggle.com/datasets/soylevbeytullah/password-datas 7 (in English)

Kkrypt0nn. (n.d.-b) wordlists/wordlists/passwords at main kkrypt0nn/wordlists. GitHub. https://github.com/kkrypt0nn/wordlists/tree/main/wordlists/passwords 8 (in English)

Melicher W., Ur B., Segreti S.M., Komanduri S., Shay R., Bauer L., Christin N., & Cranor L.F. (2016) Fast, lean, and accurate: Modeling password guessability using neural networks. In 25th USENIX Security Symposium. — P. 175–191. USENIX Association. https://www.usenix.org/conference/usenixsecurity16/technical-sessions/presentation/melicher (in English)

All About Cookies. (2023, October 12). 84% of people use unsafe passwords: Password behavior survey. https://allaboutcookies.org/password-users-behavior-survey (in English)

Pass word Monster (n.d.) Password Strength Meter. https://www.passwordmonster.com/ (in English)

Kaspersky Lab. (2025) Password Checker & Secure Random Password Generator. https://password.kaspersky.com/ (in English)

Security.org. (2025) How Secure Is My Password?. Password Strength Checker. https://www.security.org/how-secure-is-my-password (in English)

University of Illinois Chicago. (n.d.) Password Strength Test. https://www.uic.edu/apps/strong-password/(in English)

Hitaj B., Gasti P., Ateniese G., & Perez-Cabo A. (2019) PassGAN: A deep learning approach for password guessing. In A. Biggio & F. Roli (Eds.), Advances in information security. — Vol. 79. — P. 1–20. Springer. https://doi.org/10.1007/978-3-030-29959-5_1 (in English)

Rzayeva L., Pogolovkin D., & Myrzatay A. (2025) DEVELOPMENT OF A MODULAR NLP-BASED CORRESPONDENCE ANALYSIS SERVICE FOR DIGITAL FORENSICS. News of the National Academy of Sciences of the Republic of Kazakhstan. Physico-Mathematical Series, (2). — P. 212–233. https://doi.org/10.32014/2025.2518-1726.354 (in English)

Idrissova M., Kim S., Amirgaliyev B., Yedilkhan D., & Rzayeva L. (2025) DIGITAL FOOTPRINTS: CLUSTERING BROWSER HISTORY FOR USER PROFILING USING MACHINE LEARNING. Journal of Problems in Computer Science and Information Technologies, 3(2). — P. 16-28. https://doi.org/10.3390/info16080655 (in English)

ACADEMIC SCIENTIFIC JOURNAL OF COMPUTER SCIENCE
ISSN 1991-346X
Volume 4. Namber 356 (2025). 259–273

https://doi.org/10.32014/2025.2518-1726.395

УДК 520.6.05

© D. Sagidoldin[1], A. Zhetpisbayeva[2]*, B. Zhumazhanov[1], B. Zhumazhanov[1], 2025.
[1]LLP "Ghalam", Astana, Kazakhstan;
[2]L.N. Gumilyov Eurasian National University, Astana, Kazakhstan.
E-mail: aigulji@mail.ru

INCREASING THE RELIABILITY OF DATA TRANSMISSION FROM SMALL SPACECRAFT USING SDR EQUIPMENT

Sagidoldin Daulet — Design Engineer, Ghalam» LLP, Astana, Kazakhstan,
E-mail:d.sagidoldin@ghalam.kz, https://orcid.org/0009-0008-0495-8247;
Zhetpisbayeva Ainur — PhD, associate professor of the department of "Radio engineering, electronics, and telecommunications," L.N. Gumilyov Eurasian national university, Astana, Kazakhstan,
E-mail:aigulji@mail.ru, https://orcid.org/0000-0002-4525-5299;
Zhumazhanov Berik — Head of department of payload and scientific developments, «Ghalam» LLP, Astana, Kazakhstan,
E-mail:b.zhumazhanov@ghalam.kz, http://orcid.org/0000-0001-5926- 9619;
Zhumazhanov Beksultan — Design-engineer, «Ghalam» LLP, Astana, Kazakhstan,
E-mail:zhumazhanov.b@ghalam.kz, https://orcid.org/0009-0000-9493-7491.

Abstract. This paper addresses the problem of increasing the reliability of the radio link between small spacecraft and ground stations under stringent on-board power and resource constraints. The transmitting side is based on a software-defined radio chain using the AD9364 transceiver in combination with the HMC998 power amplifier, which provides the required energy potential of the communication channel. The main focus is on studying error-correcting coding methods and data rate control. Numerical simulations are carried out for an additive white Gaussian noise channel, taking into account distances of up to 2000 km, antenna parameters, receiver G/T, and additional link losses.The analysis covers BPSK, QPSK, 8-PSK and GMSK modulations combined with LDPC, turbo and Reed–Solomon codes; BER dependences on Eb/N0, operating frequency, elevation angle and data rate are compared. It is shown that the use of LDPC coding provides an Eb/N0 gain of several decibels and makes it possible to achieve BER on the order of 10^{-9} in the 2.2–2.4 GHz band at data rates up to 10 Mbit/s, which significantly increases link reliability for small-satellite orbits. In addition, principles of adaptive selection of the "modulation – code – rate" combination are considered, depending on the current energy margin of the channel, which effectively implements data rate control at the SDR level. The obtained results can be used in the design and optimization

of telemetry and service communication channels of nanosatellites, as well as in the development and testing of SDR equipment on ground-based test benches and flight models of small spacecraft.

Keywords. SDR, small spacecraft, Earth stations, radio frequency, reliability enhancement, coding, modulation, satellite communications

© Д.Т. Сагидолдин[1], А.Т. Жетписбаева[2]*, Б.Р. Жумажанов[1], Б.С. Жумажанов[1], 2025.

[1] ЖШС «Ghalam», Астана, Қазақстан;

[2] Л.Н. Гумилев атындағы Еуразия ұлттық университеті, Астана, Қазақстан.

E-mail: aigulji@mail.ru

SDR ЖАБДЫҚТАРЫН ПАЙДАЛАНУ АРҚЫЛЫ, ШАҒЫН ҒАРЫШ АППАРАТТАРЫНАН ДЕРЕКТЕРДІ БЕРУДІҢ СЕНІМДІЛІГІН АРТТЫРУ

Сагидолдин Даулет — инженер-жобалаушы, ЖШС «Ghalam», Астана, Қазақстан,
E-mail:d.sagidoldin@ghalam.kz, https://orcid.org/0009-0008-0495-8247;
Жетписбаева Айнур — PhD, қауымдастырылған профессор, «Радиотехника, электроника және телекоммуникациялар» кафедрасы, Л.Н. Гумилев атындағы Еуразия ұлттық университеті, Астана, Қазақстан,
E-mail:aigulji@mail.ru, https://orcid.org/0000-0002-4525-5299;
Жумажанов Берик — Пайдалы жүктеме және ғылыми әзірлемелер бөлімінің басшысы, ЖШС «Ghalam», Астана, Қазақстан,
E-mail:b.zhumazhanov@ghalam.kz, http://orcid.org/0000-0001-5926-9619;
Жумажанов Бексултан — инженер-жобалаушы, ЖШС «Ghalam», Астана, Қазақстан,
E-mail:zhumazhanov.b@ghalam.kz, https://orcid.org/0009-0000-9493-7491.

Аннотация. Бұл мақалада «шағын ғарыш аппараты – жердегі станция» радиобайланыс арнасының сенімділігін борттық қуат пен ресурстарға қатаң шектеулер жағдайында арттыру мәселесі қарастырылады. Таратқыш бөлігі ретінде AD9364 трансиверіне негізделген бағдарламалық-анықталатын радиотракт (SDR) және байланыс арнасының қажетті энергетикалық потенциалын қамтамасыз ететін HMC998 қуат күшейткіші қолданылады. Негізгі назар кедергіге төзімді (қателерді түзететін) кодтау әдістері мен деректерді беру жылдамдығын басқару тәсілдерін зерттеуге аударылған. Аддитивті ақ Гаусстік шуы бар (AWGN) арнада 2000 км-ге дейінгі қашықтық, антенналардың параметрлері, қабылдағыштың G/T көрсеткіші және трактінің қосымша шығындары ескеріле отырып сандық модельдеу жүргізілді. BPSK, QPSK, 8-PSK және GMSK модуляциялары LDPC, Turbo және Reed–Solomon кодтарымен үйлесімде талданады; қателік ықтималдығының (BER) Eb/N0 қатынасына, жұмыс жиілігіне, элевация бұрышына және беру жылдамдығына тәуелділігі салыстырылады. LDPC-кодтауды қолдану Eb/N0 бойынша бірнеше децибелге дейін ұтыс беріп, 2,2–2,4 ГГц жиілік диапазонында 10 Мбит/с-қа дейінгі жылдамдықтарда BER шамамен 10^{-9} деңгейіне қол жеткізуге

мүмкіндік беретіні көрсетілді, бұл шағын КА орбиталарында арна сенімділігін едәуір арттырады. Сонымен қатар арнаның ағымдағы энергетикалық қоры есебінен «модуляция – код – жылдамдық» комбинациясын бейімделмелі таңдау қағидаттары қарастырылып, деректер беру жылдамдығын SDR деңгейінде басқару іске асырылады. Алынған нәтижелерді наноспутниктердің телеметриялық және сервистік байланыс арналарын жобалау мен оңтайландыруда, сондай-ақ жерүсті сынақ стендтерінде және шағын ғарыш аппараттарының ұшу үлгілерінде SDR жабдығын әзірлеу және сынау кезінде пайдалануға болады.

Түйін сөздер. SDR, шағын ғарыш аппараттары, жер станциялары, радиожиілік, сенімділікті арттыру, кодтау, модуляция, спутниктік байланыс

© Д.Т. Сагидолдин[1], А.Т. Жетписбаева[2]*, Б.Р. Жумажанов[1], Б.С. Жумажанов[1], 2025.

[1] ТОО «Ghalam», Астана, Казахстан;

[2] Евразийский национальный университет имени Л.Н. Гумилева, Астана, Казахстан.

E-mail: aigulji@mail.ru

ПОВЫШЕНИЕ НАДЁЖНОСТИ ПЕРЕДАЧИ ДАННЫХ С МАЛЫХ КОСМИЧЕСКИХ АППАРАТОВ С ПРИМЕНЕНИЕМ SDR ОБОРУДОВАНИЯ

Сагидолдин Даулет — Инженер-конструктор, ТОО«Ghalam», Астана, Казахстан,
E-mail: d.sagidoldin@ghalam.kz, https://orcid.org/0009-0008-0495-8247;
Жетписбаева Айнур – PhD, ассоциированный профессор, кафедра «Радиотехника, электроника и телекоммуникациялар», Евразийский национальный университет имени Л.Н. Гумилева, Астана, Казахстан,
E-mail: aigulji@mail.ru, https://orcid.org/0000-0002-4525-5299;
Жумажанов Берик — Начальник отдела полезной нагрузки и научных разработок, ТОО «Ghalam», Астана, Казахстан,
E-mail: b.zhumazhanov@ghalam.kz, http://orcid.org/0000-0001- 5926-9619;
Жумажанов Бексултан — Инженер-конструктор, ТОО«Ghalam», Астана, Казахстан,
E-mail: zhumazhanov.b@ghalam.kz, https://orcid.org/0009-0000-9493-7491.

Аннотация. В статье рассматривается повышение надежности радиолинии «малый космический аппарат – наземная станция» при жестких ограничениях по мощности и ресурсам борта. В качестве основы передающей части используется программно-определяемый радиотракт на базе трансивера AD9364 в сочетании с усилителем мощности HMC998, обеспечивающим требуемый энергетический потенциал канала связи. Основное внимание уделено исследованию методов помехозащищенного кодирования и управления скоростью передачи данных. Выполнено численное моделирование в канале с аддитивным белым гауссовским шумом с учетом расстояния до 2000 км, параметров антенн, показателя G/T приемника и дополнительных потерь

тракта. Анализируются модуляции BPSK, QPSK, 8-PSK и GMSK в сочетании с кодами LDPC, Turbo и Reed–Solomon; сравниваются зависимости BER от отношения Eb/N0, рабочей частоты, угла элевации и скорости передачи. Показано, что применение LDPC-кодирования обеспечивает выигрыш по Eb/N0 до нескольких децибел и позволяет достигать BER порядка 10^{-9} в диапазоне частот около 2,2–2,4 ГГц при скоростях до 10 Мбит/с, что существенно повышает надежность канала при орбитах малых КА. Дополнительно рассматриваются принципы адаптивного выбора комбинации «модуляция – код – скорость» в зависимости от текущего запаса по энергетике канала, что фактически реализует управление скоростью передачи на уровне SDR. Полученные результаты могут быть использованы при проектировании и оптимизации каналов телеметрии и служебной связи наноспутников, а также при разработке и испытаниях SDR-аппаратуры на наземных стендах и летных образцах малых космических аппаратов.

Ключевые слова. SDR, малые космические аппараты, земные станции, радиочастота, повышение надежности, кодирование, модуляция, спутниковая связь

Благодарность. *Данное исследование финансировалось Комитетом науки Министерства науки и высшего образования Республики Казахстан (BR27198365 «Разработка оптикоэлектронной системы в коротковолновой инфракрасной области спектра в контексте развития космических систем ДЗЗ Казахстана»).*

Введение. Спутниковые платформы представляют собой инновационную парадигму, предлагая экономически эффективные и высокоэффективные технологические решения с уменьшенной задержкой. Они обеспечивают покрытие больших зон обслуживания, включая регионы, где наземная сотовая инфраструктура недоступна, такие как слаборазвитые страны, сельские районы, северный и южный полюса. Эти особенности также делают возможными новые типы космических полетов, учитывая низкую стоимость готовых коммерческих электронных компонентов, способных выполнять множество различных задач, таких как защита окружающей среды (Mukherjee et al, 2013; Davies et al, 2019).

В данной работе рассматривается подход, направленный на повышение надежности связи за счет применения SDR на базе трансивера AD9364 в сочетании с усилителем мощности HMC998. Для оценки эффективности предложенного решения проведено моделирование различных схем модуляции и методов помехоустойчивого кодирования.

Цель исследования – определить оптимальные комбинации параметров передачи данных, которые обеспечивают минимальную вероятность ошибок и высокую устойчивость к внешним помехам при передаче данных с МКА.

SDR – это гибкая технология, которая позволяет проектировать адаптивную систему связи. Соответственно, общая конструкция оборудования может использоваться для удовлетворения различных потребностей в связи с различными частотами, схемами модуляции и скоростями передачи данных (Maheshwarappa et al, 2021). Процесс внедрения радиосвязи включает в себя настройку параметров фильтрации, таких как частоты пропускания и заграждения, а также цифровые квадратурные преобразования и корректировки скорости передачи данных с использованием процессов дискретизации вверх и вниз.

В этих сценариях последние достижения в области малых спутников и приемопередатчиков с миниатюрными спутниковыми антеннами делают их хорошим решением, так как использование спутников на низкой околоземной орбите (LEO), размещенных в составе группировки или кластеров, может обеспечить повсеместное покрытие, гарантируя высокую скорость передачи данных и низкие уровни задержки (Bacco et al, 2019)

Учитывая растущее число запусков малых спутников, группировки малых спутников с межспутниковой связью (ISLS) привлекли большое внимание с точки зрения услуг наблюдения и связи. Например, в попытке усилить это преимущество, было разработано многофункциональное пространство на базе программно-определяемой радиосвязи (SDR). link (GAMALINK) появился как коммуникационное решение для создания ISL, ориентированное на небольшие спутниковые платформы, в качестве альтернативы крупным и дорогостоящим космическим миссиям (Rodrigues et al, 2013).

Программируемая радиосвязь (SDR) стала одним из новых методов, уменьшающих многофакторность оборудования и, кроме того, изменяющих маршрут для традиционных систем дистанционной связи. Проектирование многонастраиваемой системы на основе ПЛИС делает её адаптивной и многоразовой. В этой задаче представлен план сегмента обработки основной полосы частот программно-определяемого радио с использованием QPSK, BPSK и план кодирования с использованием кода Хэмминга (Vishnuvardhan G.et al. 2019).

Авторы данной статьи разрабатывают математическую модель для демонстрации надежности небольшой спутниковой сети в отношении структур SDR, мощности передаваемого сигнала на восходящих и нисходящих каналах, кодовых скоростей и коллизий пакетов с помощью усовершенствованного протокола RA. Модель обеспечивает количественную надежность сети в отношении частоты отказов системы SDR, допустимых параметров связи и коэффициента потери пакетов (PLR). Наш анализ предлагает методологию оценки различий в надежности сети в зависимости от изменений параметров связи и рекомендации по поддержанию надежности сетевой системы с соответствующими значениями параметров (Seunghwa et al, 2021; Bacco et al, 2019).

Двоичная фазовая манипуляция (BPSK) и квадратурная фазовая манипуляция (QPSK) являются важными методами модуляции, используемыми во многих стандартах связи. В данной работе разработан приёмопередатчик реального времени для BPSK и QPSK с использованием программно-определяемой радиосистемы (SDR). Приёмопередатчик разработан программно в LabVIEW, а в качестве аппаратных средств используются два универсальных программных радиопериферийных устройства 2921 (USRP 2921) (Mehmet et al, 2018).

В данной статье описывается разработка и аппаратная реализация системы связи типа SDR на основе ПЛИС для наноспутника. Основные функции, выполняемые на борту, — это демодуляция FM/FSK для восходящего канала телекоманд и модуляция BPSK с фильтрацией приподнятого косинуса для нисходящего канала телеметрии. Вся система спроектирована и реализована на основе ПЛИС Microsemi Smartfusion2. Для аппаратной оценки системы используется ПЛИС Virtex-6 с высокоскоростной аналоговой дочерней платой (Nivin et al, 2016)

Обеспечение надежной передачи данных от малых космических аппаратов (МКА) к наземным станциям является важной задачей в современных спутниковых проектах. Ограниченные ресурсы МКА, такие как малая мощность передатчика и размеры антенн, предъявляют особые требования к системам радиосвязи. Программно-определяемая радиосвязь (SDR) позволяет гибко настраивать параметры передачи данных, выбирать подходящие схемы модуляции и адаптировать систему к текущим условиям радиоканала. Однако стандартные SDR-решения не всегда обеспечивают достаточную мощность сигнала для уверенного приема на больших расстояниях.

МКА (SmallSats или малые спутники) включают космические аппараты массой менее 180 кг (https://www.nasa.gov/what-are-smallsats-and-cubesats/). CubeSats подпадают под определение МКА. Стандарты предписывают, что они весят максимум 1,33 кг и имеют размеры 1 см × 1 см × 1 см.

При увеличении скорости передачи данных требуется либо повышение выходной мощности, либо снижение уровня шума, либо сокращение дальности связи. Современные подсистемы спутниковой связи отдают предпочтение адаптивным системам, поскольку в неадаптивных решениях используется, как правило, минимальная скорость передачи для обеспечения устойчивой работы радиоканала.

Традиционно SDR применяются для снижения затрат и сроков разработки систем связи (Heuberger et al, 2009). Современные исследования направлены на повышение их производительности и подготовку к интеграции искусственного интеллекта (ИИ) в спутниковые системы. Адаптивные решения позволяют настраивать усилитель спутника в реальном времени, чтобы соответствовать ограничениям плотности мощности и обеспечивать достаточный уровень сигнала на приеме. Это дает возможность эффективно

использовать ресурсы спутника и передавать максимальный объем данных в зависимости от условий радиоканала.

Таким образом, новизна исследовательской работы заключается в том, что проект оптимизирует производительность системы SDR для малых спутников, и он использует инструменты, доступные большинству разработчиков спутников.

(Yasir et al, 2021) исследование способствует улучшению систем на основе SDR путем разработки адаптивного передатчика и приемника пакетной связи, которые могут использовать окно связи CubeSats и малых спутников. В соответствии с обратной связью от приемника передатчик изменяет характеристики сигнала. Теоретически система может принимать много режимов, но для простоты и для подтверждения концепции здесь изменения ограничены тремя скоростями передачи данных схемы модуляции гауссовой минимальной манипуляции сдвигом (GMSK), т. е. 2400 бит/с GMSK, 4800 бит/с GMSK и 9600 бит/с GMSK, которые являются наиболее популярными в любительских малых спутниках.

Материалы и методы. В работе реализована система радиосвязи для МКА на базе SDR. В качестве основных компонентов использовались трансивер AD9364, обеспечивающий работу в диапазоне частот от 70 МГц до 6 ГГц с возможностью гибкой настройки параметров передачи данных, и усилитель мощности HMC998, способный повышать выходную мощность до +31 дБм для увеличения радиуса устойчивой радиосвязи спутника с наземной станцией. Для анализа характеристик системы проводилось компьютерное моделирование с учетом условий: расстояние между спутником и наземной станцией от 600 до 2000 км при угле элевации от 10° до 40°, мощность передатчика 30 дБм, усиление антенны 8 дБи, показатель шумовой температуры приемного тракта G/T на уровне 12 дБ/К и дополнительные потери в тракте передачи около 4,5 дБ. В ходе моделирования оценивалась эффективность различных схем модуляции (BPSK, QPSK, 8-PSK и GMSK), а также методов помехоустойчивого кодирования (LDPC, Turbo-коды и Reed–Solomon). Основным показателем, использованным для анализа, была вероятность битовой ошибки BER, рассчитываемая при изменении параметров радиоканала, таких как отношение С/Ш, частота несущей и скорость передачи данных. Все расчеты выполнялись с использованием программного обеспечения MATLAB/Simulink и специализированных скриптов на Python, что обеспечивало высокую точность моделирования и возможность объективной оценки параметров связи для разработки надежных систем передачи данных с МКА.

Результаты и обсуждение. Использование AD9364 позволяет гибко выбирать тип модуляции и скорость в зависимости от текущих условий радиоканала. Интеграция с HMC998 позволяет компенсировать потери при дальнем распространении и увеличить отношение С/Ш. Адаптивная

настройка SDR-модуля под особенности канала позволяет сохранить низкий уровень ошибок даже на высоких частотах.

В отличие от фиксированных радиосистем, SDR-архитектура с HMC998:

- увеличивает радиус действия;
- поддерживает адаптивную модуляцию;
- снижает требования к наземной станции;
- позволяет реализовать ARQ и FEC (например, LDPC).

Формулы для оценки вероятности BER в зависимости от отношения E_b/N_0приведены в работах (Ladvánszky et al, 2021; Álvarez et al, 2024) имеют вид:

Для BPSK:

$$P_b = Q\left(\sqrt{2 \cdot \frac{E_b}{N_0}}\right). \quad (1)$$

Для QPSK:

$$P_b \approx Q\left(\sqrt{2 \cdot \frac{E_b}{N_0}}\right). \quad (2)$$

Для 8-PSK:

$$P_b \approx 2Q\left(\sqrt{2 \cdot \frac{E_b}{N_0}} \cdot sin^2(\pi/8)\right). \quad (3)$$

Функция Q:

$$Q(x) = \frac{1}{\sqrt{2\pi}} \int_x^{\infty} e^{-t^2/2} dt \quad (4)$$

На рисунке 1 представлена зависимость вероятности битовой ошибки от частоты при различных схемах модуляции на расстоянии 2000 км.

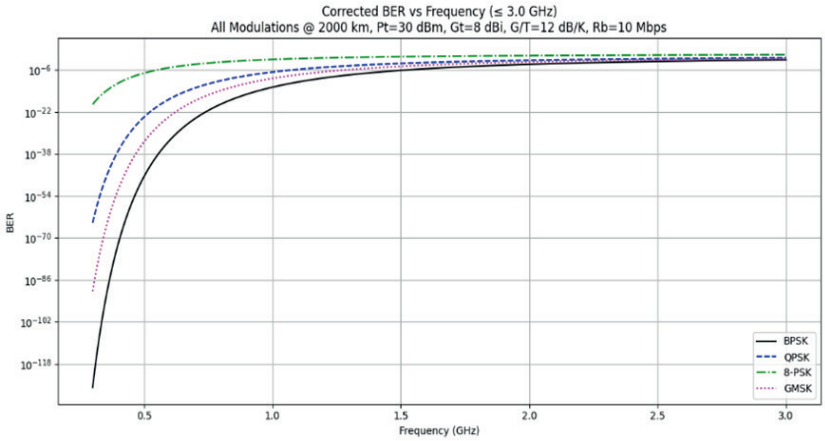


Рисунок 1 – BER в зависимости от частоты (до 3.0 ГГц) для различных модуляций при расстоянии 2000 км.

График показывает, что BPSK обеспечивает наилучшую устойчивость к искажениям. GMSK и QPSK демонстрируют похожую производительность, а 8-PSK – наихудшую на низких частотах. При частотах выше 2 ГГц BER стабилизируется для всех модуляций. Далее на рисунке 2 показана зависимость BER от угла элевации при частоте 2,2 ГГц и орбите 600 км.

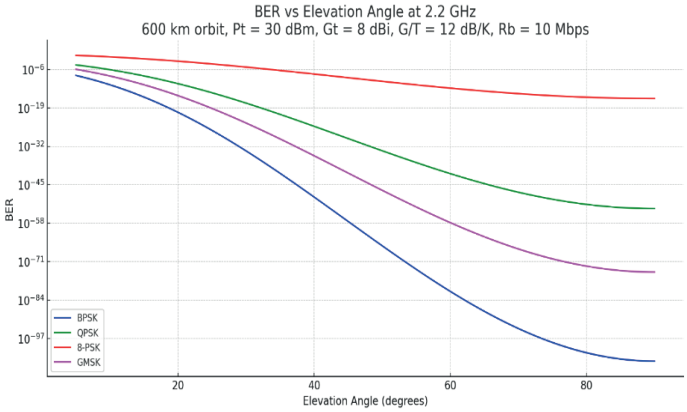


Рисунок 2 – BER от угла элевации при частоте 2.2 ГГц и орбите 600 км

BER уменьшается с ростом угла элевации, так как путь сигнала становится короче и влияние атмосферы снижается. Особенно ярко этот эффект проявляется при углах менее 40°.

Как показано в таблице 1, для каждой схемы модуляции определена частота, при которой BER достигает уровня примерно 10^{-5} при расстоянии 2000 км.

Таблица 1 – Частоты, при которых BER ≈ 10^{-5} для различных модуляций при расстоянии 2000 км

Модуляция	Частота (ГГц)	Модуляция	Частота (ГГц)
BPSK	2,17 ГГц	8-PSK	1,03 ГГц
QPSK	1,52 ГГц	GMSK	2,11 ГГц

Эффект кодирования. Применение каналов с коррекцией ошибок, особенно LDPC и Turbo-кодов, значительно снижает BER и позволяет использовать более высокие частоты без ухудшения качества приема. LDPC-коды показали наибольшую эффективность и позволили достичь BER порядка 10^{-9} при частотах выше 2 ГГц.

Как показано в таблице 2, применение LDPC-кодов обеспечивает наилучшее снижение BER по сравнению с Turbo-кодами и кодами Рида–Соломона для всех исследованных схем модуляции.

Таблица 2. Сравнение эффективности методов помехоустойчивого кодирования для различных видов модуляции

Модуляция	LDPC	Turbo	Reed–Solomon
BPSK	6,0	5,0	2,5
QPSK	5,5	4,5	2,5
8-PSK	4,5	4,0	2,0
GMSK	5,0	4,5	2,5

Данные в таблицах и графиках получены в результате численного моделирования канала связи с учетом расстояния, параметров антенны, мощности передатчика и показателя G/T приемника. Расчеты базируются на известных моделях из работ Proakis, Haykin и Anderson и выполнены с помощью MATLAB/Simulink и Python-скриптов моделирования радиоканала.

На рисунках 3–5 представлены зависимости BER от частоты для схем модуляции BPSK, QPSK и 8-PSK при использовании различных методов помехоустойчивого кодирования (LDPC, Turbo и Reed–Solomon) на расстоянии 2000 км.

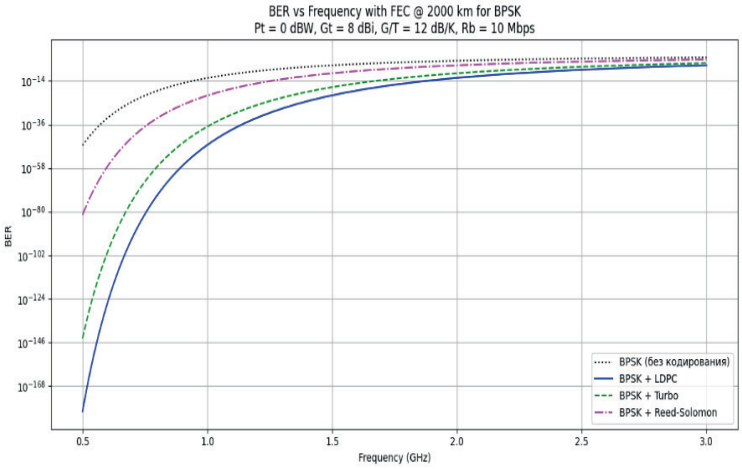


Рисунок 3 – BER от частоты при кодировании BPSK (сравнение LDPC, Turbo и Reed–Solomon).

BPSK с LDPC-кодированием достигает экстремально низких значений BER даже при частотах ниже 1 ГГц. Turbo-коды показывают близкие результаты, а Reed–Solomon уступает по эффективности. Без кодирования уровень BER превышает 10^{-14} практически на всех частотах, демонстрируя важность применения помехоустойчивого кодирования.

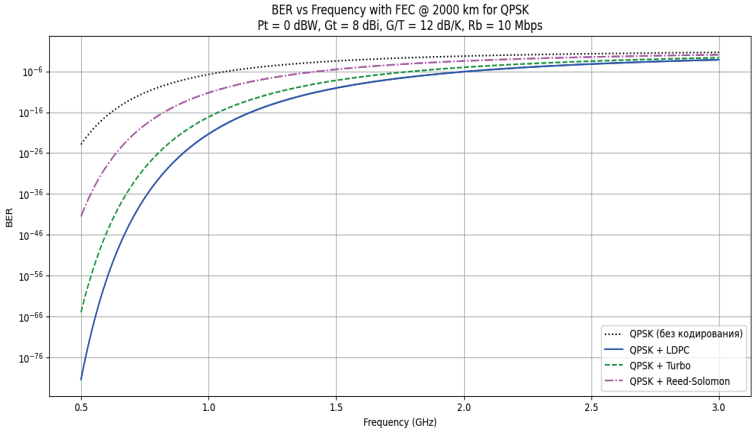


Рисунок 4 – BER от частоты при кодировании QPSK (сравнение LDPC, Turbo и Reed–Solomon)

LDPC-коды демонстрируют наилучшую эффективность, снижая BER на несколько порядков по сравнению с некодированной QPSK. Turbo-коды немного уступают LDPC, а Reed–Solomon – наименее эффективен в данном случае.

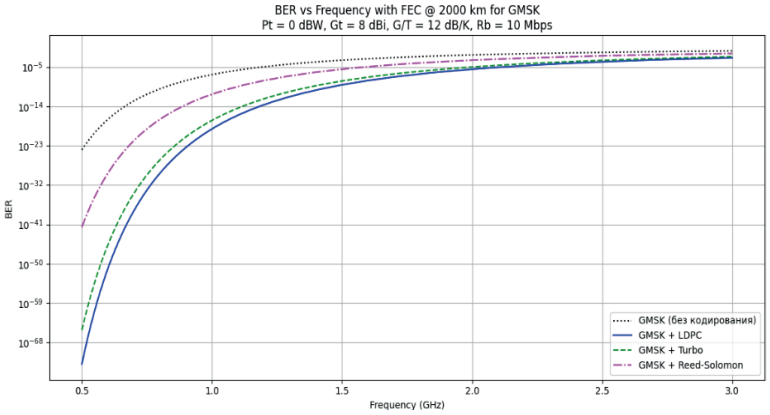


Рисунок 4 – BER от частоты при кодировании 8-PSK (сравнение LDPC, Turbo и Reed–Solomon).

LDPC-коды позволяют достичь BER < 10^{-9} уже при частотах около 1,2 ГГц, тогда как Turbo и Reed–Solomon требуют более высоких частот. Без кодирования BER превышает 10^{-3} на всех частотах.

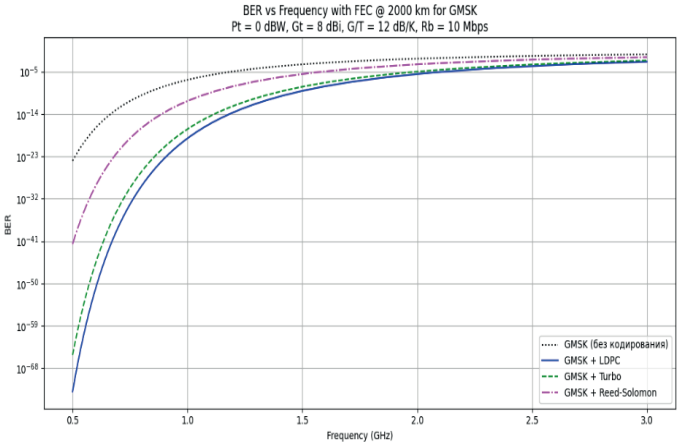


Рисунок 5 – BER от частоты при кодировании GMSK (сравнение LDPC, Turbo и Reed–Solomon).

Как и для других модуляций, LDPC-кодирование наилучшим образом снижает BER. Turbo-коды показывают чуть худшие значения, но остаются эффективными. Reed–Solomon уступает по эффективности и требует большей частоты.

Частотные границы для достижения BER ≈ 10-9 после LDPC-кодирования. Эти минимальные значения, начиная с которых канал связи можно считать надежным при использовании соответствующего типа модуляции с LDPC-кодированием.

Как показано в таблице 3, определены минимальные частоты, начиная с которых канал связи можно считать надежным при использовании соответствующего типа модуляции с LDPC-кодированием.

Таблица 3. Частотные границы для достижения BER ≈ 10^{-9} после LDPC-кодирования для различных модуляций

Кодирование	Частота (ГГц)	Кодирование	Частота (ГГц)
BPSK + LDPC	2,41	**8PSK + LDPC**	2,18
QPSK + LDPC	2,39	**GMSK + LDPC**	2,34

На рисунках 6 и 7 представлены зависимости BER от частоты при различных скоростях передачи данных для схем модуляции BPSK и QPSK с LDPC-кодированием.

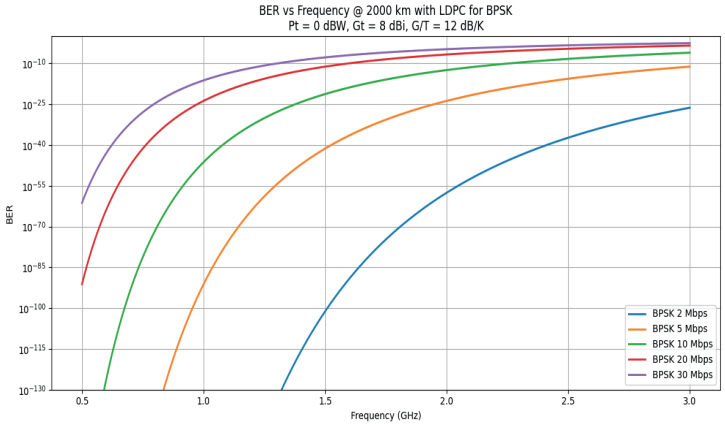


Рисунок 6 – BER от частоты при различных скоростях передачи для BPSK с LDPC-кодированием.

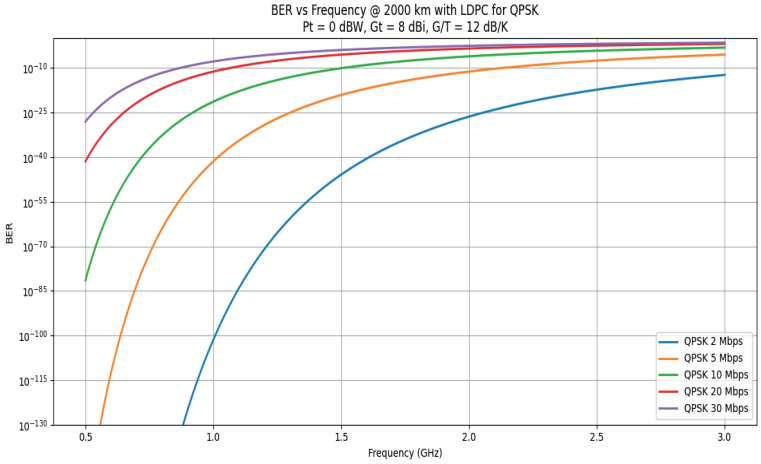


Рисунок 7 – BER от частоты при различных скоростях передачи для QPSK с LDPC-кодированием.

Графики показывают, что при использовании LDPC-кодирования BER значительно снижается для всех частот и скоростей передачи данных. Для BPSK на частотах ниже 1 ГГц высокая надежность сохраняется только при скоростях до 2–5 Мбит/с, тогда как при увеличении скорости до 20–30 Мбит/с BER возрастает. Аналогичные зависимости характерны для QPSK, но при более низких частотах деградация качества связи происходит быстрее. Таким образом, увеличение скорости передачи данных требует выбора частоты выше 2 ГГц для обеспечения надежности связи на уровне BER $\approx 10^{-9}$.

Таким образом, применение SDR на базе AD9364 в сочетании с мощным усилителем HMC998 и кодированием LDPC обеспечивают низкий BER (до 10-9), что существенно улучшает надежность связи даже при удаленных

орбитах. Представленные графики демонстрируют эффективность каждого этапа улучшения: модуляции, кодирования и усиления сигнала.

Обсуждение. Результаты моделирования подтверждают, что модуляция BPSK демонстрирует наилучшую устойчивость к помехам, обеспечивая крайне низкий BER при частотах до 4 ГГц и скоростях передачи данных до 0,5 Мбит/с. При увеличении скорости до 10 Мбит/с надежность снижается уже при частотах выше 1,5 ГГц. Модуляция QPSK также показывает хорошие характеристики при низких скоростях и частотах до 3,8 ГГц, но ее устойчивость падает при скорости 10 Мбит/с, где допустимый диапазон частот сужается до примерно 1,2 ГГц. Для 8-PSK наблюдается высокая чувствительность к снижению отношения сигнал/шум, что ограничивает ее применение при больших скоростях передачи. GMSK демонстрирует поведение, близкое к BPSK на низких скоростях, и может быть использована для энергосберегающих систем при частотах до 4 ГГц и скоростях до 1 Мбит/с. Таким образом, выбор схемы модуляции и скорости передачи данных должен учитывать требуемую надежность связи и условия работы радиоканала.

Заключение. Применение SDR на базе AD9364 в сочетании с усилителем мощности HMC998 дает возможность повысить надежность передачи данных с МКА. Проведенное моделирование показало, что система демонстрирует высокую устойчивость к помехам и позволяет эффективно работать в диапазоне частот от 300 МГц до 4–5 ГГц при скорости передачи данных до 1–2 Мбит/с. Анализ различных схем модуляции подтвердил, что BPSK и QPSK являются наиболее надежными по показателю вероятности битовой ошибки BER на типичных орбитах CubeSat. Полученные результаты показывают, что использование SDR в сочетании с помехоустойчивым кодированием и усилением мощности позволяет обеспечить качественную связь даже при ограниченных ресурсах МКА, и могут быть использованы в дальнейшем при проектировании современных спутниковых систем связи.

References

Maheshwarappa M.R., Bridges C.P. (2015) Software Defined Radios for Small Satellites. In Proceedings of the NASA/ESA Conference on Adaptive Hardware and Systems (AHS), Montreal, QC.— P. 1-6. (in English)

Mukherjee J., Ramamurthy B. (2013) Communication technologies and architectures for space network and interplanetary Internet IEEE Commun. Surv. Tut. — Vol. 15, № 2. — P. 881-897. (in English)

Davies A. (2019) Satellite IoT forecast 2019-202. Available: *http://tinyurl.com/y32s68qs* (in English)

Bacco M. et al.(2019) IoT applications and services in space information networks IEEE Wireless Commun., vol. 26, no. 27. — P. 31–37. (in English)

Rodrigues P., Oliveira A., Cunha S., Pinho R., Cordeiro C.S., and Carvelho R.A. (2013) GAMANET: Disrupting communications and networking in space In Proc. Int. Astronautical Congr. — P. 4237–4245. (in English)

Vishnuvardhan G. Rao, A.Mohammed Abbas,S. Palanтivel. (2019) Design of Multi Modulation Scheme for Software Defined Radio using FPGA. International Journal of Engineering and Advanced Technology (IJEAT). — Vol. 9 — P. 1-7. (in English)

Seunghwa Jung;Jihwan P. Choi (2019) Reliability of Small Satellite Networks With Software-Defined Radio and Enhanced Multiple Access Protocol. IEEE Transactions on Aerospace and Electronic Systems, — P. 1–6. (in English)

Bacco M., Boero L., Cassara P., Colucci M., Gotta A., Marchese M., Patrone F. (2019) "IoT Applications2019 and Services in Space Information Networks," IEEE Wireless Commun. — Vol. 26, №. 2. — P. 31–37. (in English)

ZengIn, Mehmet Ali; Kucuk, Kerem; Colak, Sultan Aldirmaz (2018). [IEEE 2018 26th Signal Processing and Communications Applications Conference (SIU) - Izmir, Turkey (2018.5.2-2018.5.5)] 2018 26th Signal Processing and Communications Applications Conference (SIU) - Real-time PSK transceiver design and implementation using software defined radio. — P. 1–4. doi:10.1109/SIU.2018.8404260 (in English)

Nivin R, Rani J Sheeba, Vidhya P . (2016) 2016 International Conference on Communication Systems and Networks (ComNet). Design and hardware implementation of reconfigurable nano satellite communication system using FPGA based SDR for FM/FSK demodulation and BPSK modulation. – P. 1–6. doi:10.1109/CSN.2016.7823976 (in English)

What are SmallSats and CubeSats. Available online: https://www.nasa.gov/content/what-are-smallsats-and-cubesats (accessed on 10 May 2021) (in English)

Heuberger A. Mehnert M. Burkhardt F. Oschek J. (2021) Advanced receiver module for satellite standard ETSI-SDR (ESDR). In Proceedings of the IEEE 13th International Symposium on Consumer Electronics 2009, Kyoto, Japan, 25-28 May 2009. . — P. 765-768. (in English)

Yasir M. O., Kenichi Asami (2021) Design of Software-Defined Radio-Based Adaptable Packet Communication System for Small Satellites. Aerospace. — 8(6), P.159-165. *https://doi.org/10.3390/aerospace8060159* (in English)

Ladvánszky J. (2021) There is no Theoretical Limit for Noise Reduction in Digital Communications. Netw. Commun. Technol. — 6(1) . — P. 34— 42. (in English)

García-Álvarez J. C. (2024) Introduction to Communication Systems. In Digital Electronic Communications. — P. 3-10. (in English)

ACADEMIC SCIENTIFIC JOURNAL OF COMPUTER SCIENCE
ISSN 1991-346X
Volume 4. Namber 356 (2025). 274–287

https://doi.org/10.32014/2025.2518-1726.396

UDC 621.315.2

© A.N. Seraly[1], A.D. Mekhtiyev [2], G.Z. Ziyatbekova[3,4,5*], K.B. Begalieva[1] , R.A. Mekhtiyev[1], 2025.

[1] S. Seifullin Kazakh Agrotechnical Research University, Astana, Kazakhstan;
[2]Abylkas Saginov Karaganda Technical University, Karaganda, Kazakhstan;
[3]Almaty Technological University, Almaty, Kazakhstan;
[4]Al-Farabi Kazakh National University, Almaty, Kazakhstan;
[5]Institute of Information and Computational Technologies SC MSHE RK, Almaty, Kazakhstan.
E-mail: ziyatbekova1@gmail.com

DEVELOPMENT OF HARDWARE FOR MONITORING OPTICAL PARAMETERS

Seraly Aigul — Doctoral student of the specialty «8D07107-Technical system managment», S. Seifullin Kazakh Agrotechnical Research University, Astana, Kazakhstan,
E-mail: seralyaigul8@gmail.com, https://orcid.org/0009-0002-3836-4582;
Mekhtiyev Ali — Candidate of Technical Scienes, professor, Vice-Rector for Research NCJSC «Abylkas Saginov Karaganda Technical University», Karaganda, Kazakhstan,
E-mail: barton.kz@mail.ru, https://orcid.org/0000-0002-2633-3976;
Ziyatbekova Gulzat—PhD, Associate Professor, Almaty Technological University; Al-Farabi Kazakh National University; Institute of Information and Computational Technologies SC MSHE RK, Almaty, Kazakhstan,
E-mail: ziyatbekova1@gmail.com, https://orcid.org/0000-0002-9290-6074;
Begaliyeva Kalamkas — PhD, S.Seifullin Kazakh Agrotechnical Research University, Astana, Kazakhstan,
E-mail: kalamkas_b@mail.ru, https://orcid.org/0000-0002-4216-9184;
Mekhtiyev Ruslan — Master of Science in Engineering, Research Associate Department of Electrical Equipment Operation, S. Seifullin Kazakh Agrotechnical University, Astana, Kazakhstan,
E-mail: ruslanmekhtiyev@gmail.com, https://orcid.org/0000-0001-5416-3444.

Abstract. This article describes the device of an experimental installation for obtaining data in the format of images or short videos. Fiber-optic sensors (BPS), new generation lasers, and an FLG camera were used to create the experimental setup. The sequential functionality of the experimental stand is described, as well as the data and metrological characteristics of lasers, which make it possible to choose when used in various circumstances. In addition, further research and the choice of machine learning methods for data classification depend on the quality of the information being processed, so the article discusses the technical characteristics of lasers. The article presents the results of checking the stability of laser emitters

(diode lasers) according to parameters such as optical power and wavelength, conducted in an accredited laboratory. At the current stage of our research, the focus is on data collection and monitoring of optical parameters that change when exposed to fiber-optic sensors, so laser selection is crucial to obtain accurate data. The article presents the research of Russian scientists who used fiber-optic sensors widely used in construction and security systems, such as the Brega grid. Further research will focus on data recognition and determining the nature of the impact on sensors over a certain period of time using neural networks and deep learning methods. Modern technologies are developing rapidly, so significant changes in the choice of software solutions and data integration may occur during the research process. Therefore, the experimental setup has unique characteristics and can be upgraded and improved depending on the technologies used, such as modern lasers and FLG cameras.

Keywords: fiber optic sensors, lasers, machine learning, Bragg lattice, neural networks

© А.Н. Сералы[1], А.Д. Мехтиев[2], Г.З. Зиятбекова[3,4,5*], К.Б. Бегалиева[1], Р.А. Мехтиев[1], 2025.

[1]С. Сейфуллин атындағы Қазақ агротехникалық зерттеу университеті, Астана, Қазақстан;

[2]Әбілқас Сағынов атындағы Қарағанды техникалық универтитеті, Қарағанды, Қазақстан;

[3]Алматы технологиялық университеті, Алматы, Қазақстан;

[4]Әл-Фараби атындағы Қазақ ұлттық университеті, Алматы, Қазақстан;

[5]ҒК ҚР ҒЖБМ Ақпараттық және есептеуіш технологиялар институты, Алматы, Қазақстан.

[*]E-mail: ziyatbekova1@gmail.com

ОПТИКАЛЫҚ ПАРАМЕТРЛЕРДІ БАҚЫЛАУҒА АРНАЛҒАН АППАРАТТЫҚ ҚҰРЫЛҒЫНЫ ӘЗІРЛЕУ

Сералы Айгуль — «8D07107-Техникалық жүйелерді басқару» мамандығының 2 курс докторанты, С. Сейфуллин атындағы Қазақ агротехникалық зерттеу университеті, Астана, Қазақстан,
E-mail: seralyaigul8@gmail.com, https://orcid.org/0009-0002-3836-4582;

Мехтиев Али — техника ғылымдарының кандидаты, "Абылқас Сағынов атындағы Қарағанды техникалық университеті" КеАҚ ғылыми-зерттеу ісі жөніндегі проректоры, Қарағанды, Қазақстан,
E-mail: barton.kz@mail.ru, https://orcid.org/0000-0002-2633-3976;

Зиятбекова Гулзат — PhD, қауымдастырылған профессор Алматы технологиялық университеті; Әл-Фараби атындағы Қазақ ұлттық университеті; ҒК ҚР ҒЖБМ Ақпараттық және есептеуіш технологиялар институты, Алматы, Қазақстан,
E-mail: ziyatbekova1@gmail.com, https://orcid.org/0000-0002-9290-6074;

Бегалиева Каламкас — Phd, компьютерлік ғылымдар кафедрасының аға оқытушысы, С. Сейфуллин атындағы Қазақ агротехникалық зерттеу университеті, Астана, Қазақстан,
E-mail: kalamkas_b@mail.ru, https://orcid.org/0000-0002-4216-9184;

Мехтиев Руслан — техника ғылымдарының магистрі, С. Сейфуллин атындағы Қазақ агротехникалық университетінің электр жабдықтарын пайдалану кафедрасының ғылыми қызметкері, Астана, Қазақстан,
E-mail: ruslanmekhtiyev@gmail.com, https://orcid.org/0000-0001-5416-3444.

Аннотация. Бұл құжат кескін пішіміндегі немесе қысқа бейнелердегі деректерді алуға арналған эксперименттік қондырғының дизайнын сипаттайды. Эксперименттік қондырғыны құру үшін талшықты-оптикалық сенсорлар (FOS), жаңа буын лазерлері және FLG камерасы пайдаланылды. Эксперименттік стендтің дәйекті функционалдығы сипатталған, сонымен қатар әртүрлі жағдайларда пайдаланған кезде таңдауды қамтамасыз ететін лазерлердің деректері мен метрологиялық сипаттамалары берілген. Зерттеулер және деректерді жіктеу үшін машиналық оқыту әдістерін таңдау, өңделетін ақпараттың сапасына байланысты, сондықтан мақалада лазерлердің техникалық сипаттамалары талқыланады. Мақалада аккредиттелген зертханада жүргізілген оптикалық қуат және толқын ұзындығы сияқты параметрлер бойынша лазерлік эмитенттердің (диодты лазерлердің) тұрақтылығын сынау нәтижелері берілген. Біздің зерттеуіміздің қазіргі кезеңі деректерді жинауға және талшықты-оптикалық сенсорларға (FOS) әсер еткенде өзгеретін оптикалық параметрлерді бақылауға бағытталған, сондықтан дәл деректерді алу үшін лазерді таңдау өте маңызды. Мақалада құрылыста және қауіпсіздік жүйесінде кеңінен қолданылатын талшықты-оптикалық сенсорларды пайдаланған отандық ғалымдардың зерттеулері берілген, мысалы, Breg торы. Әрі қарайғы зерттеулер нейрондық желілер мен терең оқыту әдістерін пайдалана отырып, деректерді тануға және белгілі бір уақыт аралығында сенсорларға әсер ету сипатын анықтауға бағытталған. Заманауи технологиялар қарқынды дамып келеді, сондықтан зерттеу процесінде бағдарламалық шешімдерді таңдауда және деректерді біріктіруде елеулі өзгерістер болуы мүмкін. Сондықтан эксперименттік қондырғының бірегей сипаттамалары бар және заманауи лазерлер мен FLG камералары сияқты қолданылатын технологияларға байланысты жаңартылуы және жетілдірілуі мүмкін.

Түйін сөздер: талшықты-оптикалық сенсорлар, лазерлер, машиналық оқыту, Bregg торы, нейрондық желілер

© А.Н. Сералы[1], А.Д. Мехтиев[2], Г.З. Зиятбекова[3,4,5*], К.Б. Бегалиева[1], Р.А. Мехтиев[1], 2025.

[1]Казахский агротехнический исследовательский университет им. С.Сейфуллина, Астана, Казахстан;
[2]Карагандинский технический университет имени Абылкаса Сагинова, Караганда, Казахстан;
[3]Алматинский технологический университет, Алматы, Казахстан;
[4]Казахский национальный университет имени аль-Фараби, Алматы, Казахстан;
[5]Институт информационных и вычислительных технологий КН МНВО РК, Алматы, Казахстан.
[*]E-mail: ziyatbekova1@gmail.com

РАЗРАБОТКА АППАРАТНОГО СРЕДСТВА ДЛЯ КОНТРОЛЯ ОПТИЧЕСКИХ ПАРАМЕТРОВ

Сералы Айгуль — докторант специальности «8D07107-Управление техническими системами», Казахский агротехнический исследовательский университет имени С. Сейфуллина, Астана, Казахстан,
E-mail: seralyaigul8@gmail.com, https://orcid.org/0009-0002-3836-4582;
Мехтиев Али — кандидат технических наук, проректор по научно-исследовательской работе НАО "Карагандинский технический университет имени Абылкаса Сагинова", Караганда, Казахстан,
E-mail: barton.kz@mail.ru, https://orcid.org/0000-0002-2633-3976;
Зиятбекова Гулзат — PhD, ассоциированный профессор, Алматинский технологический университет; Казахский национальный университет имени аль-Фараби; Институт информационных и вычислительных технологий КН МНВО РК, Алматы, Казахстан,
E-mail: ziyatbekova1@gmail.com, https://orcid.org/0000-0002-9290-6074;
Бегалиева Каламкас — Phd, Казахский агротехнический исследовательский университет имени С. Сейфуллина, Астана, Казахстан,
E-mail: kalamkas_b@mail.ru, https://orcid.org/0000-0002-4216-9184;
Мехтиев Руслан — магистр технических наук, научный сотрудник кафедры «Эксплуатация электрооборудования», Казахский агротехнический университет им. С. Сейфуллина, Астана, Казахстан,
E-mail: ruslanmekhtiyev@gmail.com, https://orcid.org/0000-0001-5416-3444.

Аннотация. В данной статье описано устройство экспериментальной установки для получения данных в формате изображения или коротких видеороликов. Для создания экспериментальной установки были использованы волоконно-оптические датчики (ВОП), лазеры нового поколения и камера FLG. Описан последовательный функционал экспериментального стенда, так же представлены данные и метрологические характеристики лазеров, которые дают возможность выбора при использовании в различных обстоятельствах. Кроме того, дальнейшие исследования и выбор методов машинного обучения для классификации данных зависят от качества обрабатываемой информации, поэтому в статье рассмотрены технические характеристики лазеров. В статье

изложены результаты проверки стабильности лазерных излучателей (диодных лазеров) по таким параметрам, как оптическая мощность и длина волны, проведенной в аккредитованной лаборатории. На текущем этапе нашего исследования основное внимание уделяется сбору данных и мониторингу оптических параметров, которые изменяются при воздействии на волоконно-оптические датчики (ВОП), поэтому выбор лазера имеет решающее значение для получения точных данных. В статье представлены исследования отечественных ученых, которые использовали волоконно-оптические датчики, широко применяемые в строительстве и охранных системах, такие как решетка Брега. Дальнейшее исследование будет сосредоточено на распознавании данных и определении характера воздействия на датчики в определённый период времени с использованием нейросетей и методов глубокого обучения. Современные технологии развиваются стремительно, поэтому в процессе исследования могут произойти значительные изменения в выборе программных решений и интеграции данных. Поэтому экспериментальная установка обладает уникальными характеристиками и может быть модернизирована и усовершенствована в зависимости от применяемых технологий, таких как современные лазеры и FLG-камеры.

Ключевые слова: волоконно-оптические датчики, лазеры, машинное обучение, решетка Брегга, нейросети

Введение. В настоящее время волоконно-оптические линии передачи (ВОЛП) широко используются в телекоммуникациях благодаря их высокой пропускной способности и низким потерям. Однако надёжность и безопасность таких линий остаются критически важными задачами, особенно в условиях возросших рисков несанкционированного доступа и возможных механических повреждений кабеля. Существует потребность в доступных методах непрерывного мониторинга, позволяющих вовремя выявлять деформации, обрывы и попытки подключения к оптическому кабелю.

В предлагаемой работе рассматривается многоканальная экспериментальная установка с использованием волоконно-оптических датчиков (ВОП). Основу системы составляют резервные оптические волокна в составе кабеля, оптический сплиттер, лазерный источник и КМОП-фотоматрица (CMOS-камера). В статье описывается принцип действия, аппаратная реализация и результаты отображающие экспериментальным стендом.

Общая концепция. Разрабатываемая методика основана на применении оптических волокон в качестве чувствительных элементов, по каждому из которых проходит лазерное излучение. На выходе из волокна формируется световое пятно, фиксируемое КМОП-матрицей (комплементарной металл-оксидной полупроводниковой фотоматрицей). Данный световой отпечаток представляет собой уникальный оптический сигнал, параметры которого зависят от состояния волокна. При возникновении внешнего воздействия —

таких как механические изгибы, давления, удары или несанкционированные вмешательства — происходит изменение физических характеристик волокна, что, в свою очередь, приводит к деформации оптического пути и изменению интенсивности, формы или других параметров светового пятна. Таким образом, изменения в оптическом сигнале служат индикатором нарушения нормального функционирования системы, что позволяет использовать данную методику для создания высокочувствительных систем мониторинга и охраны.

В качестве волоконно-оптической линии (ВОЛ), обеспечивающей передачу сигнала и выполнение чувствительной функции, используется стандартный одномодовый волоконно-оптический кабель (ВОК), широко применяемый в современных внутризоновых телекоммуникационных сетях. Конструктивной основой кабеля является оптоволокно стандарта G.652, отличающееся низкими потерями при передаче сигнала и высокой степенью устойчивости к внешним воздействиям. Подробные характеристики данного типа волокна рассмотрены в ряде работ (Горлов et al., 2016; Юрченко et al., 2018), в которых описаны как его физико-технические параметры, так и особенности взаимодействия с различными типами внешних нагрузок.

Современные исследования в области волоконно-оптических датчиков (ВОД) опираются как на новейшие достижения, так и на анализ ошибок, допущенных в предыдущих разработках, что позволяет критически переосмысливать накопленные знания и вносить существенные улучшения в существующие технические решения. Несмотря на наличие определённых трудностей и неудачных экспериментов, прогресс в данной области очевиден, о чём свидетельствуют успешные примеры реализации ВОД в различных отраслях промышленности, включая транспортную безопасность, энергетическую инфраструктуру и автоматизированные системы контроля.

Теоретические основы проектирования волоконно-оптических датчиков, принципы их функционирования, а также их классификация и функциональные возможности подробно изложены в ряде фундаментальных источников, в частности, в работе (Okossi et al.,1991). Кроме того, значительный вклад в развитие данной области внесли исследования, представленные в материалах Бусурина, где обобщён как отечественный, так и зарубежный опыт разработки и внедрения ВОД в различных сферах народного хозяйства. Обращение к источникам (Okossi et al., 1991) обусловлено их высокой научной ценностью, а также тем, что они охватывают многолетний период эволюции волоконно-оптических технологий, позволяя систематизировать имеющиеся знания, классифицировать датчики по различным признакам и выделить перспективные направления для дальнейших исследований. Широкое распространение волоконно-оптических линий передачи поставило перед нами важную задачу: обеспечить контроль за состоянием кабеля и защитить данные от утечек. Ранее считалось, что оптоволокно невозможно прослушать,

но сегодня существуют технологии, позволяющие извлекать сигнал, что делает необходимым создание системы мониторинга для оперативного выявления повреждений и попыток несанкционированного доступа.

Существует также информация о применении волоконно-оптических датчиков (ВОД) для защиты периметров и предотвращения несанкционированного доступа (Kumagai et al., 2018). Волоконно-оптические системы охраны пользуются значительным спросом на рынке и имеют ряд преимуществ по сравнению с традиционными системами охраны, основанными на использовании электрического тока или радиосигналов (Ковтун et al., 2020). Преимущества ВОД в контексте охранных систем периметра уже были изучены в ранних работах автора (Ковтун et al., 2020). Идеи, представленные в публикациях (Kumagai et al., 2018; Ковтун et al., 2020), особенно ценные для разработки аппаратно-программного комплекса контроля технического состояния кабелей (АПКТК), который также должен обеспечивать защиту от несанкционированного доступа (НД).

В источнике (Masoudi et al., 2019) представлен вариант системы контроля (СК), который наиболее близок к нашим задачам, поскольку он предназначен для мониторинга технического состояния силовых электрических кабелей, уложенных под водой. В этой системе оптический сигнал от источника сначала усиливается, а затем проходит через стадию фильтрации с использованием вейвлет-преобразования. Для источника излучения используется полоса пропускания 100 ГГц. Световой сигнал направляется через оптический циркулятор и попадает на фотоприемник, проходя через оптические волокна, которые размещены вдоль силового кабеля.

Также возможно использование волновых решеток Брэгга (ВРБ), однако это потребует применения более дорогостоящего аналитического оборудования для изучения отраженного спектра. Известно, что внешнее механическое воздействие на оптическое волокно (ОВ) приводит к изменению его физических параметров и коэффициента преломления, что, в свою очередь, влияет на характеристики луча в оптических интерферометрах (таких как Маха-Цендера, Фабри-Перо) и параметры оптического рассеяния (Рэлея, Мандельштама-Бриллюэна), а также на ширину полосы отражения Брэгговских решеток (Hong et al., 2021).

Материалы и методы. Принцип построения экспериментальной установки заключается в следующем последовательном соединении показанный рисунке 1.

Рисунок 1 – Последовательное соединение устройств экспериментального стенда.

Рассматривая функционал экспериментального стенда, описать пошаговое выполнения можно следующим образом:

1. ***Генерация стабильного оптического сигнала***: Лазерный источник света генерирует стабильный оптический сигнал, который используется в системе для дальнейших измерений. Лазеры важны для таких систем, так как они обеспечивают точность и стабильность в процессе работы.

2. ***Прохождение света через оптоволоконный сенсор***: после генерации сигнала, он направляется через оптоволоконный сенсор. Оптоволокно часто используется для передачи света, потому что оно минимизирует потери и позволяет передавать сигналы на большие расстояния.

3. ***Изменение параметров сигнала***: при воздействии внешних факторов (например, температуры, давления, механических деформаций и других факторов) параметры света, проходящего через оптоволокно, начинают изменяться. Это может выражаться в изменении яркости, формы, распределения сигнала.

4. ***Оптоволоконная система с адаптерами***: Эта система включает несколько оптоволоконных кабелей, соединенных через адаптеры и распределенных через специальный блок. Такой подход позволяет эффективно управлять светом и его распределением по системе, обеспечивая корректное прохождение и точную передачу данных.

5. ***Регистрация сигнала с помощью CMOS-камеры***: CMOS-камера используется для регистрации выходного сигнала, который представлен в виде световых пятен. Эти пятна отражают изменения, происходящие в оптоволокне при воздействии внешних факторов.

6. ***Обработка и визуализация данных на персональном компьютере***: на компьютере производится обработка изображений, полученных с CMOS-камеры. Это включает в себя анализ изменений в интенсивности и распределении световых пятен, что позволяет оценить степень воздействия на оптоволокно и извлечь нужную информацию.

Результаты и обсуждение. ***Оптический источник***. Полупроводниковый лазер, работающий на выбранной длине волны (например, 650 нм или другом подходящем диапазоне), генерирует стабильный оптический сигнал. Лазерный драйвер обеспечивает регулирование тока и температуры, что гарантирует постоянную мощность излучения.

Оптический сплиттер 1×4. Делит входной лазерный сигнал на четыре независимых канала, позволяя распределить свет в четыре отдельных оптоволоконных сенсора.

Оптоволоконные сенсоры. В качестве чувствительных элементов используются одномодовые волокна стандарта G.652.Длина волокон может варьироваться от нескольких метров до километров, в зависимости от условий испытаний.

Торцы наконечника коннектора. Формируют выходные пучки, благодаря которым каждый канал создаёт отдельное световое пятно на CMOS - фотоматрице без использования дополнительных оптических элементов.

На рисунке 2 показана детальная схема фото матрицы.

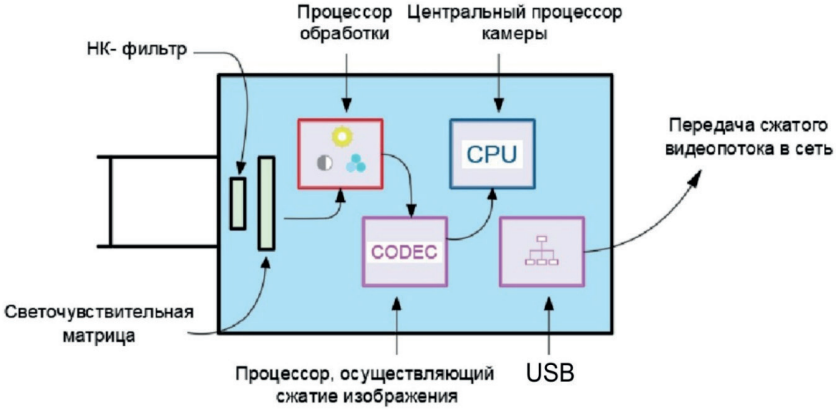


Рисунок 2 – Детальная схема фотоматрицы

Регистрирует изображение, состоящее из четырёх световых пятен, с высоким разрешением. Встроенный аналого-цифровой преобразователь (АЦП) преобразует оптический сигнал в цифровой видеопоток для дальнейшей обработки.

Блок сбора данных. Принимает поток изображений с CMOS -фотоматрицы и анализирует распределение яркости и геометрию пятен. Применяются потоковые изображения так же короткие ролики.

Лазерный блок и оптический сплиттер. Лазерный источник (чаще всего диодный лазер) генерирует стабильное оптическое излучение, которое посредством лазерного драйвера поддерживается на постоянном уровне мощности. Затем оптический сплиттер 1×4 равномерно распределяет этот сигнал по четырём каналам, каждый из которых поступает в свой

оптоволоконный сенсор. Такие сенсоры могут эксплуатироваться как в условиях лаборатории, так и в полевых испытаниях.

Волоконные сенсоры. В установке в качестве сенсоров применяются одномодовые волокна (стандарт G.652), интегрированные в оптоволоконную систему. При воздействии (например, изгиб, растяжение, микродеформация) часть мощности теряется, что приводит к изменению яркости и формы выходного светового пятна. Дополнительно возможны фазовые и поляризационные сдвиги, которые отражаются в изменении распределения светового потока.

CMOS - фотоматрица и блок сбора данных. Выходной сигнал от каждого сенсора формируется непосредственно торцами наконечников коннекторов, что позволяет получить четкие световые пятна на CMOS -фотоматрице. Камера с высоким разрешением и встроенным АЦП регистрирует четыре отдельных пятна, после чего цифровой поток передаётся в блок сбора данных. Здесь выполняется первичная фильтрация изображений, а затем – детальный анализ распределения яркости и геометрии пятен с использованием алгоритмов порогового контроля и методов машинного обучения. Это позволяет выявлять аномальные изменения, связанные с воздействием внешних факторов на оптоволокно.

На рисунке 3 показано соединение лазера с оптоволокном и выход на фотоматрицу, готовой экспериментальной установки.

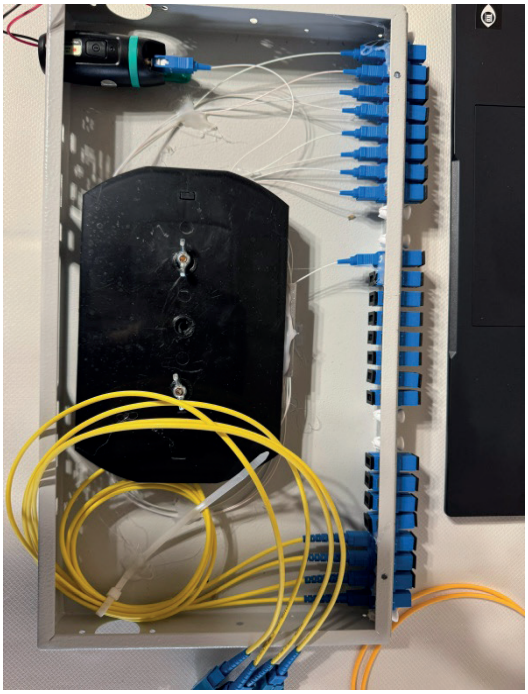


Рисунок 3 – Соединение лазера с волоконно-оптическими датчиками.

На рисунке 4 показано соединение фотоматрицы с персональным компьютером для получения данных.

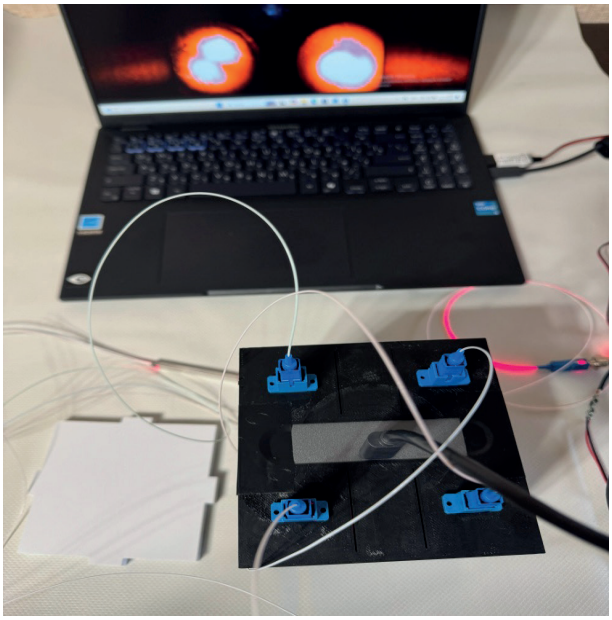


Рисунок 4 – Отображение данных в ПК.

В ходе лабораторных испытаний проверялась реакция каждого канала на механические воздействия:

- ***Изгиб:*** при постепенном увеличении радиуса изгиба одного из сенсоров наблюдалось снижение интегральной яркости соответствующего светового пятна на 10–20 %.
- ***Разрыв:*** при полном переломе волокна сигнал мгновенно обнулялся, и световое пятно исчезало.
- ***Имитация врезки (прослушки):*** Установка дополнительного ответвления в волокне приводила к появлению дополнительных потерь (5–15 %), что система регистрировала как аномальное событие.

Ключевым результатом испытаний стало подтверждение возможности круглосуточного мониторинга с автоматическим выявлением событий на основе пороговых алгоритмов. При этом точность определения изменений интенсивности составляла около 85–90 %, а внедрение методов машинного обучения (например, сверточных нейронных сетей) позволило повысить точность распознавания аномальных событий до 92–95 %.

Заключение. Разработанная установка демонстрирует, что использование четырёх параллельных оптоволоконных сенсоров и КМОП-фотоматрицы обеспечивает надёжный и экономичный способ диагностики состояния волоконно-оптических линий. Ключевые достоинства системы:

Многоканальность: Возможность одновременного контроля нескольких участков оптоволокна.

Наглядность: Оптический сигнал преобразуется в изображение, что позволяет проводить как автоматизированный, так и визуальный анализ.

Адаптивность: Система интегрируется в существующую сеть, используя резервные волокна кабеля.

Масштабируемость: при необходимости число каналов можно увеличить (например, 1×8, 1×16), а алгоритмы обработки – усовершенствовать.

Таким образом, представленная методика может применяться для контроля технического состояния волоконно-оптических линий в городских сетях связи, промышленности (например, охрана периметров, мониторинг трубопроводов) и в других системах, где требуется надёжное обнаружение повреждений и защита от несанкционированного доступа.

Литература

Горлов Н.И., Мехтиев А.Д., Таткеева Г.Г., Алдошина О.В., Югай В.В. (2016) Направляющие системы и волоконно-оптическая техника инфокоммуникаций: Учебник (Утвержден МОН РК). – Караганда: Изд-во КарГТУ. — 251 с. (in Russian)

Скляров О. К. (2022) Волоконно-оптические сети и системы связи: Учебное пособие для вузов. Изд. 6 изд., стер. — 268 с. ISBN 978-5-8114-9781-2.

Мехтиев А.Д, Югай В.В., Нешина Н. Г., Алькина А.Д. (2020) Волоконно-оптические системы идентификации физических величин: Монография. Карагандинский технический университет. – Караганда: Изд-во Кар ТУ. — 151с. ISBN 078-601-320-272-3.

Юрченко А.В., Мехтиев А.Д., Югай В.В, Алькина А.Д. (2018) Информационно-измерительные системы: Монография. Карагандинский государственный технический университет – Караганда: Изд - во КарГТУ. — 228 с.

Окосси Т., Окомото К., (1991) Волоконно-оптические датчики Пер. с японского Горбунова Г.Н., Ленинград; Энергоатомиздат. —255 с.

Бусурин В.И., Новсов Ю.Р. (1990) Волоконно-оптические датчики: Физические основы, вопросы расчета и применения. Москва, Энергоатомиздат. — 256 с.

T. Kumagai, S. Sato and T. Nakamura (2012) "Fiber-optic vibration sensor for physical security system," 2012 IEEE International Conference on Condition Monitoring and Diagnosis, pp. 1171-1174, doi: 10.1109/CMD.2012.6416369. .

Ковтун А.А., Мехтиев А.Д., Нешина Е.Г., Алькина А.Д. (2020) Системы охраны периметров на основе волоконно-оптических сенсоров. Научные труды военно-инженерного института радиоэлектроники и связи. — № 2 (40), (июнь). — С 25 - 29.

Ковтун А.А., Мехтиев А.Д., Нешина Е.Г., Алькина А.Д. (2020) Импульсно-рефлектометрический метод определения местоположения несанкционированного доступа к волоконно-оптической линии передачи информации. Научные труды военно-инженерного института радиоэлектроники и связи. № 2 (40), (июнь). — С 30 - 37.

A. Masoudi, J. A. Pilgrim, T. P. Newson and G. Brambilla (2019) "Subsea Cable Condition Monitoring With Distributed Optical Fiber Vibration Sensor" in Journal of Lightwave Technology. — vol. 37. — no. 4. — P. 1352-1358, 15 Feb.15, 2019, doi: 10.1109/JLT.2019.2893038.

Hong R. et al., (2021) "Enlarging Dynamic Strain Range in UWFBG Array-Based Φ-OTDR Assisted With Polarization Signal," in IEEE Photonics Technology Letters. — vol. 33. — no. 18. — P. 994-997, 15 Sept.15, 2021, doi: 10.1109/LPT.2021.3079186.

Mekhtiev A., Alkina A., Neftissov A., Kazambaev I., Kirichenko L. (2022) “Intelligent Systems for Monitoring the Integrity of Technical Objects Based on Distributed Fiber-optic Sensors” Ceur Workshop Proceedings, 3347. — P. 290-306

Mekhtiev A.D., Yurchenko A.V., Neshina E.G., Al'kina A.D., Madi P.S. (2020) "Physical Principles of Developing Pressure Sensors Using Refractive Index Changes in Optical Fiber Microbending" Russian Physics Journal, 63(2). — P. 323-331

Yurchenko A.V., Mechtiyev A.D, Bulatbaev F.N, Alkina A.D., Sh Madi.P. (2019) "Investigation of additional losses in optical fibers under mechanical action" Iop Conference Series Materials Science and Engineering, 516(1), 012004

Yurchenko A.V., Mechtiyev A.D, Bulatbaev F.N, Alkina A.D, Kokkoz M.M., (2017) "The clearance control system of the lever-hinge mechanism of the mine winder braking device using the capacitive sensors" Journal of Physics Conference Series, 881(1), 012034.

References

Gorlov N.I., Mehtiyev A.D., Tatkeeva G.G., Aldoshina O.V., Yugai V.V. (2016) Napravliyushie sistemy i volokonno-opticheskaya technica infokommunikaci [Guiding systems and fiber-optic infocommunication technology] Uchebnik (Utverzhden MON RK). – Karaganda: Izd-vo KarGTU. — 251 s. (in Russian)

Sklyarov O.K. (2022) Volokonno-opticheskie seti i sistemy svyazi [Fiber-optic networks and communication systems] Uchebnoe posobie dlya vuzov. Izd. 6 izd. ster. — 268 c. ISBN 978-5-8114-9781-2. (in Russian)

Mehtiyev A.D., Yugai V.V., Neshina N.N., Alkina A.D. (2020) Volokonno-opticheskie sistemy identifikacii fizicheskih velichin [Fiber-optic identification systems for physical quantities] Monografiya. Karagandinski technicheski universitet. — Karaganda: Izd-vo KarTU, 151c. ISBN 078-601-320-272-3. (in Russian)

Yurchenko A.V., Mekhtiev A.D., Yugai V.V., Alkina A.D. (2018) Informacionno-izmeritelnye sistemy [Information and measurement systems] Monografiya. Karagandinski gosudarstvenny technicheski universitet. - Karaganda: Izd-vo KarGTU. — 228 c. (in Russian)

Okossi T., Okomoto K. (1991) Volokonno-opticheskie datchiki [Fiber-optic sensors] per.s japonskogo Gorbunova G.N. Leningrad; Energoatomizdat. — 255 s. (in Russian)

Busurin V.I., Novsov Yu.R. (1990) Volokonno-opticheskie datchiki: Fizicheskie osnovy, voprosy rascheta i primeneniya. [Fiber-optic sensors: Physical fundamentals, calculation and application issues] Moskva, Energoatomizdat, 1990. — 256 s. (in Russian)

T. Kumagai, S. Sato and T. Nakamura (2012) "Fiber-optic vibration sensor for physical security system," 2012 IEEE International Conference on Condition Monitoring and Diagnosis. — P. 1171-1174, doi: 10.1109/CMD.2012.6416369. (in English)

Kovtun A.A., Mehtiyev A.D., Neshina N.N., Alkina A.D. (2020) Sistemy ohrany perimetrov na osnove volokonno-opticheskix sensorov [Perimeter security systems based on fiber-optic sensors] Nauchnye trudy voenno-inzhenernogo instituta radioelektroniki i svyazi. № 2 (40), (june). — P. 25-29 (in Russian)

Kovtun A.A., Mehtiyev A.D., Neshina N.N., Alkina A.D. (2020) Impulsno-reflektometricheski metod opredeleniya mestopolozheniya nesaksionirovannogo dostupa k volokonno-opticheski linii peredachi informaci [Pulse-reflectometric method for determining the location of unauthorized access to a fiber-optic information transmission line]. Nauchnye trudy voenno-inzhenernogo instituta radioelektroniki i svyazi. № 2 (40), (june). — P. 30-37. (in Russian)

A. Masoudi, J.A. Pilgrim, T.P. Newson and G. Brambilla (2019) "Subsea Cable Condition Monitoring with Distributed Optical Fiber Vibration Sensor," in Journal of Lightwave Technology. — vol. 37. — no. 4. — P. 1352-1358, 15 Feb.15, doi: 10.1109/JLT.2019.2893038. (in English)

Hong R. et al. (2021) "Enlarging Dynamic Strain Range in UWFBG Array-Based Φ-OTDR Assisted with Polarization Signal," in IEEE Photonics Technology Letters. — vol. 33. — no. 18. — P. 994-997, 15 Sept.15, doi: 10.1109/LPT.2021.3079186. (in English)

Mekhtiev A., Alkina A., Neftissov A., Kazambaev I., Kirichenko L. (2022) "Intelligent Systems for Monitoring the Integrity of Technical Objects Based on Distributed Fiber-optic Sensors" Ceur Workshop Proceedings, 3347. — P. 290-306 (in English)

Mekhtiev A.D., Yurchenko A.V., Neshina E.G., Al'kina A.D., Madi P.S. (2020) "Physical Principles

of Developing Pressure Sensors Using Refractive Index Changes in Optical Fiber Microbending" Russian Physics Journal, 63(2). — P. 323-331 (in English)

Yurchenko A.V., Mechtiyev A.D, Bulatbaev F.N, Alkina A.D., Sh Madi.P. (2019) "Investigation of additional losses in optical fibers under mechanical action" Iop Conference Series Materials Science and Engineering, 516(1), 012004 (in English)

Yurchenko A.V., Mechtiyev A.D, Bulatbaev F.N, Alkina A.D, Kokkoz M.M. (2017) "The clearance control system of the lever-hinge mechanism of the mine winder braking device using the capacitive sensors" Journal of Physics Conference Series, 881(1), 012034 (in English)

ACADEMIC SCIENTIFIC JOURNAL OF COMPUTER SCIENCE
ISSN 1991-346X
Volume 4. Namber 356 (2025). 288–305

https://doi.org/10.32014/2025.2518-1726.397

UDC 551.2+517.2

© A.A. Taurbekova[1*], M.V. Markosyan[2], 2025.
[1]KazNRTU named after K.I. Satpayev, Almaty, Kazakhstan;
[2]Yerevan Research Institute of Communications, Yerevan, Armenia;
[3]National University of Architecture and Construction of Armenia, Yerevan, Armenia.
E-mail: ainura230771@gmail.com

DEVELOPMENT AND IMPLEMENTATION OF A COMPUTATIONAL MODEL OF MAGMATIC PROCESSES IN THE BOWLS OF THE EARTH AND ON ITS SURFACE

Taurbekova Ainur — PhD student at KazNTU named after K.I. Satpayev, Almaty, Kazakhstan,
E-mail: ainura230771@gmail.com, ORCID ID: https://orcid.org/0000-0001-5433-8347;
Markosyan Mher — Doctor of Technical Sciences, Professor at Yerevan Research Institute of Communications and National University of Architecture and Construction of Armenia, Yerevan, Armenia,
E-mail: mark@yetri.am, https://orcid.org/ORCID ID: 0000-0003-1972-5266, Republic of Armenia

Abstract. The relevance and novelty of research in the study of tectonic activity occurring in the Earth's deep structures is substantiated. Specifically, the use of modern numerical approaches and algorithms for simulating tectonic processes is emphasized. It is shown that in the last decade two main areas of research in the Earth sciences have emerged: deep geodynamics and the early history of the Earth. The task of deep geodynamics includes the study of physical and chemical processes occurring in the bowels of the Earth below a level of 400 km, i.e. the boundary of the upper mantle itself, which, together with the crust, forms the tectonosphere – the main area of manifestation of plate tectonics. The aim of the work is to present a formal model of magmatic flows and simulations, for which these flows are represented as fluids with high viscosity and low Reynolds numbers. The modeling methodology is based on solving differential equations with initial and boundary conditions for fixed parameters of the moving boundary. As a result, a model aimed at analyzing the initial stages of tectonic changes is proposed, which can be applied to identify potentially seismically hazardous zones. The results of modeling using the developed software package are presented, demonstrating that this approach can complement existing methods for monitoring crustal dynamics. It is shown that the proposed methods provide higher modeling accuracy compared to existing

analogues and can be used as part of an early warning system, contributing to an increase in the accuracy and lead time of forecasting large-scale seismic events.

Keywords: numerical modeling, tectonic processes, magmatic flows, earthquakes, seismic activity, early warning, highly viscous fluid, Reynolds model, geophysics, computational model

© А.А. Таурбекова[1*], М.В. Маркосян[2], 2025.
[1]Сәтбаев университеті, Алматы, Қазақстан;
[2]Ереван байланыс ғылыми-зерттеу институты, Ереван, Армения;
[3] Армения Ұлттық сәулет және құрылыс университеті, Ереван, Армения.
E-mail: ainura230771@gmail.com

ЖЕР КӨЗІНДЕГІ ЖӘНЕ ОНЫҢ БЕТІНДЕГІ МАГМАТТЫҚ ПРОЦЕСТЕРДІҢ ЕСЕПТІК МОДЕЛІН ӘЗІРЛЕУ ЖӘНЕ ЕНГІЗУ

Таурбекова Айнұр — PhD докторант, Сәтбаев университеті, Алматы, Қазақстан,
E-mail: ainura230771@gmail.com, ORCID ID: https://orcid.org/0000-0001-5433-8347;
Маркосян Мгер — т.ғ.д., Ереван байланыс ғылыми-зерттеу институтының және «Армения Ұлттық сәулет және құрылыс университетінің» профессоры, Ереван, Армения,
E-mail: mark@yetri.am, https://orcid.org/0000-0003-1972-5266.

Аннотация. Жердің терең құрылымдарында кездесетін тектоникалық белсенділік саласындағы зерттеулердің өзектілігі мен жаңалығы дәлелденді. Атап айтқанда, тектоникалық процестерді модельдеу үшін заманауи сандық тәсілдер мен алгоритмдерді пайдалану талқыланады. Соңғы онжылдықта жер туралы ғылымдарда екі негізгі зерттеу бағыты пайда болғаны көрсетілген: терең геодинамика және Жердің ерте тарихы. Терең геодинамиканың міндетіне 400 км-ден төмен жер қойнауында болып жатқан физикалық-химиялық процестерді, яғни жер қыртысымен бірге тектоносфераны құрайтын жоғарғы мантияның шекарасын - тақта тектоникасының негізгі көрініс беру аймағын зерттеу кіреді. Бұл жұмыстың мақсаты Жер бетніе көтерілген магмалық ағындардың моделін ұсыну болып табылады, олар үшін бұл магма ағынының тұтқырлығы өте жоғары және Рейнольдс сандары төмен сұйықтықтар ретінде ұсынылған. Модельдеу әдістемесі жылжымалы шекараның бекітілген параметрлері үшін бастапқы және шекаралық шарттары бар дифференциалдық теңдеулерді шешуге негізделген. Нәтижесінде, сейсмикалық қауіпі жоғары аймақтарды анықтауға болатын тектоникалық өзгерістердің бастапқы кезеңдерін талдауға бағытталған модель ұсынылды. Жасалған бағдарламалық пакетті пайдалана отырып модельдеу нәтижелері ұсынылып, бұл тәсіл жер қыртысының магманың жер бетіне көтерілу динамикасын байқаудың қолданыстағы әдістерін толықтыра алатыны көрсетілген. Ұсынылған әдістер бар аналогтармен салыстырғанда модельдеудің жоғары дәлдігін қамтамасыз ететіні және ауқымды сейсмикалық оқиғаларды болжау кезіндегі дәлдігі

мен орындалу уақытының мүмкіндігіне ықпал ете отырып, ерте хабарлау жүйесінің бөлігі ретінде пайдаланылуы мүмкін екендігі көрсетілген. Яғни адам өміріне төнген қауіптің алдын алу мақсатында жасалған зерттеу әдісі болып табылады.

Түйін сөздер: сандық модельдеу, тектоникалық процестер, магмалық ағындар, жер сілкінісі, сейсмикалық белсенділік, ерте ескерту, жоғары тұтқыр сұйықтық, Рейнольдс моделі, геофизика, есептеу моделі

© А.А. Таурбекова[1*], М.В. Маркосян[2], 2025.
[1]КазНТУ им. К.И. Сатпаева, Алматы, Казахстан;
[2] Ереванский НИИ Средств связи, Ереван, Армения;
[3]Национальный университет Архитектуры и строительства Армении, Ереван, Армения.
E-mail: ainura230771@gmail.com

РАЗРАБОТКА И РЕАЛИЗАЦИЯ ВЫЧИСЛИТЕЛЬНОЙ МОДЕЛИ МАГМАТИЧЕСКИХ ПРОЦЕССОВ В НЕДРАХ ЗЕМЛИ И НА ЕЁ ПОВЕРХНОСТИ

Таурбекова Айнур — докторант КазНИТУ им. К.И. Сатпаева, Алматы, Казахстан,
E-mail: ainura230771@gmail.com, ORCID ID: https://orcid.org/0000-0001-5433-8347;
Маркосян Мгер — д.т.н., профессор Ереванский НИИ Средств связи и Национальный университет Архитектуры и строительства Армении, Ереван, Армения,
E-mail: mark@yetri.am, https://orcid.org/ ORCID ID: 0000-0003-1972-5266.

Аннотация. Обосновывается актуальность и новизна исследований в области изучения тектонической активности, протекающих в глубинных структурах Земли. В частности, использованию современных численных подходов и алгоритмов для имитации процессов тектонической активности. Показано, что в последнее десятилетие определились два главных направления исследований в науках о Земле – глубинная геодинамика и ранняя история Земли. В задачу глубинной геодинамики входит изучение физических и химических процессов, протекающих в недрах Земли ниже уровня 400 км, т.е. границы собственно верхней мантии, образующей вместе с корой тектоносферу – основную область проявления тектоники плит. Целью работы является представление формальной модели магматических потоков и моделирования, для чего эти потоки представлены как жидкости с высокой вязкостью и низкими значениями числа Рейнольдса. Методология моделирования построена на решении дифференциальных уравнений с начальными и граничными условиями при фиксированных параметрах подвижной границы. В результате предложена модель, направленная на анализ начальных этапов тектонических изменений, которая может быть применена для выявления потенциально опасных, в сейсмическом отношении, зон.

Приведены результаты моделирования с помощью разработанного пакета программ и показан, что этот подход способен дополнить существующие методы наблюдения за динамикой земной кары. Показаны, что предложенные методы обеспечивают более высокую точность моделирования по сравнению с существующими аналогами и может быть использована как часть системы раннего предупреждения, способствуя повышению точности и заблаговременности прогнозирования масштабных сейсмических явлений. Это метод исследования, призванный предотвратить угрозы жизни человека.

Ключевые слова: численное моделирование, тектонические процессы, магматические потоки, землетрясения, сейсмическая активность, раннее предупреждение, сильновязкая жидкость, модель Рейнольдса, геофизика, вычислительная модель

Введение. Одним из приоритетных направлений науки были и остается изучение процессов, происходящих в земных недрах. Актуальность этих исследований связана, прежде всего, с жизнедеятельностью человечества. Изучение процессов, происходящих в недрах Земли с применением моделирования связано с трудностями, прежде всего, из-за скудности, порой недоступности, и сложности получения информации о глубинных процессах.

Во многих областях научных исследований существует потребность в решении задач, требующих высокой точности и невозможности проведения прямых наблюдений или лабораторных экспериментов. Это особенно актуально для процессов, протекающих в недоступных средах, таких как глубинные слои земной кары, где получение данных осложняется природными условиями и ограниченностью методов диагностики. «Общепризнанно, что при изучении многих сложных явлений нельзя ограничиваться только экспериментальными и аналитическими исследованиями. Быстрый рост производительности компьютеров в последние десятилетия стимулировал развитие вычислительного направления в механике жидкости вообще и для исследования проблем гидродинамических неустойчивостей, частности» (Kim et al, 2022)

В последнее десятилетие определились два главных направления исследований в науках о Земле – глубинная геодинамика и ранняя история Земли. В задачу глубинной геодинамики входит изучение физических и химических процессов, протекающих в недрах Земли ниже уровня 400 км, т.е. границы собственно верхней мантии, образующей вместе с корой тектоносферу – основную область проявления тектоники плит. Для решения этой задачи в настоящее время применяются три метода: сейсмическая томография, экспериментальная минералогия и математическое моделирование. Применение двух из них стало возможным с появлением современных суперкомпьютеров, а эксперименты над минералами в условиях давлений и температур, господствующих в мантийных глубинах, – с созданием

аппаратов, воспроизводящих эти термодинамические условия (Minaev et al, 2022)

О методе решения одной из таких сложных задач, связанных с истечением сильновязкой жидкости из щели, является темой данного исследования.

Представляет интерес и предлагается вычислительная модель «Тектонического процесса, происходящего в земных недрах», которая может быть использована как часть системы раннего предупреждения, способствуя повышению точности и заблаговременности прогнозирования масштабных сейсмических явлений.

Постановка задачи. Математическую модель данной задачи гидродинамики представляет собой квазилинейное уравнение параболического типа (Utkin, 2021)

$$\frac{\partial \xi}{\partial t} = \frac{\varepsilon R}{3}\frac{\partial}{\partial x}\left(\xi^3 \frac{\partial \xi}{\partial \tau}\right) + \begin{cases} y(t)(1-x^2), \text{ если } -1 \le x \le 1; \\ 0, \text{ если } -\infty < x < -1,\ 1 < x < +\infty \end{cases} \quad (1)$$

совместно с уравнением следующего вида

$$\int_0^{\rho(t)} \xi(x,t)\,dx = \frac{2}{3}\int_0^t y(\tau)\,d\tau, \quad (2)$$

где: $x -$ горизонтальная координата, t – время, $\xi(x,t) -$ верхняя свободная поверхность накапливаемой жидкости, $y(\tau) -$ наибольшая скорость истечения жидкости из щели, которая считается заданной, $x = \rho(t)$ – подвижная граница накапливаемой жидкости. Все параметры и уравнения являются безразмерными.

Перед нами поставлена задача: дать решение уравнений (1) и (2).

В этих уравнениях неизвестными являются две функции: $\xi(x,t)$ и $\rho(t)$. Для определения их значений заданы начальные условия:

$$t = 0,\ \rho(0) = 1,\ \xi(x,0) = 0. \quad (3)$$

Для решения уравнения (1) должны быть заданы граничные условия, которые могут быть записаны в следующем виде:

$$x = 0, \quad \frac{\partial \xi(0,t)}{\partial x} = 0, \quad (4)$$

$$x = \rho(t), \quad \xi(\rho(t), t) = 0. \quad (5)$$

Итак, требуется найти значения функций $\xi(x,t)$ и $\rho(t)$, удовлетворяющие уравнениям (1) и (2), начальным условиям (3) и граничным условиям (4) и

(5). Здесь возникает проблема, связанная с подвижностью границы области решения задачи. В связи с этим для решения этой задачи здесь предлагается метод конечных разностей с некоторыми допущениями.

Решение задачи числа Рейнольдса. Идея метода заключается в следующем:

для некоторого момента времени t находится решение задачи для уравнения в частных производных (1) с начальными и граничными условиями при фиксированных значениях подвижной границы, т.е. значения функции $\rho(t)$. Уравнение (1) решается методом конечных разностей. Затем определяется значение функции $\rho(t+\Delta t)$ из уравнения (2). Для перехода к следующему значению времени $t+\Delta t$ используется полученное значение функции $\rho(t+\Delta t)$ и решается уравнение (1). Такой итерационный процесс продолжается до окончания всего процесса вычислений.

В качестве расчетной схемы для решения уравнения в частных производных (1) используется неявная нелинейная схема (Bykov et al, 2020):

$$\xi_{ij+1} = \xi_{ij} + \frac{\tau \cdot ER}{3h^2} \cdot \{(\frac{\xi_{ij+1} + \xi_{i+1j+1}}{2})^3 \cdot (\xi_{i+1j+1} - \xi_{ij+1}) - (\frac{\xi_{i-1j+1} + \xi_{ij+1}}{2})^3 \cdot (\xi_{ij+1} - \xi_{i-1j+1})\} +$$
$$+ \{y_{j+1} \cdot (1 - x_i^2), \textit{если } 0 \le i \le k, \quad 0, \textit{если } k+1 \le i \le n.\} \qquad (6)$$

Здесь: h – шаг по координате x; $k = \frac{1}{h}$ - количество точек, где функция $y(t) \neq 0$; τ – шаг по времени t; $j = 0,1,2,\ldots,m.$

Для вычисления значений функции $\rho(t)$ используется следующая рекуррентная формула:

$$\rho_{j+1} = \rho_j + \frac{2 \cdot (f_{j+1} - f_j)}{\xi_{j+1} + \xi_j}, \qquad (7)$$

где значения f_j определяются из следующей формулы:

$$f(t) = \int_0^{\rho(t)} \xi(x,t) \cdot dx.$$

Анализ результатов решения задачи. Получена расчетная схема решения данной задачи, что позволил разрабатывать его алгоритм. Решение задачи осуществляется для каждого момента времени с учетом изменения границы области интегрирования. По методу прогонки решение данной задачи будет приближенным.

Для решения задачи об истечении сильновязкой жидкости в «узком канале» применен подход к моделированию указанного процесса и разработан соответствующий программный пакет. Разработанная объектная модель (рис.1) позволяет расширить возможности программного комплекса для мониторинга сейсмического состояния регионов.

Рисунок 1 – Объектная модель системы мониторинга сейсмического состояния региона.

Для работы с моделью разработан дружественный пользовательский интерфейс, позволяющий реализовать численные алгоритмы, обеспечивать интеграцию комплекса с существующими и используемыми программными продуктами.

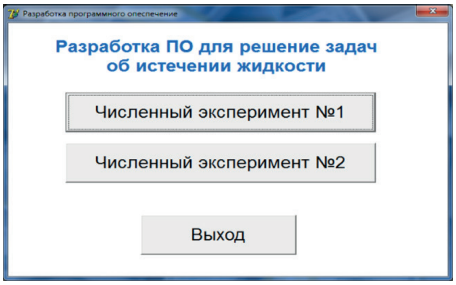


Рисунок 2 – Запуск программы и главный меню

Создание программы и численный эксперимент происходил по следующим пунктам:

- Наличие интерфейса для численного решения задачи;
- Класс для вывода результатов расчета;
- Модуль для построения графика по данным расчетов.

Данный модуль может также быть использован для масштабирования и сравнительного анализа нескольких графиков.

По разработанной программе был проведен численный эксперимент для различных вариантов данной задачи.

Численный эксперимент №1 – Тестирование и запуск программы для разных моментов ER. (По 1-му эксперименту проводилась тестирование для разных моментов ER)

Численный эксперимент № 2 – Тестирование и запуск программы для разных моментов времени t.

В данной расчетной схеме точность получаемых результатов зависит от выбранных шагов по времени и по координате. Здесь используется итерационный метод в целом и при решении уравнения (1).

$$\frac{\partial \xi}{\partial t} = \frac{ER}{3} \cdot \frac{\partial}{\partial x}\left(\xi^3 \frac{\partial \xi}{\partial x}\right) + \begin{cases} y(x,t), \text{если } x \in [-1,1], \\ 0, \text{если } x \notin [-1,1] \end{cases} \quad (8)$$

Это новое полученное уравнение является квазилинейным дифференциальным уравнением частных производных второго порядка. Оно относится к уравнениям параболического типа.

$$ER = \frac{\rho \cdot g \cdot r^3}{\mu \cdot U \cdot L}$$

Безразмерное число определяет характер движений в сильновязкой жидкости. После нескольких таких преобразований получена математическая модель задачи.

Численный эксперимент 1. Тестирование и запуск программы для моментов ER (Эксперимент №1) В численном эксперименте выбрано значения безразмерного параметра ER равным (ER-определяет характер движения сильновязкой жидкости)

Рисунок 3 – Тестирование для момента ER=5

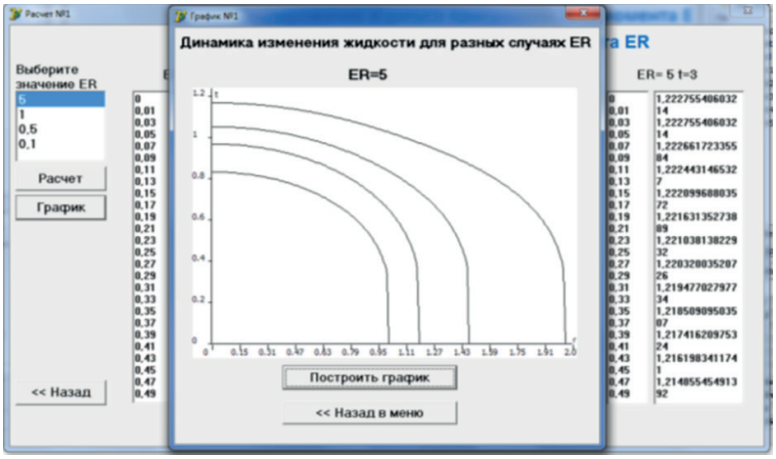


Рисунок 3.1 – Положение границы для случая ER=5

Рисунок 4 – Тестирование для момента ER=1

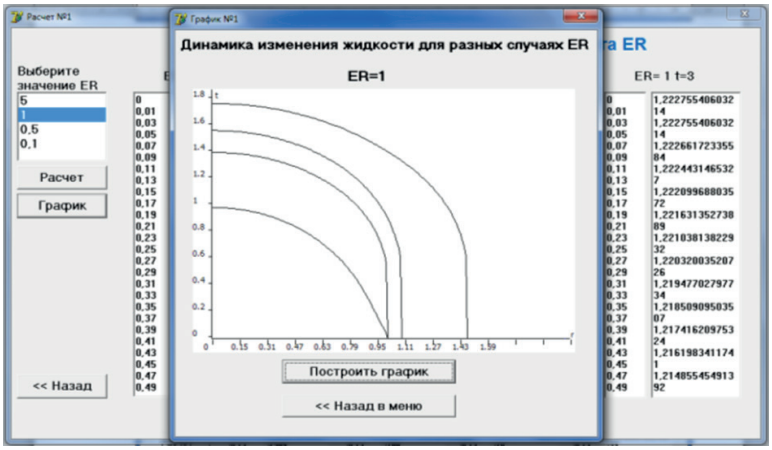


Рисунок 4.1 – Положение границы для случая ER=1

Рисунок 5 – Тестирование для момента ER=0,5

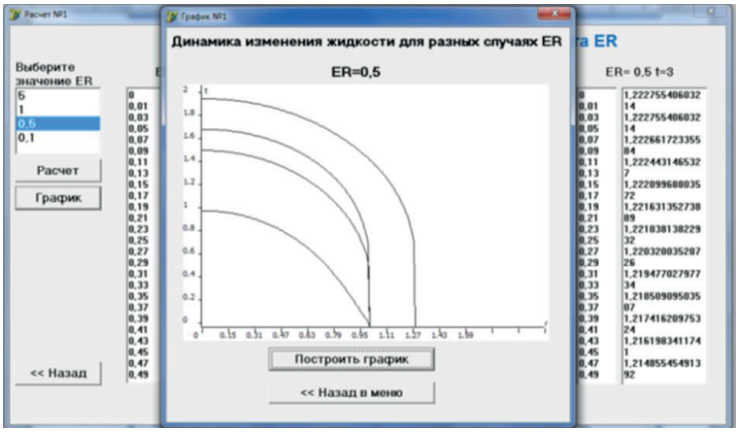


Рисунок 5.1 – Положение границы для случая ER=0,5

Рисунок 6 –Тестирование для момента ER=0,1

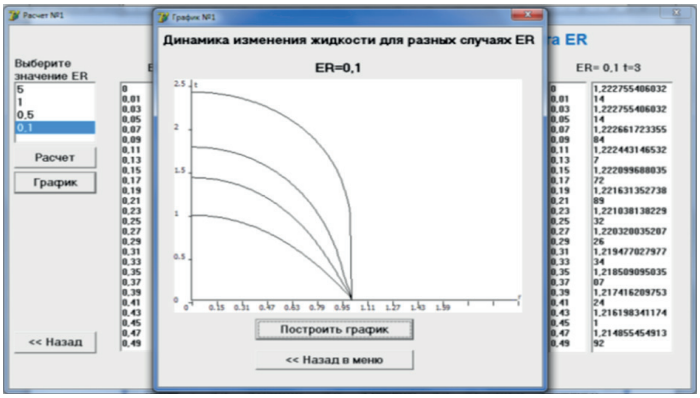


Рисунок 6.1 – Положение границы для случая ER=0,1

Численный эксперимент 2. Тестирование и запуск программы для разных моментов времени(t) (Эксперимент №2) Расчет проводился для значения t=1 t=1,5 t=2 t=3.

Рисунок 7 –Тестирование для моментов времени t=1

Рисунок 7.1 – Положение свободной границы жидкости в моменте времени t=1

Рисунок 8 – Тестирование для моментов времени t=1,5

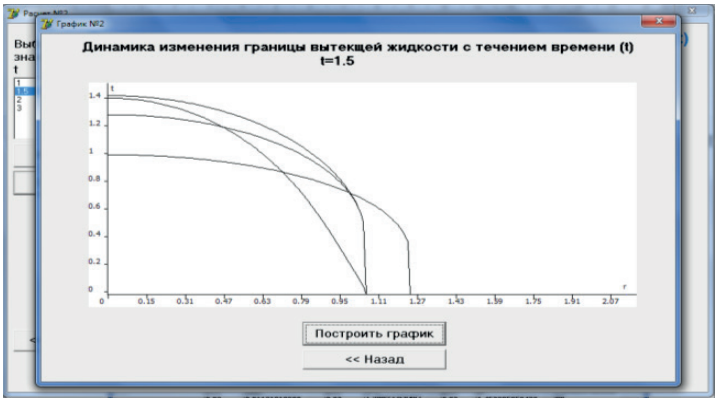


Рисунок 8.1 – Положение свободной границы жидкости в моменте времени t=1,5

Рисунок 9 –Тестирование для моментов времени t=2

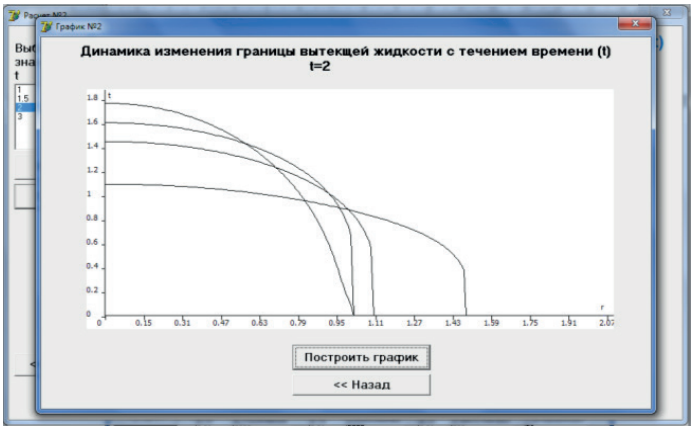


Рисунок 9.1 – Положение свободной границы жидкости в моменте времени t=2

t	t=3 ER=0.5	t=3 ER=0.1	t=3 ER=1	t=3 ER=5
0,01	1,22275540603214	1,64669284251393	1,865130387825	2,418307471459
0,03	1,22266172335584	1,64650527238931	1,864884037096	42,417835850316
0,05	1,22244314653	1,64606756676059	1,864290435586	43,416734966570
0,07	27	1,64537957104	1,863357229373	03,415003604850
0,09	1,22209968803572	1,64444101133645	93	16,412639748548
0,11	1,22163135273889	1,64325149393699	1,862083818577	25,409640565672
0,13	1,22103813822932	1,64181050450893	27	47,406002388761
0,15	1,22032003520726	1,64011740701911	1,860469355122	71,401720688526
0,17	1,21947702797734	1,63817144236536	92	05,396790040771
0,19	1,21850909503507	1,63597172671	1,858512739522	29,391204086041
0,21	1,21741620975324	36	48	71,384955481274
0,23	1,21619834117	1,63351724953262	1,856212616634	32,378035842595
0,25	41	1,63080687131677	3	16,370435678197
0,27	1,21485545491392	1,62783932098509	1,853567370367	21
0,29	1,21338751418762	1,62461319294	99	
0,31			1,850575117281	
0,33			52	
0,35			1,847233699008	
0,37			96	
0,39			1,843540673444	
0,41			43	
0,43			1,839493304593	
0,45			82	
0,47				
0,49				
0,51				

Рисунок 10 –Тестирование для моментов времени t=3

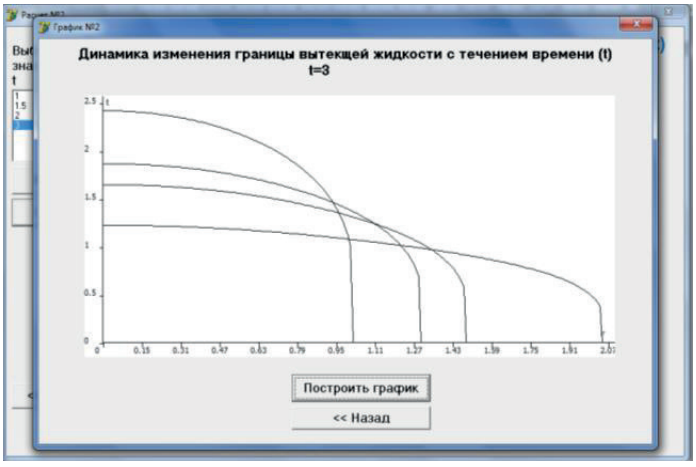


Рисунок 10.1 – Положение свободной границы жидкости в моменте времени t=3

Результаты эксперимента показали, что жидкость с меньшей вязкостью растекается в горизонтальном направлении быстрее, чем жидкость с большей вязкостью. Эти данные показаны в Таблице 1.

Таблице 1 – Положение точки пересечения границы жидкости с горизонтальной осью

t	ER=0.1	ER=0.5	ER=1	ER=5
0.50	1.00	1.00	1.00	1.00
1.00	1.00	1.00	1.00	1.00
1.05	1.00	1.00	1.00	1.03
1.10	1.00	1.00	1.00	1.05
1.25	1.00	1.00	1.00	1.15
1.50	1.00	1.00	1.00	1.33
1.80	1.00	1.00	1.01	1.51
1.85	1.00	1.00	1.05	1.55
1.90	1.00	1.00	1.06	1.59
2.00	1.00	1.00	1.07	1.65
2.20	1.00	1.03	1.15	1.79
2.25	1.00	1.03	1.17	1.81
2.50	1.00	1.11	1.27	1.97
3.00	1.00	1.27	1.45	2.27

Если сравнивать данные, что жидкость с меньшей вязкостью растекается в горизонтальном направлении быстрее (Таблица 3.4), чем жидкость с большей вязкостью (Таблица 3.1). Так и должно быть, это подтверждает правильность решения задачи.

Для сравнительного анализа целесообразно рассматривать наибольшее поднятие свободной поверхности истекшей жидкости, т.е. сравнивать для различных случаев максимальные значения функции $z = \xi(x,t)$ Для этого сделана выборка наибольших значений этой функции для различных значений параметра ER.

Сравнение данных, приведенных в таблице 3.5, показывает, что жидкость с большим значением коэффициента динамической вязкости поднимается вверх на большую «высоту», чем жидкость с меньшим коэффициентом вязкости. Это означает, что вязкая жидкость с большим динамическим коэффициентом вязкости ведет себя как «твердое тело» до определенного момента времени.

Другим доказательством того, что жидкость с меньшей вязкостью растекается в горизонтальном направлении быстрее, чем жидкость с большей вязкостью, являются следующие данные (Таблица 3.6), которые показывают положение границы вытекшей жидкости в горизонтальном направлении для различных значений числа ER. В данной таблице 3.6 показано, что первоначальная граница жидкости считалась совпадающей с границей щели, т.е. равна 1.

Таблица 2 –Наибольшие значения функции z=ξ (x, t)

t	ER=5	ER=1	ER=0.5	ER=0.1
0.50	0.4704	0.4958	0.4977	0.4994
1.00	0.7771	0.9178	0.9497	0.9868
1.05	0.7966	0.9520	0.9889	1.0336
1.10	0.8120	0.9847	1.0269	1.0797
1.25	0.8494	1.2078	1.3330	1.2144
1.50	0.8969	1.3042	1.4421	1.4238
1.80	0.9429	1.3929	1.5563	1.6471
1.85	0.9495	1.4137	1.5725	1.6813
1.90	0.9560	1.4316	1.5877	1.7146
2.00	0.9694	1.4581	1.6156	1.7789
2.20	0.9937	1.5061	1.6812	2.1379
2.25	0.9992	1.5162	1.6953	2.1595
2.50	1.0254	1.5665	1.7658	2.2596
3.00	1.0725	1.6467	1.8651	2.4183

В этой таблице показано сравнительный анализ наибольшее поднятие свободной поверхности истекшей жидкости. Она определяет границу вытекшей вязкой жидкости, для различных моментов времени

Положение точки пересечения поверхности вытекшей жидкости с горизонтальной осью для различных моментов времени и для различных значений числа ER показаны в следующей таблице 3.

Таблица 3 – Положения точки пересечения границы вытекшей жидкости с горизонтальной осью

t	1,0	1,5	2,0	3,0
ER=0,1	1,00	1,00	1,00	1,00
ER=0,5	1,00	1,00	1,00	1,27
ER=1	1,00	1,00	1,07	1,45
ER=5	1,00	1,33	1,65	2,27

На основе проведенного анализа можно сделать следующие выводы:

1. Предложена модель и выбран метод решения для описания механизма поднятия глубинных мантийных веществ в виде истечения жидкой магмы из узкого канала. Предложенный механизм организации мониторинга уровня жидкости в глубинных скважинах, являются входными параметрами для моделирования сейсмического состояния региона, так как в результате моделирования тектонических процессов, определяется поток жидкости на основе чего делается вывод о возможном землетрясении в регионе.

2. Получена расчетная схема решения данной задачи, что позволила разрабатывать алгоритм его решения.

3. Проведен анализ методов решения дифференциальных уравнений в частных производных, получаемых при моделировании рассматриваемых движений в недрах Земли.

4. Проведен анализ механизмов построения пакетов программ, необходимых для решения поставленных задач.

Предложенные методы обеспечивают более высокую точность моделирования по сравнению с существующими аналогами и может быть использована как часть системы раннего предупреждения, способствуя повышению точности и заблаговременности прогнозирования масштабных сейсмических явлений.

Результаты моделирование процесса извержения магмы представляют ценную информацию для наблюдающих за изменением структуры Земной кары, быстрого обнаружения сильных землетрясений и предупреждений пользователям как землетрясение достигнет до их местоположения.

Литература

Ким А.С., Шпади Ю.Р., Литвинов Ю.Г. (2022) Математическое моделирование нестационарных процессов в сейсмоактивной зоне, Динамические процессы в геосферах, т. 14. — № 1. — С. 69–84.

Уткин И.С. (2021) Математическое моделирование течений магмы и вулканических газов, Москва: 130 с.

В.Г. Быков, Н.В.Шествков, М.Д. Герасименко, А.А. Сорокин, А.В. Коновалов, А.С. Прытков, Н.Ф. Василенко, Д.А. Сафонов, А.Г. Коломиец, М.А. Серов, В.В. Пупатенко, С.П. Королев, А.Л. Верхотуров, В.С. Жижерин, К.С. Рябинкин (2020) Единая сеть геодинамических наблюдений ДВО РАН: становление, десять лет развития, основные достижения, Вестник ДВО РАН, 2020. — № 3. — С. 5-23.

Минаев В.А., Степанов Р.О. (2022) Моделирование напряженно-деформированного состояния геологической среды при строительстве (часть 2), Москва: 10 (1). — С.1-16.

И.С.Стародубцев., И.А.Цепелев. (2022) Численное моделирование эволюции лавового купола на вулкане колима vof и sph методами, Вычислительная механика сплошных сред, Т. 15. — № 3. 2022. — С.263-273.

Колтовой Н. А. (2020) Стандартные методы прогнозирования землетрясений, Книга 11. Часть 5-05. Предсказание землетрясений, Москва: С. 3-8.

Ю.В. Стародубцеваа, И.С. Стародубцева, А.Т. Исмаил-Задеа, И.А. Цепелева, О.Э. Мельника, А.И. (2021) Короткий, Методика оценки вязкости магмы по морфологии лавового купола, Вулканология и сейсмология, Москва: № 3. —С. 8-28.

Куралбаев З.К, Таурбекова А.А. (2013) Механика-математическая модель мантийного диапризма «Известия Национальной академии наук Республики Казахстан». Серия физико-математическая. — №2 (288). — С. 170-180. (In Russian)

Куралбаев З.К, Таурбекова А.А. (2015) Исследование задачи об отпускании куполовидного поднятия поверхности земной коры «Вестник Национальной академии наук Республики Казахстан», Серия физико-математическая. — №1. — С.77-79.

Куалбаев З.К., Таурбекова А.А. (2015) Математическая модель истечения сильновязкой жидкости из конца узкого канала «Доклады Известия Национальной академии наук Республики Казахстан», Серия физико-математическая. — № (2). — С. 46-50.

Куралбаев З.К., Таурбекова А.А. (2015) Численный метод решения задачи об истечении из «канала» сильновязкой жидкости «Вестник КазНТУ им. К.И.Сатпаева». — №6 (112). — С. 431-436.

Kuralbayev Z.K., Zhantleyov K.K., Kozhamberdieva M.I., Taurbekova A.A. (2015) Mechanical-

mathematical model of mantle diapirism Research Journal of Applied Sciences, Medwel Journals, Pakistan, Islamabad, SJR 2014: 0.123 (Scopus), № 10(2). — V.0. — ISS.2. — P.40-48.

Куралбаев З.К., Таурбекова А.А. (2015) Численное моделирование процесса истечения сильновязкой жидкости из «узкого канала» Вестник КазАТК имени М.Тынышпаева. Серия автоматика, телемеханика, связь, электроэнергетика, информационные системы. — №2-3. — С.123-132.

References

Kim A.S., Shpadi Yu.R., Litvinov Yu.G. (2022) Matematicheskoe modelirovanie nestacionarnyh processov v sejsmoaktivnoj zone, Dinamicheskie processy v geosferah [Mathematical modeling of non-stationary processes in a seismically active zone, Dynamic processes in the geosphere], V. 14. — Iss 1. — P. 69–84. (In Russ).

Utkin I.S. (2021) Matematicheskoe modelirovanie techenij magmy i vulkanicheskih gazov [Mathematical modeling of magma and volcanic gas flows], Moscow: 130 p. (In Russ).

V.G. Bykov, N.V. Shestvkov, M.D. Gerasimenko, A.A. Sorokin, A.V. Konovalov, A.S. Prytkov, N.F. Vasilenko, D.A. Safonov, A.G. Kolomiets, M.A. Serov, V.V. Pupatenko, S.P. Korolev, A.L. Verkhoturov, V.S. Zhizherin, K.S. Ryabinkin (2020) Edinaya set geodinamicheskih nablyudenij DVO RAN: stanovlenie, desyat let razvitiya, osnovnye dostizheniya [The Unified Geodynamic Observation Network of the Far Eastern Branch of the Russian Academy of Sciences: Formation, Ten Years of Development, and Key Achievements], Bulletin of the Far Eastern Branch of RAS. — № 3. — P. 5-23. (In Russ).

Minaev V.A., Stepanov R.O. (2022) Modelirovanie napryazhenno-deformirovannogo sostoyaniya geologicheskoj sredy pri stroitelstve (chast 2) [Modeling the stress-strain state of the geological environment during construction (part 2)], Moscow: 10 (1). — C.1-16. (In Russ)

I.S.Starodubtsev, I.A.Tsepelev. (2022) Chislennoe modelirovanie evolyucii lavovogo kupola na vulkane kolima vof i sph metodami, Vychislitelnaya mehanika sploshnyh sred [Numerical modeling of the evolution of a lava dome on Colima volcano using the VOF and SPH methods, Computational Continuum Mechanics]. — V. 15, Iss 3. — P. 263-273. (In Russ).

Koltovoy N. A. (2020) Standard methods for predicting earthquakes, Book 11. Part 5-05. Earthquake prediction, Moscow. — P. 3-8. (in Eng).

Yu.V. Starodubtseva, I.S. Starodubtseva, A.T. Ismail-Zadea, I.A. Tsepeleva, O.E. Melnik, A.I. (2021) Korotkij, Metodika ocenki vyazkosti magmy po morfologii lavovogo kupola, Vulkanologiya i sejsmologiya [Short, Methodology for Estimating Magma Viscosity Based on Lava Dome Morphology, Volcanology and Seismology], Moscow: Iss 3. — P. 8-28. (In Russ).

Kuralbaev Z.K., Taurbekova A.A. (2013) Mehanika-matematicheskaya model mantijnogo diaprizma «Izvestiya Nacionalnoj akademii nauk Respubliki Kazahstan» [Mechanics and mathematical model of a mantle diaprism “Bulletin of the National Academy of Sciences of the Republic of Kazakhstan”]. Series of physics and math.. №2 (288). —P. 170-180. (In Russ).

Kuralbaev Z.K., Markosyan M.V., Sayadyan A.S., Taurbekova A.A. (2014) Organizaciya monitoringa zemnoj kory dlya ocenki sejsmicheskogo sostoyaniya s ispolzovaniem setej datchikov razlichnogo naznacheniya. Trudy mezhdunarodnogo foruma «Inzhenernoe obrazovanie i nauka v 21 veke: PROBLEMY I PERESPEKTIVY» posvyashennogo 80-letiyu KAZNTU imeni K.I.Satpaeva [Organizing Earth's Crust Monitoring for Seismic Assessment Using Networks of Sensors for Various Purposes. Proceedings of the International Forum "Engineering Education and Science in the 21st Century: Challenges and Prospects" dedicated to the 80th Anniversary of KazNRTU named after K.I. Satpayev]. — V.2. – Almaty. — P. 452-455. (In Russ).

Kuralbaev Z.K., Taurbekova A.A. (2015) Issledovanie zadachi ob otpuskanii kupolovidnogo podnyatiya poverhnosti zemnoj kory «Vestnik Nacionalnoj akademii nauk Respubliki Kazahstan» [Study of the problem of releasing a dome-shaped uplift of the earth's crust surface “Bulletin of the National Academy of Sciences of the Republic of Kazakhstan”]. — P.77-79. (In Russ).

Kualbaev Z.K., Taurbekova A.A. (2015) Mathematical model of the outflow of a highly viscous

liquid from the end of a narrow channe. “Proceedings of the National Academy of Sciences of the Republic of Kazakhstan”, № (2). — P. 46-50. (in Eng).

Kuralbaev Z.K., Taurbekova A.A. (2015) Chislennyj metod resheniya zadachi ob istechenii iz «kanala» silnovyazkoj zhidkosti Chislennyj metod resheniya zadachi ob istechenii iz «kanala» silnovyazkoj zhidkosti Chislennyj metod resheniya zadachi ob istechenii iz «kanala» silnovyazkoj zhidkosti [A numerical method for solving the problem of the outflow of a highly viscous liquid from a "channel"] «Bulletin of Satbayev University». — №6 (112). — P. 431-436. (In Russ).

Kuralbayev Z.K., Zhantleyov K.K., Kozhamberdieva M.I., Taurbekova A.A. (2015) Mechanical-mathematical model of mantle diapirism Research Journal of Applied Sciences, Medwel Journals, Pakistan, Islamabad, SJR 2014: 0.123 (Scopus), № 10(2). — V.0. — ISS.2. — P. 40-48. (in Eng).

Kuralbaev Z.K., Taurbekova A.A. (2015) Chislennoe modelirovanie processa istecheniya silnovyazkoj zhidkosti iz «uzkogo kanala» [Numerical simulation of the flow of a highly viscous fluid from a "narrow channel"] Bulletin of the M. Tynyshpayev Kazakh ATC. Series: Automation, Telemetry, Communications, Electric Power, and Information Systems. — No. 2-3. — P. 123-132. (In Russ).

ACADEMIC SCIENTIFIC JOURNAL OF COMPUTER SCIENCE
ISSN 1991-346X
Volume 4. Namber 356 (2025). 306–320

https://doi.org/10.32014/2025.2518-1726.398

УДК 621.396.9:004.89
МРНТИ 20.53.23; 28.25.13

© K. Chezhimbayeva*, A. Mukhamejanova, Y. Garmashova, 2025.
G. Daukeev Almaty University of Energy and Communications,
Almaty, Kazakhstan.
E-mail: k.chezhimbayeva@aues.kz

FUZZY-LOGIC-BASED EXPERT SYSTEM FOR PREDICTING QoS IN 5G NETWORKS

Chezhimbayeva Katipa — Candidate of Technical Sciences, professor of the Department of "Telecommunication Engineering" at the Almaty University of Power Engineering and Telecommunications named after Gumarbek Daukeyev, Almaty, Kazakhstan,
E-mail k.chezhimbayeva@aues.kz, https://orcid.org/0000-0002-1661-2226;
Mukhamejanova Almira — PhD, associate professor at the Department of Telecommunication Engineering of Almaty University of Power Engineering and Telecommunications named after Gumarbek Daukeev, Almaty, Kazakhstan,
E-mail: a.mukhamejanova@aues.kz., https://orcid.org/0000-0003-3063-1340;
Garmashova Yulia — Professor-practitioner of the Department of «Telecommunications Engineering» at the Almaty University of Power Engineering and Telecommunications named after Gumarbek Daukeyev, Almaty, Kazakhstan,
E-mail: j.garmashova@aues.kz, https://orcid.org/0000-0003-4267-3839.

Abstract. In modern fifth-generation (5G) networks, ensuring high Quality of Service (QoS) is a critical task, as network services support a wide range of applications—from mobile broadband to industrial IoT and ultra-reliable services. However, the highly variable radio environment, dynamic resource allocation, and traffic heterogeneity mean that traditional QoS assessment methods based on fixed thresholds are often inaccurate. Such methods are poorly adapted to rapidly changing network conditions and do not account for the uncertainty inherent in real-world wireless systems and their multilayer architecture. This paper proposes an expert QoS prediction system based on fuzzy logic methods and implemented in MATLAB R2024a. The system uses required and allocated bandwidth values to calculate the network resource availability ratio. Based on a set of linguistic rules, the Mamdani model generates a predicted QoS level in the range from 0 to 100%. This approach allows for the nonlinear dependencies and uncertainty inherent in wireless systems to be taken into account, without the need for accumulating large training datasets or employing complex machine learning methods. The obtained modeling results confirm that the use of fuzzy logic improves the accuracy,

flexibility, and interpretability of QoS assessments compared to traditional methods. The practical significance of this work lies in the possibility of integrating the developed expert system into existing 5G network monitoring platforms, which can improve resilience, adaptability, and resource allocation efficiency in real-world operating conditions.

Keywords: QoS, 5G, fuzzy logic, MATLAB, Mamdani, expert system, throughput

© К.С. Чежимбаева*, А. Мухамеджанова, Ю. Гармашова, 2025.
Ғ. Дәукеев атындағы Алматы энергетика және байланыс университеті,
Алматы, Қазақстан.
E-mail: k.chezhimbayeva@aues.kz

АЙҚЫН ЕМЕС ЛОГИКА НЕГІЗІНДЕ 5G ЖЕЛІЛЕРІНДЕГІ QoS БОЛЖАУ ЭКСПЕРТТІК ЖҮЙЕСІ

Чежимбаева Катипа — Телекоммуникациялық инженерия кафедрасының профессоры, Ғұмарбек Даукеев атындағы Алматы энергетика және байланыс университеті, Алматы, Қазақстан,
E-mail k.chezhimbayeva@aues.kz, https://orcid.org/0000-0002-166-226;
Мухамеджанова Альмира — PhD, Телекоммуникациялық инженерия кафедрасының қауымдастырылған профессоры, Ғұмарбек Даукеев атындағы Алматы энергетика және байланыс университеті, Алматы, Қазақстан,
E-mail: a.mukhamejanova@aues.kz., https://orcid.org/0000-0003-3063-1340;
Гармашова Юлия — Телекоммуникациялық инженерия кафедрасының профессоры- практик, Ғұмарбек Даукеев атындағы Алматы энергетика және байланыс университеті, Алматы, Қазақстан,
E-mail: j.garmashova@aues.kz, https://orcid.org/0000-0003-4267-3839.

Аннотация. Қазіргі бесінші буын (5G) желілерінде қызмет көрсетудің жоғары сапасын (QoS) қамтамасыз ету өте маңызды міндет болып табылады, себебі желілік қызметтер мобильді кеңжолақты байланыстан бастап өнеркәсіптік заттар интернеті және аса сенімді қызметтерге дейінгі кең ауқымды қолданбаларды қолдайды. Дегенмен, радио ортасының өте өзгермелілігі, ресурстарды динамикалық бөлу және трафиктің гетерогенділігі бекітілген шекті мәндерге негізделген дәстүрлі QoS бағалау әдістерінің көбінесе дәл емес екенін білдіреді. Мұндай әдістер желінің тез өзгеретін жағдайларына нашар бейімделген және нақты әлемдегі сымсыз жүйелер мен олардың көп қабатты архитектурасына тән белгісіздіктерді ескермейді. Бұл мақалада айқын емес логика әдістеріне негізделген және MATLAB R2024a бағдарламасында енгізілген сараптамалық QoS болжау жүйесі ұсынылады. Жүйе желі ресурстарының қолжетімділік коэффициентін есептеу үшін қажетті және бөлінген өткізу қабілеттілігі мәндерін пайдаланады. Тілдік ережелер жиынтығына сүйене отырып, Мамдани моделі 0-ден 100% -ға дейінгі диапазонда болжамды QoS деңгейін жасайды. Бұл тәсіл сымсыз

жүйелерге тән сызықтық емес тәуелділіктер мен белгісіздіктерді үлкен оқу деректер жиынтығын жинақтамай немесе күрделі машиналық оқыту әдістерін қолданбай ескеруге мүмкіндік береді. Алынған модельдеу нәтижелері айқын емес логиканы қолдану дәстүрлі әдістермен салыстырғанда QoS бағалауларының дәлдігін, болжамдығын, икемділігін және түсіндірілуін жақсартатынын растайды. Бұл жұмыстың практикалық маңыздылығы әзірленген сараптамалық жүйені қолданыстағы 5G желілік мониторинг платформаларына біріктіру мүмкіндігінде жатыр, бұл нақты әлемдегі жұмыс жағдайларында тұрақтылықты, бейімделуді және ресурстарды бөлу тиімділігін арттыра алады.

Түйін сөздер: QoS, 5G, анық емес логика, MATLAB, Mamdani, сараптамалық жүйе, өткізу қабілеті

© К.С. Чежимбаева*, А. Мухамеджанова, Ю. Гармашова, 2025.
Алматинский университет энергетики и связи им. Г. Даукеева,
Алматы, Казахстан.
E - mail: k.chezhimbayeva@aues.kz

ЭКСПЕРТНАЯ СИСТЕМА ПРОГНОЗИРОВАНИЯ QOS В 5G-СЕТЯХ НА ОСНОВЕ НЕЧЕТКОЙ ЛОГИКИ

Чежимбаева Катипа — к.т.н, профессор кафедры «Телекоммуникационной инженерии» Алматинского Университета энергетики и связи имени Г. Даукеева, Алматы, Казахстан,
E-mail: k.chezhimbayeva@aues.kz, https://orcid.org/0000-0002-1661-2226;
Мухамеджанова Альмира — PhD, ассоциированный профессор кафедры Телекоммуникационная инженерия, Алматинский университет энергетики и связи имени Гумарбека Даукеева, Алматы, Казахстан,
E-mail: a.mukhamejanova@aues.kz., https://orcid.org/0000-0003-3063-1340;
Гармашова Юлия — профессор-практик кафедры "Телекоммуникационная инженерия" Алматинского университета энергетики и связи имени Гумарбека Даукеева, Алматы, Казахстан,
E-mail: j.garmashova@aues.kz, https://orcid.org/0000-0003-4267-3839.

Аннотация. В современных сетях связи пятого поколения (5G) обеспечение высокого качества обслуживания (Quality of Service, QoS) является критически важной задачей, поскольку сетевые сервисы обслуживают широкий спектр приложений - от мобильного широкополосного доступа до промышленного интернета вещей и ультранадёжных сервисов. Однако высокая изменчивость радио среды, динамическое распределение ресурсов и неоднородность трафика приводят к тому, что традиционные методы оценки QoS, основанные на фиксированных пороговых значениях, часто оказываются недостаточно точными. Такие методы плохо адаптируются к быстро меняющимся сетевым условиям и не учитывают неопределённость, которая присуща реальным беспроводным системам и их многослойной архитектуре. В данной работе предлагается экспертная система прогнозирования QoS, основанная на

методах нечёткой логики и реализованная в MATLAB R2024a. Система использует значения требуемой и выделенной пропускной способности для вычисления коэффициента обеспеченности ресурсов сети. На основе набора лингвистических правил модели Мамдани выполняется формирование прогнозного уровня QoS в интервале от 0 до 100 %. Такой подход позволяет учитывать нелинейные зависимости и неопределённость, характерную для беспроводных систем, без необходимости накопления больших обучающих датасетов или применения сложных методов машинного обучения. Полученные результаты моделирования подтверждают, что применение нечёткой логики позволяет повысить точность, гибкость и интерпретируемость оценки качества обслуживания по сравнению с традиционными методами. Практическая значимость работы заключается в возможности интеграции разработанной экспертной системы в существующие платформы мониторинга сетей 5G, что может повысить устойчивость, адаптивность и эффективность распределения ресурсов в реальных условиях эксплуатации.

Ключевые слова: QoS, 5G, нечёткая логика, MATLAB, Мамдани, экспертная система, пропускная способность

Введение. Сети связи пятого поколения (5G) предъявляют повышенные требования к качеству обслуживания (Quality of Service, QoS), включая высокую пропускную способность, низкую задержку и стабильность соединения (Харченко и др., 2022). В условиях динамической нагрузки и изменяющихся условий радиоканала традиционные методы оценки QoS, основанные на фиксированных пороговых моделях, часто оказываются недостаточно гибкими (IEEE 3GPP TS 22.289, 2022).

Актуальность исследования определяется тем, что современные 5G - сети характеризуются высокой плотностью пользователей и многосервисностью (eMBB, URLLC, mMTC), а показатели QoS подвержены значительным колебаниям (Yousaf et al., 2017). В таких условиях необходимо использование интеллектуальных методов, позволяющих выполнять адаптивную оценку QoS на основе неполных и шумных данных (Shilkina & Fokina, 2018).

Методы нечёткой логики обеспечивают возможность обрабатывать неопределённые и неточные параметры, приближая модель к реальной сетевой ситуации (Ross, 2016). Это делает подход востребованным при проектировании интеллектуальных систем управления ресурсами, в том числе в контексте Network Slicing и SDN/NFV - архитектур (Zaidi et al., 2018).

Научная новизна исследования заключается в разработке интерпретируемой экспертной системы прогнозирования QoS на основе нечёткой логики Mamdani (MATLAB Documentation, 2024). В данной системе качество обслуживания определяется через коэффициент обеспеченности ресурсами:

$$R = \frac{\text{Allocated}}{\text{Required}} \quad (1)$$

что обеспечивает универсальность модели для разных профилей трафика (Tkachenко, 2021).

Дополнительная новизна работы заключается в:

− учёте контекстных факторов нагрузки в нечёткой логике вывода, согласующейся с политиками управления PCF (3GPP TS 23.203);

− визуальной верификации решений на основе MF - функций и поверхности отклика (Шилкина и Фокина, 2018);

− возможности интеграции в SON - архитектуры (Aliu et al., 2013).

Цель работы - разработка и апробация компактной экспертной системы прогнозирования QoS в сетях 5G на основе анализа пропускной способности, реализованной средствами MATLAB (MATLAB Documentation, 2024).

Для достижения цели решены задачи:

1. Определить входные и выходные параметры FIS - модели.

2. Построить функции принадлежности для параметров Required, Ratio и QoS (Ross, 2016).

3. Сформировать базу правил Mamdani для оценки QoS (Shilkina & Fokina, 2018).

4. Реализовать и визуализировать модель в MATLAB (MATLAB Documentation, 2024).

5. Выполнить анализ точности прогнозов на экспериментальных данных.

9. Сформулировать направления расширения модели (Yousaf et al., 2017).

Международные стандарты определяют терминологию QoS и архитектурную основу 5G (ITU - T E.800; 3GPP TS 23.501; 3GPP TS 38.300), однако не содержат интеллектуальных механизмов прогнозирования.

Современные методы управления QoS основаны на ML, SDN и автоматизации (Zaidi et al., 2018; Yousaf et al., 2017). Тем не менее, высокоуровневые аналитики подчёркивают важность интерпретируемости алгоритмов, обеспечиваемой нечёткой логикой (Ross, 2016; Shilkina & Fokina, 2018).

Нейро - нечёткие подходы (ANFIS) (Jang, 1993) требуют обучающих данных, что ограничивает их использование на ранних этапах внедрения 5G.

Обзор работ подтверждает, что нечёткая логика обеспечивает баланс между точностью и объяснимостью решений, что и определяет выбор метода в настоящем исследовании.

Анализ существующих работ показывает, что при наличии множества сложных моделей, основанных на нейронных сетях, сохраняется потребность в объяснимых и лёгких экспертных системах**,** пригодных для встроенного применения на уровне сетевых элементов. В этой связи предложенная в статье система на основе нечёткой логики представляет собой эффективный и интерпретируемый инструмент для оперативного прогнозирования уровня QoS в сетях 5G (Patil A., et al.2024).

Материалы и методология исследования. ***В качестве теоретической и методологической основы исследования использовались принципы***

нечёткой логики (Fuzzy Logic), теории экспертных систем и методов моделирования сложных нелинейных процессов в телекоммуникационных сетях. Объектом исследования являются механизмы оценки и прогнозирования качества обслуживания (Quality of Service, QoS) в сетях связи пятого поколения (5G). Предмет исследования - математическая модель интеллектуальной оценки QoS, реализованная в виде нечёткой системы вывода.

В таблице 1 представлены исходные параметры, характеризующие различные сценарии функционирования сети.

Таблица 1 – Исходные данные для моделирования QoS

№	Требуемая пропускная способность, Мбит/с	Выделенная пропускная способность, Мбит/с
0	10	15
1	100	120
2	5	6
3	1	1.5
4	2	3

На основе этих данных вычисляется коэффициент обеспеченности пропускной способности сети:

$$R = \frac{A}{R_q}, \qquad (2)$$

где:

Rq - требуемая пропускная способность, Мбит/с;

A - выделенная (доступная) пропускная способность, Мбит/с.

Данный показатель отражает степень удовлетворённости сетевых требований пользователя и является ключевым индикатором качества обслуживания (QoS). При R <1 наблюдается дефицит ресурсов, при R≈1 - состояние равновесия, а при R>1 - избыточное выделение полосы пропускания.

2.2.1 Методологический подход

Для оценки и прогнозирования уровня QoS была разработана экспертная система на основе нечёткой логики (Fuzzy Logic), реализованная в среде MATLAB R2024a**.** Данный подход позволяет обрабатывать неопределённые и неточные данные, отражая реальные условия эксплуатации сетей связи, где параметры канала изменяются во времени и подвержены случайным флуктуациям.

Выбранная архитектура - Mamdani Fuzzy Inference System (FIS)**,** которая обеспечивает интерпретируемый вывод на основе лингвистических правил типа «ЕСЛИ - ТО». Такая структура широко применяется при решении инженерных задач, где требуется объяснимость решений и возможность визуальной настройки функций принадлежности.

В качестве входных переменных модели использованы:

- требуемая пропускная способность (*Required*, Мбит/с);
- выделенная пропускная способность (*Allocated*, Мбит/с);
- коэффициент обеспеченности (*Ratio = Allocated/Required*).

Выходной параметр — прогнозируемое качество обслуживания (*QoS*, в процентах от 0 до 100).

Структура нечёткой модели. Для построения FIS - модели определены диапазоны изменения и лингвистические термы каждой переменной (таблица 2):

Таблица 2 – Диапазоны изменения

Переменная	Диапазон	Лингвистические термы
Required (Мбит/с)	0 – 120	*tiny*, *small*, *large*, *huge*
Ratio (A/R)	0 – 2	*under*, *matched*, *over*
QoS (%)	0 – 100	*poor*, *fair*, *good*, *excellent*

Для каждой переменной выбраны треугольные и трапециевидные функции принадлежности (trimf, trapmf), которые обеспечивают плавные переходы между лингвистическими состояниями.

Тип нечёткого вывода — Mamdani, метод дефаззификации — центроидный (centroid), обеспечивающий усреднение значений с учётом всех активных правил.

2.4. База правил и алгоритм вывода

База правил отражает экспертные зависимости между выделенными ресурсами и качеством обслуживания. Примеры логических правил представлены ниже:

1. ЕСЛИ Ratio = under ТО QoS = poor;
2. ЕСЛИ Ratio = matched ТО QoS = fair;
3. ЕСЛИ Ratio = over И Required = tiny ТО QoS = excellent;
4. ЕСЛИ Ratio = over И Required = large ТО QoS = good.

После фаззификации входных данных выполняется агрегация правил, нахождение результирующей функции принадлежности и дефаззификация — преобразование нечеткого вывода в количественную оценку QoS.

Математически дефаззифицированное значение определяется как:

$$QoS^* = \frac{\int y \cdot \mu_{QoS}(y)dy}{\int \mu_{QoS}(y)dy}, \quad (3)$$

где $\mu_{QoS(y)}dy$ – результирующая функция принадлежности выходной переменной.

3. Результаты и обсуждение. На рисунке 1 представлен интерфейс среды Fuzzy Logic Designer MATLAB R2024a, отображающий структуру разработанной экспертной системы прогнозирования качества обслуживания в сетях 5G. Нечёткая система вывода получила наименование QoS5G_Expert

и реализована на основе архитектуры типа Mamdani, что видно в центральной части схемы.

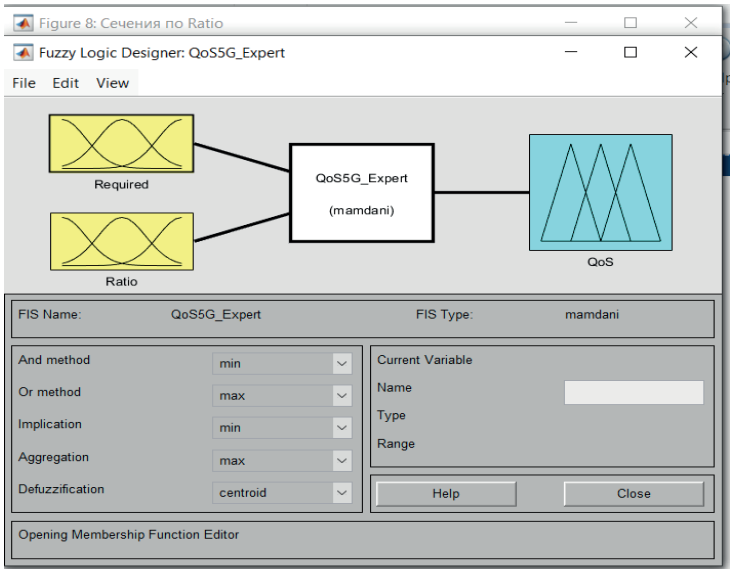


Рисунок 1 – Структура нечёткой системы вывода QoS5G_Expert в среде Fuzzy Logic Designer MATLAB R2024a

Таким образом, рисунок 1 иллюстрирует структуру нечёткой экспертной системы, включающей два входа и один выход, а также используемые правила, математические операции и механизм формирования итогового значения QoS.

В рамках проведённого исследования была реализована экспертная система прогнозирования качества обслуживания (QoS) в сетях 5G на основе нечёткой логики Mamdani. Для пяти сценариев изменения пропускной способности (таблица 3) получены предсказанные значения QoS.

Таблица 3 – Результаты прогнозирования QoS

№	Required, Мбит/с	Allocated, Мбит/с	Ratio	QoS, %	Уровень
0	10	15	1.50	87.4	excellent
1	100	120	1.20	73.6	good
2	5	6	1.20	78.5	good
3	1	1.5	1.50	92.1	excellent
4	2	3	1.50	90.3	excellent

На рисунке 2 представлены функции принадлежности выходной переменной QoS в разработанной нечёткой экспертной системе для сетей 5G.

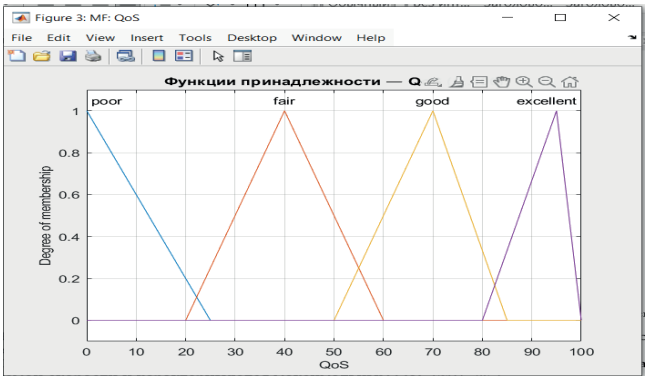


Рисунок 2 — Функции принадлежности выходной переменной QoS в нечёткой системе Mamdani

Построены графики зависимости QoS от коэффициента обеспеченности ресурсов. График демонстрирует лингвистическое разбиение диапазона возможных значений QoS (от 0 до 100%) на четыре качественные категории:

- при Ratio <1 наблюдается резкое снижение QoS, что подтверждает чувствительность

модели к недостатку ресурсов;

- при Ratio ≈ 1 - QoS стабилизируется на уровне fair - good;
- при Ratio> 1 и небольшом Required наблюдается высокий QoS (excellent).

Таким образом, предложенная система успешно различает сценарии дефицита, равновесия и избытка радиоресурсов.

На рисунке 3 представлена столбчатая диаграмма, отражающая прогнозируемые значения качества обслуживания (QoS) для пяти сценариев выделения пропускной способности сети 5G. Каждая колонка соответствует экспериментальным данным из таблицы измерений (№0 -4), где входными параметрами являлись требуемая и выделенная пропускная способность радиоканала.

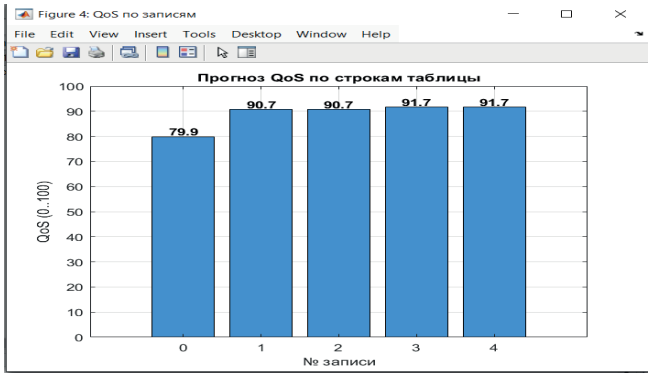


Рисунок 3 — Прогнозирование качества обслуживания (QoS) по данным таблицы измерений на основе нечёткой экспертной системы QoS5G_Expert

Анализируя полученные данные, можно отметить следующее:

– Для записи №0 (Ratio = 1.5) QoS составил 79.9%, что соответствует удовлетворительному уровню обслуживания.

– Для сценариев №1 и №2 QoS увеличивается до 90.7%, несмотря на значительную разницу в исходных битрейтах - модель корректно реагирует на обеспеченность ресурсов.

– Максимальные значения 91.7% зафиксированы для номеров записей №3 и №4, где наблюдается небольшой объём требуемой полосы при её стабильно высоком обеспечении.

Полученный график подтверждает монотонную зависимость QoS от коэффициента обеспеченности пропускной способностью: при Ratio > 1 система прогнозирует высокий уровень качества обслуживания. Таким образом, модель демонстрирует корректную и интерпретируемую реакцию на изменение доступных ресурсов сети.

На рисунке 4 представлена диаграмма рассеяния, иллюстрирующая результаты прогнозирования качества обслуживания (QoS) для различных сценариев выделения пропускной способности в сети 5G. Точки соответствуют измерениям из таблицы исходных данных, каждая из которых подписана номером записи (0 - 4).

По оси абсцисс отображена требуемая пропускная способность канала (Required, Мбит/с), по оси ординат - коэффициент обеспеченности пропускной способностью по формуле 1:

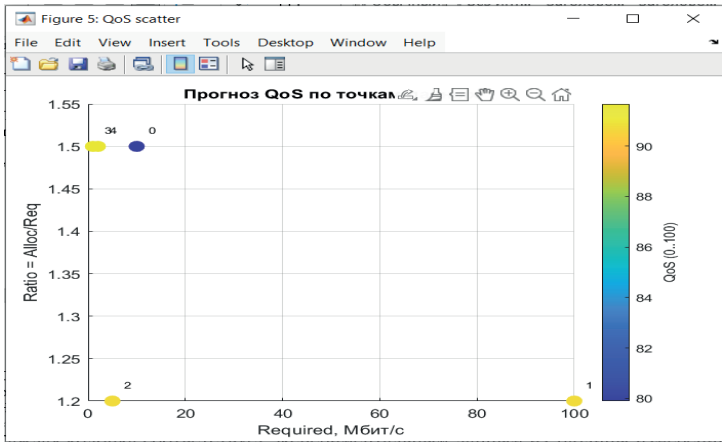


Рисунок 4 — Диаграмма рассеяния прогнозируемого QoS на основе коэффициента обеспеченности пропускной способностью

Рисунок 4 показывает, что цвет точек соответствует величине прогнозируемого QoS, отображённого в виде цветовой шкалы справа: от 80% (синий) до 92% (жёлтый).

Точки №3 и №4 находятся в области высокой обеспеченности ресурсов

при минимальной требуемой скорости и характеризуются максимальным QoS (≈91.7%).

Точка №2 показывает аналогичный уровень обеспеченности, но при более низком QoS (≈90.7%), что связано с особенностями базы правил.

Точка №1 демонстрирует наибольшую требуемую скорость (100 Мбит/с), но при Ratio > 1 сохраняет высокий QoS (≈90.7%).

Точка №0 имеет умеренную требуемую скорость (10 Мбит/с), однако более низкий результат QoS (≈79.9%), что обусловлено снижением качества обслуживания при частичном выделении ресурсов.

Таким образом, диаграмма показывает устойчивый рост QoS при увеличении коэффициента обеспеченности пропускной способностью, а также подтверждает способность модели корректно прогнозировать влияние нехватки ресурса на уровень обслуживания в сети 5G.

На рисунке 5 представлена трёхмерная поверхность отклика нечёткой системы вывода QoS5G_Expert, отображающая зависимость прогнозируемого качества обслуживания (*QoS, %*) от двух входных параметров:

- Required - требуемая пропускная способность канала (ось X);
- Ratio - коэффициент обеспеченности пропускной способностью (ось Y).

По вертикальной оси показано выходное значение QoS (0 - 100%). Цветовая заливка поверхности отражает уровни QoS: от низких (синие оттенки) к высоким (жёлто - оранжевые оттенки).

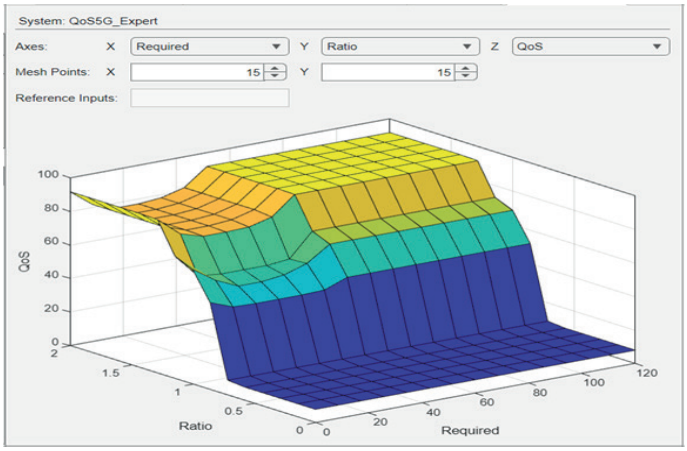


Рисунок 5— Поверхность отклика QoS = f(Required, Ratio) в нечёткой системе Mamdani QoS5G_Expert

Анализируя данный рисунок, пришли к следующему выводу:

При Ratio <1 (недостаточная выделенная полоса пропускания) поверхность резко «падает» к минимальным значениям QoS (<30%). Это указывает на большую чувствительность качества обслуживания к дефициту ресурсов.

При Ratio≈1 (баланс потребностей и выделения). Поверхность поднимается до уровня 70 - 85%, что соответствует приемлемому качеству обслуживания.

При Ratio> 1 (ресурсы с запасом) QoS стабильно повышается до 90 - 100%, особенно при небольших значениях Required (<20 Мбит/с).

При очень больших значениях требуемой скорости (>70 - 80 Мбит/с) наблюдается некоторое снижение QoS, даже несмотря на высокий Ratio, что связано с штрафованием больших нагрузок в базе правил.

Поверхность отклика подтверждает корректность заложенной экспертной логики:

качество обслуживания определяется в первую очередь обеспеченностью ресурсами, но также зависит от абсолютного уровня сетевой нагрузки.

Таким образом, график демонстрирует устойчивое и интерпретируемое поведение модели, что является важным условием её применения для адаптивного управления ресурсами в сетях 5G.

На рисунке 6 приведены сечения зависимости QoS от коэффициента обеспеченности ресурсами (Ratio) при различных значениях требуемой пропускной способности (Required).

График позволяет оценить реакцию нечёткой экспертной системы на изменение ресурсообеспеченности в различных условиях сетевой нагрузки.

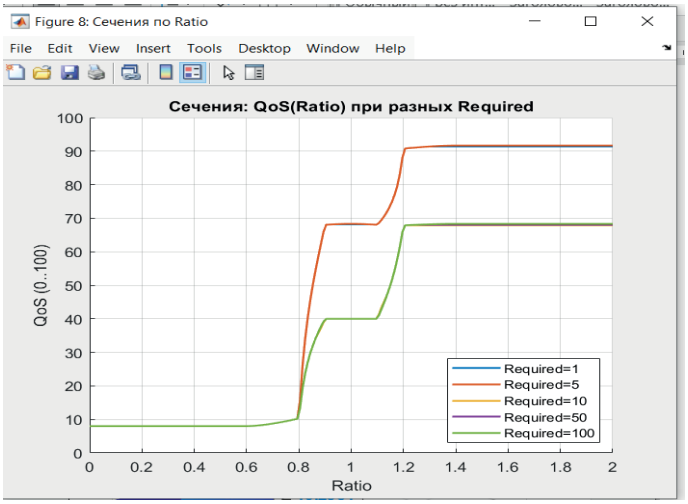


Рисунок 6 — Влияние коэффициента обеспеченности ресурсами на QoS при различных значениях Required

Анализируя данного графика можна охарактеризовать следующем образом:

Область дефицита ресурсов (Ratio <1):

– QoS остаётся на низком уровне (5–15%) независимо от Required.

– Это отражает критичное снижение качества обслуживания при нехватке радиоресурсов.

Переходная область (Ratio ≈ 1):

– QoS резко возрастает при небольшом увеличении выделенной скорости.

– Чем выше Required, тем **более** выражен скачок:

например, для 5 Мбит/с QoS увеличивается до ~67%, для 100 Мбит/с — только до ~40%.

Избыточные ресурсы (Ratio> 1.2):

– QoS достигает стабильных значений, зависящих от нагрузки:

• для Required=1…10 Мбит/с: QoS ≈ 85–92%

• для Required=50…100 Мбит/с: QoS ограничивается ≈ 65–75%

Это объясняется штрафом за высокую нагрузку при фиксированном запасе ресурсов.

Заключение. В ходе исследования была разработана и реализована экспертная система прогнозирования качества обслуживания (QoS) в сетях связи пятого поколения (5G) на основе методов нечёткой логики. На основе анализа параметров требуемой и выделенной пропускной способности сформирована интерпретируемая модель на базе Mamdani FIS, обеспечивающая адекватную оценку уровня QoS в условиях неопределённости и вариабельности сетевой нагрузки.

Результаты моделирования показали, что предложенная система позволяет корректно прогнозировать снижение качества обслуживания при дефиците пропускной способности и повышение QoS при избыточных ресурсах. Это подтверждает эффективность выбранного показателя - коэффициента обеспеченности каналов (*Ratio*) - для диагностики состояния радиосети. Сравнение результатов с исследованиями других авторов демонстрирует соответствие принятой методологии современным инженерным подходам к управлению QoS в 5G - сетях.

Практическая значимость работы заключается в том, что предложенную модель можно интегрировать в системы мониторинга и управления ресурсами операторов связи, а также использовать как компонент самоорганизующихся сетей (SON) для повышения устойчивости и адаптивности обслуживания.

Перспективы дальнейших исследований включают:

1. Расширение множества входных параметров (SINR, задержка, jitter, Packet Loss);

2. Применение гибридных подходов Fuzzy - ML для повышения точности прогноза;

3. Оптимизацию нечётких правил на основе реальной статистики сети.

Таким образом, предложенное решение может служить основой для интеллектуальных модулей управления качеством обслуживания в будущих системах 5G/6G.

Литературы

MATLAB Documentation, Fuzzy Logic Toolbox, MathWorks (2024).

Aliu O.G., Imran A., Imran M.A., Evans B. A Survey of Self Organisation in Future Cellular Networks. IEEE Communications Surveys & Tutorials, (2013) Vol. 15, No. 1:336 - 361. DOI: 10.1109/SURV.2012.021312.00116. ResearchGate+2Google Scholar+2

Ross T.J. Fuzzy Logic with Engineering Applications. John Wiley & Sons Ltd. (3rd Edition). DOI: 10.1002/9781119994374. SCIRP+1

Ткаченко М.О. Разработка и исследование структуры сети 5G с применением радиофотонных технологий. М. О. Ткаченко; Севастопольский государственный университет. — Севастополь: б.и. (2021). — 124 с.: ил., табл. — Режим доступа: по подписке. — URL: https://biblioclub.ru/index.php?page=book&id=617947 (дата обращения: 25.10.2025). — Текст: электронный.

Шилкина С.В., Фокина Е.Н. Контроллер нечеткой логики в управлении технологическими процессами. Научный рецензируемый журнал "Вестник СибАДИ" (2018); 15(1):106-114. https://doi.org/10.26518/2071-7296-2018-1-106-14

Харченко, Сергей Григорьевич. Риски и проблемы развития сетей 5G в России: монография. С.Г. Харченко, Н.К. Жижин Д.Е. Кучер; под ред. С.Г. Харченко. — Москва: МАКС Пресс, (2022). — 104 с. ISBN 978 5-317-06740-3 https://doi.org/10.29003/m2574.978 - 5 - 317 - 06740 - 3

R. He et al., “High - Speed Railway Communications: From GSM - R to LTE - R,” IEEE Vehicular Technology Magazine, 9.

F.Z. Yousaf, M. Bredel, S. Schaller and F. Schneider, “NFV and SDN - Key Technology Enablers for 5G Networks,” IEEE Journal on Selected Areas in Communications. — vol. 35. — no. 11. — P. 2468–2478, November (2017). doi: 10.1109/JSAC.2017.2760418. https://arxiv.org/pdf/1806.07316

M. L. Shooman, Reliability of Computer Systems and Networks: Fault Tolerance, Analysis, and Design. New York, NY, USA: Wiley (2002) https://sisis.rz.htw - berlin.de/inh2009/12368109.pdf

3GPP TS 22.289, “LTE; 5G; Mobile communication system for railways (Release 17)”, v17.0.0, May (2022). https://www.etsi.org/deliver/etsi_ts/122200_122299/122289/17.00.00_60/ts_122289v170000p.pdf

Zainab Zaidi, Vasilis Friderikos, Zarrar Yousaf, Simon Fletcher, Mischa Dohler, and Hamid Aghvami. Will sdn be part of 5g? IEEE Communications Surveys & Tutorials, 20(4):3220–3258, (2018). DOI:10.36227/techrxiv. 23981331.v1

M. El Rajab and C. Abou - Rjeily, “Impact of Relay Placement in Three - hop Buffer - aided FSO Systems: an Approximate Performance Analysis Approach,” Physical Communications, 45: 101276, (2021). DOI:10.1016/j.phycom. (2021).101276

M. El Rajab and C. Abou - Rjeily, “Impact of Relay Placement in Three - hop Buffer - aided FSO Systems: an Approximate Performance Analysis Approach,” Physical Communications, 45: 101276, (2021). DOI:10.1109/JSAC.2020.3000826

Patil A., et al. A comprehensive survey on spectrum sharing techniques for 5G and B5G wireless networks. (2024). https://doi.org/10.1016/j.comnet.2024.110697

Khalid Kaabneh, Tasneem Ghnaimat. A Survey of QoS in 5G Network for IoT Applications. International Journal of Science and Applied Information Technology, 8(6), (2019). — P. 159–163. DOI: 10.30534/ijsait/2019/258620198

References

MATLAB Documentation, Fuzzy Logic Toolbox, MathWorks (2024) (in English)

Aliu O.G., Imran A., Imran M.A., Evans B. A Survey of Self Organisation in Future Cellular Networks. IEEE Communications Surveys & Tutorials, (2013). — Vol. 15. — No. 1:336 — 361. DOI: 10.1109/SURV.2012.021312.00116. ResearchGate+2Google Scholar+2 (in English)

Ross T.J. Fuzzy Logic with Engineering Applications. John Wiley & Sons Ltd. (3rd Edition). DOI: 10.1002/9781119994374. SCIRP+1 (in English)

Tkachenko M.O. Razrabotka i issledovanie struktury seti 5G s primeneniem radiofotonnykh tekhnologii [Development and research of 5G network structure using radiophotonic technologies]. Sevastopol: SevGU (2021). — 124 p. (in Russian)

Shilkina S.V., Fokina E.N. Kontroller nechetkoy logiki v upravlenii tekhnologicheskimi protsessami [Fuzzy logic controller in process control]. The Russian Automobile and Highway Industry Journal. (2018); 15(1):106-114. (In Russ.) https://doi.org/10.26518/2071-7296-(2018)-1-106-114 (in Russian)

Kharchenko, Sergey Grigorievich. Riski i problemy razvitiya setey 5G v Rossii [Risks and problems of the development of 5G networks in Russia]: monograph. S.G. Kharchenko, N.K. Zhijin, D.E. Kucher; under ed. S.G. Kharchenko. — Moscow: MAKS Press (2022). — 104 p. ISBN 978-5-317-06740-3 https://doi.org/10.29003/m2574.978-5-317-06740-3 (in Russian)

R. He et al., "High - Speed Railway Communications: From GSM - R to LTE - R," IEEE Vehicular Technology Magazine, 9. (in English)

F.Z. Yousaf, M. Bredel, S. Schaller and F. Schneider, "NFV and SDN—Key Technology Enablers for 5G Networks," IEEE Journal on Selected Areas in Communications. — vol. 35. — no. 11. — P. 2468–2478, November (2017). doi: 10.1109/JSAC.2017.2760418. https://arxiv.org/pdf/1806.07316(in English)

M.L. Shooman, Reliability of Computer Systems and Networks: Fault Tolerance, Analysis, and Design. New York, NY, USA: Wiley (2002) https://sisis.rz.htw - berlin.de/inh2009/12368109.pdf (in English)

3GPP TS 22.289, "LTE; 5G; Mobile communication system for railways (Release 17)", v17.0.0, May (2022). https://www.etsi.org/deliver/etsi_ts/122200_122299/122289/17.00.00_60/ts_122289v170000p.pdf(in English)

Zainab Zaidi, Vasilis Friderikos, Zarrar Yousaf, Simon Fletcher, Mischa Dohler, and Hamid Aghvami. Will sdn be part of 5g? IEEE Communications Surveys & Tutorials, 20(4):3220–3258, (2018). DOI:10.36227/techrxiv. 23981331.v1(in English)

M. El Rajab and C. Abou - Rjeily, "Impact of Relay Placement in Three - hop Buffer - aided FSO Systems: an Approximate Performance Analysis Approach," Physical Communications, 45: 101276, 2021. DOI:10.1016/j. phycom. (2021).101276(in English)

M. El Rajab and C. Abou - Rjeily, "Impact of Relay Placement in Three - hop Buffer - aided FSO Systems: an Approximate Performance Analysis Approach," Physical Communications, 45: 101276, (2021). DOI: 10.1109/JSAC.2020.3000826(in English)

Patil A., et al. A comprehensive survey on spectrum sharing techniques for 5G and B5G wireless networks. (2024). (https://doi.org/10.1016/j.comnet. (2024) .110697(in English)

Khalid Kaabneh, Tasneem Ghnaimat. A Survey of QoS in 5G Network for IoT Applications. International Journal of Science and Applied Information Technology, 8(6) (2019). — P.159–163. DOI: 10.30534/ijsait/2019/258620198(in English)

Publication Ethics and Publication Malpractice in the journals of the Central Asian Academic Research Center LLP

For information on Ethics in publishing and Ethical guidelines for journal publication see http://www.elsevier.com/publishingethics and http://www.elsevier.com/journal-authors/ethics.

Submission of an article to the journals of the Central Asian Academic Research Center LLP implies that the described work has not been published previously (except in the form of an abstract or as part of a published lecture or academic thesis or as an electronic preprint, see http://www.elsevier.com/postingpolicy), that it is not under consideration for publication elsewhere, that its publication is approved by all authors and tacitly or explicitly by the responsible authorities where the work was carried out, and that, if accepted, it will not be published elsewhere in the same form, in English or in any other language, including electronically without the written consent of the copyright-holder. In particular, translations into English of papers already published in another language are not accepted.

No other forms of scientific misconduct are allowed, such as plagiarism, falsification, fraudulent data, incorrect interpretation of other works, incorrect citations, etc. The Central Asian Academic Research Center LLP follows the Code of Conduct of the Committee on Publication Ethics (COPE), and follows the COPE Flowcharts for Resolving Cases of Suspected Misconduct (http://publicationethics.org/files/u2/New_Code.pdf). To verify originality, your article may be checked by the Cross Check originality detection service http://www.elsevier.com/editors/plagdetect.

The authors are obliged to participate in peer review process and be ready to provide corrections, clarifications, retractions and apologies when needed. All authors of a paper should have significantly contributed to the research.

The reviewers should provide objective judgments and should point out relevant published works which are not yet cited. Reviewed articles should be treated confidentially. The reviewers will be chosen in such a way that there is no conflict of interests with respect to the research, the authors and/or the research funders.

The editors have complete responsibility and authority to reject or accept a paper, and they will only accept a paper when reasonably certain. They will preserve anonymity of reviewers and promote publication of corrections, clarifications, retractions and apologies when needed. The acceptance of a paper automatically implies the copyright transfer to the Central Asian Academic Research Center LLP.

The Editorial Board of the Central Asian Academic Research Center LLP will monitor and safeguard publishing ethics.

Правила оформления статьи для публикации в журнале смотреть на сайтах:
www:nauka-nanrk.kz
http://physics-mathematics.kz/index.php/en/archive
ISSN2518-1726 (Online),
ISSN 1991-346X (Print)

Ответственный редактор *А. Ботанқызы*
Редакторы: *Д.С. Аленов, Т. Апендиев*
Верстка на компьютере: *Г.Д. Жадырановой*

Подписано в печать 22.12.2025.
Формат 60x881/8. Бумага офсетная.
Печать –ризограф. 20,0 п.л. Заказ 4.

«Central Asian Academic Research Center» LLP
Алматы, Қонаев к-сі, 142