

ISSN 1998-6688 (Print)
ISSN 2959-8109 (Online)

**ҚАЗАҚСТАН-БРИТАН ТЕХНИКАЛЫҚ
УНИВЕРСИТЕТІНІҢ
ХАБАРШЫСЫ**

**HERALD
OF THE KAZAKH-BRITISH
TECHNICAL UNIVERSITY**

**ВЕСТНИК
КАЗАХСТАНСКО-БРИТАНСКОГО
ТЕХНИЧЕСКОГО УНИВЕРСИТЕТА**

Volume 22, Issue 4
October–December, 2025

**ҚАЗАҚСТАН-БРИТАН
ТЕХНИКАЛЫҚ УНИВЕРСИТЕТІНІҢ
ХАБАРШЫСЫ**

**HERALD
OF THE KAZAKH-BRITISH
TECHNICAL UNIVERSITY**

**ВЕСТНИК
КАЗАХСТАНСКО-БРИТАНСКОГО
ТЕХНИЧЕСКОГО УНИВЕРСИТЕТА**

Алматы
№ 4 (75) 2025 г.

Издание зарегистрировано Министерством культуры и информации Республики Казахстан. Свидетельство о постановке на учет СМИ № 9757 – Ж от 03.12.2008 г.

Журнал зарегистрирован в Международном центре по регистрации сериальных изданий ISSN (ЮНЕСКО, г. Париж, Франция)

В журнале публикуются научные статьи на английском, казахском и русском языках по направлениям: компьютерные науки, математические науки, физические науки, нефтегазовая инженерия и геология, экономика и бизнес.

Подписной индекс – 74206

Издается с 2004 года. Выходит 4 раза в год.

УЧРЕДИТЕЛЬ
Казахстанско-Британский технический университет

ISSN 1998-6688 (Print)
ISSN 2959-8109 (Online)

© Казахстанско-Британский технический университет, 2025

Editor-in-chief – **B. Kulpeshov**, Doctor of Physical and Mathematical Sciences, academician of the National Academy of Sciences of the Republic of Kazakhstan, Professor, School of Applied Mathematics of JSC “Kazakh-British Technical University”, Almaty, Kazakhstan

EDITORIAL BOARD

Section “Computer science”

- Akzhalova A.Zh.** - Candidate of Physical and Mathematical sciences, PhD, Professor, School of Information Technology and Engineering of JSC “Kazakh-British Technical University”, Almaty, Kazakhstan;
- Pak A.A.** - Candidate of Technical Sciences, Associate professor, School of Information Technology and Engineering of JSC “Kazakh-British Technical University”, Almaty, Kazakhstan;
- Samigulina Z.I.** - PhD, Professor, School of Information Technology and Engineering of JSC “Kazakh-British Technical University”, Almaty, Kazakhstan;
- Ilyas Muhammad** - PhD, Professor, School of Information Technology and Engineering of JSC “Kazakh-British Technical University”, Almaty, Kazakhstan.
- Uvalieva I.M.** - PhD, Professor at the Faculty of Computer Science, EKTU named after D. Serikbayev, Ust-Kamenogorsk, Kazakhstan

Section “Mathematical Sciences”

- Kalmurzayev B.S.** - PhD, Associate professor, School of Applied Mathematics of JSC “Kazakh-British Technical University”, Almaty, Kazakhstan;
- Sudoplatov S.V.** - Doctor of Physical and Mathematical Sciences, Chief Researcher, Institute of Mathematics SB RAS; Head of Department, Novosibirsk State Technical University, Novosibirsk, Russia;
- Bazhenov N.A.** - Candidate of Physical and Mathematical Sciences, Senior Researcher, Institute of Mathematics SB RAS, Novosibirsk, Russia.

Section “Physical Sciences”

- Abdirakhmanov A.R.** - PhD, Assistant Professor JSC “KBTU”, Almaty, Kazakhstan; University of Mons, Mons, Belgium;
- Beysen Khanov N.B.** - Doctor of Physical and Mathematical Sciences, Dean, Professor, School of Materials Science and Green Technologies JSC “KBTU”, Almaty, Kazakhstan;
- Laifa Boufendi** - PhD, Professor, Orlean University, France

Section “Oil and Gas Engineering and Geology”

- Ismailov A.A.** - Candidate of Technical Sciences, Dean, Professor, School of Energy and Oil and Gas Industry of JSC “Kazakh-British Technical University”, Almaty, Kazakhstan;
- Korobkin V.V.** - Candidate of Geological and Mineralogical Sciences, Professor, School of Geology of JSC “Kazakh-British Technical University”, Almaty, Kazakhstan;
- Kang Wanli** - Professor, School of Petroleum Engineering, China University of Petroleum (East China), Qingdao, PR China;
- Sarsenbekuly B.** - PhD, Associate professor, School of Energy and Petroleum Engineering of JSC “Kazakh-British Technical University”, Almaty, Kazakhstan;
- Moldabaeva G.Zh.** - Doctor of Technical Sciences, Professor, Department of Petroleum Engineering of K. Turysov Institute of Geology and Petroleum Engineering, KazNTU, Almaty, Kazakhstan.

Section “Economics and Business”

- Narbaev T.S.** - PhD, Professor, Business School JSC “KBTU”, Almaty, Kazakhstan;
- Kozhakhmetova A.K.** - PhD, Assistant Professor, Business School JSC “KBTU”, Almaty, Kazakhstan;
- Aldashev A.** - Professor, International School of Economics JSC “KBTU”, Almaty, Kazakhstan;
- Akhmedov E.T.** - PhD, Associate Professor, Business School, JSC “Kazakh-British Technical University”, Almaty, Kazakhstan;
- Seitkazieva A.M.** - Doctor of Economics, Professor/Business School JSC “KBTU”, Almaty, Kazakhstan;
- Gabriel Castelblanco** - PhD, Assistant-Professor, University of Florida, Florida, USA;
- Dubina I.** - Doctor of Economic Sciences, Professor, Novosibirsk National Research State University (NSU);
- Osei-Kyei Robert** - PhD, Associate Professor, University of Western Sydney, Penrith, Australia;
- Montag Josef** - PhD, Senior Lecturer, Charles University, Prague, Czech Republic..

Бас редактор – **Кулпешов Б.Ш.**, ф.-м.ғ.д., ҚР ҰҒА академигі, профессор,
«ҚБТУ» АҚ Қолданбалы математика мектебі, Алматы қ., Қазақстан

РЕДАКЦИЯ АЛҚАСЫ

«Компьютерлік ғылымдар» бөлімі

- Акжолова А.Ж.** - ф.-м.ғ.к, PhD, профессор, «ҚБТУ» АҚ Ақпараттық технологиялар және инженерия мектебі, Алматы қ., Қазақстан;
Пак А.А. - т.ғ.к., қауымдастырылған профессор, «ҚБТУ» АҚ Ақпараттық технологиялар және инженерия мектебі, Алматы қ., Қазақстан;
Самигулина З.И. - PhD, профессор, «ҚБТУ» АҚ Ақпараттық технологиялар және инженерия мектебі, Алматы қ., Қазақстан;
Pyas Muhammad - PhD, профессор, «ҚБТУ» АҚ Ақпараттық технологиялар және инженерия мектебі, Алматы қ., Қазақстан.
Увалиева И.М. - PhD, компьютерлік ғылымдар факультетінің профессоры, Д. Серікбаев атындағы ШҚТУ, Өскемен қ., Қазақстан

«Математикалық ғылымдар» бөлімі

- Калмурзаев Б.С.** - PhD, қауымдастырылған профессор, «ҚБТУ» АҚ Қолданбалы математика мектебі, Алматы қ., Қазақстан;
Судоплатов С.В. - ф.-м.ғ.д., профессор, РҒА СБ математика институтының бас ғылыми қызметкері; Новосібір мемлекеттік техникалық университеті, кафедра меңгерушісі, Новосібір қ., Ресей;
Баженов Н.А. - ф.-м.ғ.к., РҒА СБ математика институтының аға ғылыми қызметкері, Новосібір қ., Ресей.

«Физикалық ғылымдар» бөлімі

- Әбдірахманов А.Р.** - PhD, ассистент-профессор, «ҚБТУ» АҚ, Алматы қ., Қазақстан; Монс университеті, Монс қ., Бельгия;
Бейсенханов Н.Б. - ф.-м.ғ.д., декан, профессор, Материалтану және жасыл технологиялар мектебі «ҚБТУ» АҚ, Алматы қ., Қазақстан;
Laifa Boufendi - PhD, профессор, Орлеан университеті, Франция.

«Мұнай газ инженериясы және геология» бөлімі

- Исмаилов А.А.** - т.ғ.к., декан, профессор, «ҚБТУ» АҚ Энергетика және мұнай-газ индустриясы мектебі, Алматы қ., Қазақстан;
Коробкин В.В. - геол.-минерал.ғ.к., профессор, «ҚБТУ» АҚ Геология мектебі, Алматы қ., Қазақстан;
Канг Ванли - профессор, Мұнай инженериясы мектебі, Қытай мұнай университеті (Шығыс Қытай), Циндао қ., ҚХР;
Сарсенбекұлы Б. - PhD, қауымдастырылған профессор, «ҚБТУ» АҚ Энергетика және мұнай-газ индустриясы мектебі, Алматы қ., Қазақстан;
Молдабаева Г.Ж. - тех.ғ.д., профессор, мұнай инженериясы кафедрасы, Қ. Тұрысов атындағы геология және мұнай-газ институты, ҚазҰТУ, Алматы қ., Қазақстан.

«Экономика және бизнес» бөлімі

- Нарбаев Т.С.** - PhD, профессор, Бизнес мектебі «ҚБТУ» АҚ, Алматы қ., Қазақстан;
Кожаметова А.К. - PhD, ассистент-профессор, Бизнес мектебі, «ҚБТУ» АҚ, Алматы қ., Қазақстан;
Алдашев А. - PhD, профессор, Халықаралық экономика мектебі, «ҚБТУ» АҚ, Алматы қ., Қазақстан;
Ахмедов Е.Т. - PhD, қауымдастырылған профессор, Бизнес мектебі, «Қазақстан-Британ техникалық университеті» АҚ, Алматы қ., Қазақстан;
Сейтказиева А.М. - э.ғ.к., профессор, Бизнес мектебі «ҚБТУ» АҚ, Алматы қ., Қазақстан;
Gabriel Castelblanco - PhD, ассистент-профессор, Флорида университеті, Флорида, АҚШ
Дубина И. - э.ғ.д., профессор, Новосибирск ұлттық зерттеу мемлекеттік университеті (НМУ);
Osei-Kyei Robert - PhD, қауымдастырылған профессор, Батыс Сидней университеті, Пенрит қ., Австралия;
Montag Josef - PhD, сениор-лектор, Карлов университеті, Прага қ., Чехия Республикасы.

Главный редактор – **Кулпешов Б.Ш.**, д.ф.-м.н., академик НАН РК, профессор,
Школа прикладной математики АО «КБТУ», г. Алматы, Казахстан

ЧЛЕНЫ РЕДКОЛЛЕГИИ

Раздел «Компьютерные науки»

- Акжалова А.Ж.** - к.ф.-м.н., PhD, профессор, Школа информационных технологий и инженерии АО «КБТУ», г. Алматы, Казахстан;
- Пак А.А.** - к.т.н., ассоциированный профессор, Школа информационных технологий и инженерии АО «КБТУ», г. Алматы, Казахстан;
- Самигулина З.И.** - PhD, профессор, Школа информационных технологий и инженерии АО «КБТУ», г. Алматы, Казахстан;
- Puys Muhammad** - PhD, профессор, Школа информационных технологий и инженерии АО «КБТУ», г. Алматы, Казахстан.
- Увалиева И.М.** - PhD, профессор факультета компьютерных наук, ВКТУ им. Д.Серикбаева, г. Усть-Каменогорск, Казахстан

Раздел «Математические науки»

- Калмурзаев Б.С.** - PhD, ассоциированный профессор, Школа прикладной математики АО «КБТУ», г. Алматы, Казахстан;
- Судоплатов С.В.** - д.ф.-м.н., профессор, главный научный сотрудник, Институт математики СО РАН; зав. кафедрой, Новосибирский государственный технический университет, г. Новосибирск, Россия;
- Баженов Н.А.** - к.ф.-м.н., старший научный сотрудник, Институт математики СО РАН, г. Новосибирск, Россия;

Раздел «Физические науки»

- Әбдірахманов А.Р.** - PhD, ассистент-профессор АО «КБТУ», г. Алматы, Казахстан; Университет Монса, г. Монс, Бельгия;
- Бейсенханов Н.Б.** - д.ф.-м.н., декан, профессор, Школа материаловедения и зеленых технологий АО «КБТУ», г. Алматы, Казахстан;
- Laifa Boufendi** - PhD, профессор, Университет Орлеана, Франция.

Раздел «Нефтегазовая инженерия и геология»

- Исмаилов А.А.** - к.т.н., декан, профессор, Школа энергетики и нефтегазовой индустрии АО «КБТУ», г. Алматы, Казахстан;
- Коробкин В.В.** - к.геол.-минерал.н., профессор, Школа геологии АО «КБТУ», г. Алматы, Казахстан;
- Канг Ванли** - профессор, Школа нефтяной инженерии, Китайский нефтяной университет (Восточный Китай), г. Циндао, КНР;
- Сарсенбекұлы Б.** - PhD, ассоциированный профессор, Школа энергетики и нефтегазовой индустрии АО «КБТУ», г. Алматы, Казахстан;
- Молдабаева Г.Ж.** - д.тех.н., профессор, кафедра нефтяной инженерии, Институт геологии и нефтегазового дела им. К. Турысова, КазНТУ, г. Алматы, Казахстан.

Раздел «Экономика и бизнес»

- Нарбаев Т.С.** - PhD, профессор, Бизнес-школа АО «КБТУ», г. Алматы, Казахстан;
- Кожаметова А.К.** - PhD, ассистент-профессор, Бизнес-школа АО «КБТУ», г. Алматы, Казахстан;
- Алдашев А.** - PhD, профессор, Международная школа экономики АО «КБТУ», г. Алматы, Казахстан;
- Ахмедов Е.Т.** - PhD, ассоциированный профессор, Бизнес-школа АО «Казахстанско-Британский технический университет», г. Алматы, Казахстан;
- Сейтказиева А.М.** - д.э.н., профессор, Бизнес-школа АО «КБТУ», г. Алматы, Казахстан;
- Gabriel Castelblanco** - PhD, ассистент профессора, Университет Флориды, Флорида, США.
- Дубина И.** - д.э.н., профессор, Новосибирский национальный исследовательский государственный университет (НГУ);
- Osei-Kyei Robert** - PhD, ассоциированный профессор, Университет Западного Сиднея, г. Пенрит, Австралия;
- Montag Josef** - PhD, сениор-лектор, Карлов университет, Прага, Чешская Республика.

МАЗМҰНЫ – СОДЕРЖАНИЕ

1 КОМПЬЮТЕРЛІК ҒЫЛЫМДАР КОМПЬЮТЕРНЫЕ НАУКИ

Еримбетова А., Бержанова У., Дайырбаева Э., Сәкенов Б., Самбетбаева М. MediaPipe және терең оқыту әдістерін пайдалану арқылы нақты уақыт режимінде қазақ ым тілін тану үшін модель жасау	10
Оралбекова Д., Мамырбаев О., Еримбетова А., Бекарыстанқызы А., Тұрдалыұлы М. Автогрессиялық емес декодтаудың қазақ тіліндегі сөйлеуді тануда қолдану	23
Әбжанова А.Е., Танирбергенев А.Ж., Тасуов Б., Тасжурекова Ж.К., Серикбаева С.К. Применение гибридной модели машинного обучения для классификации типов почв.....	31
Жукабаева Т.К., Марденов Е.М., Танирбергенев А.Ж. Комплексная методика выявления и анализа инцидентов безопасности в беспроводных сенсорных сетях	40
Samigulina Z.I., Shegentay I.T. Development of automation and control system of the technological process for sulfur production	60
Kikbayev N.E., Zhexebay D.M., Xin Y., Tynymbayev S.T., Aitmagambetov A.Z., Abdizhalilova L.B., Skabylov A.A. Comparative study of machine learning methods for detecting anomalies in network traffic.....	79
Abdildaeva A.A., Nurtugan G.B., Wojtkiewicz K. Development and optimization of neural network models with attention mechanisms for intraday price forecasting for EUR/USD	97
Самуйлова А. Средства выявления уязвимостей информационной безопасности на основании данных тематических интернет-ресурсов	107
Есмұхамедов Н.С., Сапақова С.З., Қожамқұлова Ж.Ж., Даниярова Д.Р., Арманқызы Р. Диабеттік ретинопатияны анықтауда көз түбі суреттерін алдын ала өңдеу және талдау әдістері.....	119
Gorlov L.V., Seilova N.A., Okhrimenko T.A. On the diffusion layer generation method.....	131
Нұртас М., Абдикаликова З.Т., Адиат Л. Өзіне-өзі ұқсас шешімдерге арналған физикаға негізделген нейрондық желі (pinn) әдісі.....	143
Амангелді Н., Еримбетова А., Газизова Н.Е., Тұрсынова Н.А., Болатбекқызы К. Глоссарлық қабатты қолдану арқылы қазақ тіліндегі мәтінді формализациялау.....	155
Ashim Zh., Botanov A., Abdoldina F., Serek A. Salary prediction from job descriptions using attention-based nlp models	168
Юсупов Р.В., Құлбекова Ш.А., Аманжолова Р.Ш., Сагин Ж.М. Ауыл шаруашылығына арналған жылжымалы (мобильді) датчиктер	178
Salim T., Kuatbayeva A.A. Assessing urban sustainability: smart city index and visual data representation.....	196
Майлыбаев Е.К., Сейдалиева У.О. Проектирование защищенного сетевого шлюза для облачных приложений	209
Semenyuk V.V., Kurmashev I.G., Serbin V.V., Kurmasheva L.B., Moldagulova A.N. Development of a real-time UAV recognition model based on YOLOv10 neural network	219
Mukhsimbayev B., Pak A., Kuralbayev A. A computational pipeline for lexical and thematic analysis of the code of administrative offenses of the Republic of Kazakhstan.....	227
Satybaldiyeva R.Zh., Kalpeyeva Z.B., Uskenbayeva R.K., Sarkambayeva Sh.G., Razaque A., Elle V.Zh. Benchmarking best practices of top-ranked cities for high livability and citizen participation	244
Kaidullayev M.A., Akzhalova A.Z. A system architecture for disaster management system	254

2 МАТЕМАТИКАЛЫҚ ҒЫЛЫМДАР МАТЕМАТИЧЕСКИЕ НАУКИ

Tamabay D., Temirbekov A., Tleulesova A., Mukhanova T. Alpert wavelet-based galerkin method for first-kind Fredholm integral equations	266
Майер Ф.Ф., Тастанов М.Г., Утемисова А.А., Калаков Б.А. О некоторых подклассах классов почти выпуклых и дважды почти выпуклых функций.....	279

Әбек А.Н., Гогатишвили А., Бокаев Н.А., Унвер Т.	
Весовые неравенства для суперпозиции трех операторов.....	295
Baizhanov B., Sargulova F.	
Uniformly externally definable expansion.....	306
Басаров С.Ж., Нурсултанов Е.Д.	
Восстанавливающая функция для пространств Соболева с доминирующей смешанной производной	313
Төлеубай А.М., Кәріпова А.Т.	
Перфорацияланған облыста Чэфи-Инфанте тендеулерінің аттракторларының орташалануы.....	324

3 ФИЗИКАЛЫҚ ҒЫЛЫМДАР ФИЗИЧЕСКИЕ НАУКИ

Нодяров А.С., Расылова А.Е., Омар А.Ж.	
Спектральный анализ звезды MWC 645.....	340
Ismagambetova T.N., Muratov M.M., Ussenov Y.A., Gabdullin M.T.	
Effect of ionic core on the properties of non-isothermal plasmas	354
Шонгалова А.Қ., Токмолдин Н., Өміртай М.	
Получение и исследование пленки селенида сурьмы путем селенизации	365
Ескермесов Д.К., Лукос С.С., Биркетт М., Пазылбек С.А., Ескермесова А.С., Төлеуханова Ж.Т.	
Физика-механикалық сипаттамалары жақсартылған биомедициналық қолдануға арналған титан-алтын (Ti-Au) негізіндегі жұқа жабындарды алу және зерттеу	374
Коданова С.К., Исанова М.К., Бастыкова Н.Х., Кенжебекова А.И., Хоффманн Д.	
Исследование оптических и динамических свойств водородной плазмы при высоких плотностях	386
Zhukeshov A.M., Abdybai U.B., Salimov E.E., Carlos M.	
Comparative performance characteristics of metal-air batteries with cast and porous electrodes.....	401

4 МҮНАЙ ГАЗ ИНЖЕНЕРИЯСЫ ЖӘНЕ ГЕОЛОГИЯ НЕФТЕГАЗОВАЯ ИНЖЕНЕРИЯ И ГЕОЛОГИЯ

Базарбай Д.Қ.	
Изучение пластов-коллекторов верхнекумкольской подсвиты месторождения Южный Кумколь	411
Liang Peng, Hongbin Yang, Zixiang Yi, Shuhe Zhang, Xin Chen, Haobin Shi, Yubin Zhang, Wanli Kang	
Research progress on threshold pressure gradient in heavy oil reservoirs	419
Болтанов А.К., Тулемисова Ж.С.	
Перспективы выявления новых месторождений полезных ископаемых на лисаковской площади по результатам геологического доизучения	427
Haobin Shi, Hongbin Yang, Xin Chen, Liang Peng, Shuhe Zhang, Wanli Kang, Abdusamig Kadirov	
A review on anti-adsorption behaviors of partially hydrolyzed polyacrylamide in reservoir formations.....	432

5 ЭКОНОМИКА ЖӘНЕ БИЗНЕС ЭКОНОМИКА И БИЗНЕС

Sadykov E., Yesmurzaeva A.	
Modernization of production capacities of JSC “NAC “Kazatomprom”: strategic analysis	444
Shah N., Ali T.	
Adapting sustainability metrics into traditional monitoring and control systems of construction projects: an earned value management approach	456

CONTENTS

1 COMPUTER SCIENCE

Yerimbetova A., Berzhanova U.G., Daiyrbayeva E., Sakenov B., Sambetbayeva M. Development of a model for real-time recognition of Kazakh sign language using MediaPipe.....	10
Oralbekova D., Mamyrbayev O., Yerimbetova A., Bekarystankyzy A., Turdalyuly M. Application of non-autoregressive decoding for Kazakh speech recognition	23
Abzhanova A., Tanirbergenov A., Tassuov B., Taszhurekova Zh., Serikbayeva S. Application of a hybrid machine learning model for soil type classification	31
Zhukabayeva T.K., Mardenov E.M., Tanirbergenov A. Complex technique for detection and analysis of security incidents in wireless sensor networks	40
Samigulina Z.I., Shegentay I.T. Development of automation and control system of the technological process for sulfur production	60
Kikbayev N.E., Zhexebay D.M., Xin Y., Tynymbayev S.T., Aitmagambetov A.Z., Abdizhalilova L.B., Skabylov A.A. Comparative study of machine learning methods for detecting anomalies in network traffic.....	79
Abdildaeva A.A., Nurtugan G.B., Wojtkiewicz K. Development and optimization of neural network models with attention mechanisms for intraday price forecasting for EUR/USD	97
Samuilova A. Tools for identifying information security vulnerabilities based on data from internet resources.....	107
Yesmukhamedov N.S., Sapakova S.Z., Kozhamkulova Zh.Zh., Daniyarova D., Armankyzy R. Methods for pre-processing and analysis of fund images for detection of diabetic retinopathy	119
Gorlov L.V., Seilova N.A., Okhrimenko T.A. On the diffusion layer generation method.....	131
Nurtas M., Abdikalikova Z.T., Adiat L. Physics-informed neural network (pinn) method based on self-similar solutions.....	143
Amangeldy N., Yerimbetova A., Gazizova N., Tursynova N., Bolatbekkyzy K. Development of a model for real-time recognition of Kazakh sign language using mediapipe.....	155
Ashim Zh., Botanov A., Abdoldina F., Serek A. Salary prediction from job descriptions using attention-based nlp models	168
Yussupov R.V., Kulbekova S.A., Amanzholova R.S., Sagin J. Mobile sensors farm support.....	178
Salim T., Kuatbayeva A.A. Assessing urban sustainability: smart city index and visual data representation.....	196
Mailybayev Y., Seidaliyeva U. Design of a secure network gateway for cloud applications.....	209
Semenyuk V.V., Kurmashev I.G., Serbin V.V., Kurmasheva L.B., Moldagulova A.N. Development of a real-time UAV recognition model based on YOLOv10 neural network	219
Mukhsimbayev B., Pak A., Kuralbayev A. A computational pipeline for lexical and thematic analysis of the code of administrative offenses of the Republic of Kazakhstan.....	227
Satybaldiyeva R.Zh., Kalpeyeva Z.B., Uskenbayeva R.K., Sarkambayeva Sh.G., Razaque A., Elle V.Zh. Benchmarking best practices of top-ranked cities for high livability and citizen participation.....	244
Kaidullayev M.A., Akzhalova A.Z. A system architecture for disaster management system	254

2 MATHEMATICAL SCIENCES

Tamabay D., Temirbekov A., Tleulesova A., Mukhanova T. Alpert wavelet-based galerkin method for first-kind Fredholm integral equations	266
Maiyer F.F., Tastanov M.G., Utemissova A.A., Kalakov B.A. About some subclasses of classes of close-to-convex and doubly close-to-convex functions	279
Abek A.N., Gogatishvili A., Bokayev N.A., Ünver T. Weighted inequalities for a superposition of the three operators	295
Baizhanov B., Sargulova F. Uniformly externally definable expansion	306
Bassarov S.Zh., Nursultanov E.D., Recovery function for Sobolev spaces with dominating mixed derivative	313
Toleubay A.M., Karipova A.T. Homogenization of attractors of the Chafee–Infante equations in a perforated domain	324

3 PHYSICAL SCIENCES

Nodyarov A.S., Rassylova A.E., Omar A.Zh. Spectral analysis of the star MWC 645	340
Ismagambetova T.N., Muratov M.M., Ussenov Y.A., Gabdullin M.T. Effect of ionic core on the properties of non-isothermal plasmas	354
Shongalova A.K., Tokmoldin N., Omirtay M. Preparation and investigation of antimony selenide film by selenization	365
Yeskermessov D.K., Lukose C.C., Birkett M., Pazylbek S.A., Yeskermessova A.S., Toleukhanova Zh.T. Synthesis and investigation of titanium-gold (ti-au) based thin coatings for biomedical applications with enhanced physical-mechanical properties	374
Kodanova S., Issanova M., Bastykova N., Kenzhebekova A., Hoffmann D. Study of optical and dynamic properties of hydrogen plasma at high densities	386
Zhukeshov A.M., Abdybai U.B., Salimov E.E., Carlos M. Comparative performance characteristics of metal-air batteries with cast and porous electrodes	401

4 OIL AND GAS ENGINEERING, GEOLOGY

Bazarbay D.K. Study of reservoir formations of the Upper Kumkol Subformation of the South Kumkol field	411
Liang Peng, Hongbin Yang, Zixiang Yi, Shuhe Zhang, Xin Chen, Haobin Shi, Yubin Zhang, Wanli Kang Research progress on threshold pressure gradient in heavy oil reservoirs	419
Boltanov A.K., Tulemisova Zh.S. Potential for new deposits in the Lisakovsk area based on the results of extended geological exploration	427
Haobin Shi, Hongbin Yang, Xin Chen, Liang Peng, Shuhe Zhang, Wanli Kang, Abdusamig Kadirov A review on anti-adsorption behaviors of partially hydrolyzed polyacrylamide in reservoir formations	432

5 ECONOMY AND BUSINESS

Sadykov E., Yesmurzaeva A. Modernization of production capacities of JSC “NAC “Kazatomprom”: strategic analysis	444
Shah N., Ali T. Adapting sustainability metrics into traditional monitoring and control systems of construction projects: an earned value management approach	456

**COMPUTER SCIENCE
КОМПЬЮТЕРЛІК ҒЫЛЫМДАР
КОМПЬЮТЕРНЫЕ НАУКИ**

ӘОЖ 681.327.12
FTAXP 65.011.56

<https://doi.org/10.55452/1998-6688-2025-22-4-10-22>

^{1,2,3}**Еримбетова А.,**

т.ғ.к., қауымдастырылған профессор, PhD, ORCID ID: 0000-0002-2013-1513,
e-mail: aigerian8888@gmail.com

^{2*}**Бержанова У.,**

докторант, ORCID ID: 0009-0000-2467-5721,
*e-mail: berzhanovaulmekenn@gmail.com

^{1,4}**Дайырбаева Ә.,**

оқытушы, ORCID ID: 0000-0002-4255-5456,
e-mail: nurbekkyzy_e@mail.ru

¹**Сәкенов Б.,**

инженер-программист, ORCID ID: 0000-0002-9849-6176,
e-mail: sbakzhan22@gmail.com

⁵**Самбетбаева М.,**

PhD, ORCID ID: 0000-0001-9358-1614,
e-mail: madina_jgtu@mail.ru

¹Ақпараттық және есептеуіш технологиялар институты, Алматы қ., Қазақстан

²Әл-Фараби атындағы Қазақ ұлттық университеті, Алматы қ., Қазақстан

³Еуразиялық технологиялық университеті, Алматы қ., Қазақстан

⁴Қ.И. Сәтбаев атындағы Қазақ ұлттық зерттеу техникалық университеті,
Алматы қ., Қазақстан

⁵Л.Н. Гумилев атындағы Еуразия ұлттық университеті, Астана қ., Қазақстан

**MEDIAPIPE ЖӘНЕ ТЕРЕҢ ОҚЫТУ ӘДІСТЕРІН
ПАЙДАЛАНУ АРҚЫЛЫ НАҚТЫ УАҚЫТ РЕЖИМІНДЕ
ҚАЗАҚ ҒЫМ ТІЛІН ТАНУ ҮШІН МОДЕЛЬ ЖАСАУ**

Аңдатпа

Бұл мақалада MediaPipe платформасын қолдана отырып, қазақ ғым тілін тану жүйесін әзірлеу процесі қарастырылған. Платформа ғым тілі қимылдарын нақты уақытта дәл анықтауға мүмкіндік береді. Негізгі назар ғым қимылдарын тану модельдерін құруға, нейрондық желілерді оқытуға және MediaPipe платформасымен интеграциялауға бөлінген. Маңызды аспектілердің бірі – нейрондық желілердің архитектурасын пайдалану арқылы қимылдарды өңдеудің жоғары дәлдігі мен жылдамдығына қол жеткізу. Жүйе аннотацияланған

қимылдардан алынған үлкен деректер жиынтығы негізінде оқытылған, бұл танудың сапасын айтарлықтай жақсартты. Қазақ ым тілі қимылдарын тану үшін LSTM нейрондық желісі қолданылды, себебі ол уақыттық қатарлар мен деректердің реттіліктерімен тиімді жұмыс істейді. Модель қазақ ым тілінің 30 қимылына оқытылып, ым қимылдарын нақты уақытта мәтінге айналдыруға мүмкіндік береді. Бұл әдіс есту және сөйлеу қабілеті шектеулі адамдармен қарым-қатынасты жеңілдетіп, инклюзивтілікті арттыруға ықпал етеді. Сонымен қатар, пайдаланушымен өзара әрекеттесу үшін нейрондық желіні ым тілі қимылдарын тануға арналған қосымшалармен оңай интеграциялауға мүмкіндік беретін ыңғайлы веб-интерфейс жасалды. Жұмыстың негізгі аспектілерінің бірі – тану дәлдігін жақсарту мақсатында деректерді аннотациялау және өңдеу әдістерін жетілдіру. Жүйені болашақта дамыту ым тілі қимылдарының дерекқорын кеңейту мен веб-қосымшалармен интеграциялауды көздейді. Бұл есту және сөйлеу қабілеті шектеулі адамдар үшін әлеуметтік қамтуды жақсартып, кең ауқымды және қолжетімді платформаны құруға мүмкіндік береді.

Тірек сөздер: қазақ ым тілі, MediaPipe, терең оқыту, жасанды нейрондық желі, нақты уақыт режимі.

Кіріспе

Қазіргі заманғы технологиялар мүмкіндігі шектеулі жандардың өмір сүру сапасын жақсартуға айтарлықтай үлес қосуда. Солардың бірі – ым тілі қимылдарын тану жүйелері, олар есту және сөйлеу қабілеті нашар адамдар үшін аса маңызды рөл атқарады. Мұндай жүйелер ым-ишараларды мәтін немесе сөйлемдерге түрлендіру арқылы қарым-қатынасты жеңілдетіп, инклюзивтілікті арттыруға мүмкіндік береді.

Көптеген зерттеушілер [1–4] CNN (Convolutional Neural Networks) негізінде терең оқыту әдістерін қолдана отырып, өздерінің тәсілдері мен әдістемелерін ұсынған. Бұл жұмыстарда CNN модельдері бейнелерден ерекшеліктерді автоматты түрде бөліп алып, оларды тиімді өңдеу мүмкіндігін көрсетеді. Бірінші және төртінші зерттеулер тек қана CNN модельдеріне сүйенсе, екінші және үшінші зерттеулер бірнеше үлгілердің комбинациясын пайдаланған. Зерттеулерде жақсы нәтижелерге жету көбіне қолданылған деректердің сапасы мен көлеміне байланысты болған.

Ым тілін тану бойынша YOLO (You Only Look Once) объектілерді анықтау моделінің әртүрлі нұсқаларының пайдаланылуы [5–8] зерттеу жұмыстарында ұсынылды. Бұл модельдер ым тілін тануға арналған жүйелерде қолданылып, тез және дәл анықтау үшін дамытылған. Алғашқы жұмыста, YOLO-ның түрлі нұсқаларын қолдану арқылы ым тілін тану үшін арнайы алгоритм ұсынылып, нәтижесінде қолданыстағы модельдерден жоғары көрсеткіштер алынған; содан кейінгі зерттеу жұмысында YOLO-ға негізделген трансформерлік архитектура зерттеліп, ым тілін тану кластарын және шекараларын тікелей болжау мүмкіндігі қарастырылған; үшінші жұмыста Raspberry Pi мен YOLOv3 негізінде автоматты ым тілін тану жүйесі жасалып, кескіндерді өңдеу және тану дәлдігін арттыруға бағытталған; соңғысында, YOLOv5 моделінің графикалық интерфейсі арқылы видеолардың нақты уақытта ым тілін тануын жоғары дәлдікпен жүзеге асыру мүмкіндігі көрсетілген. YOLO моделінің қолданылуы осы төрт жұмыстың нәтижелерін жақсартуда шешуші рөл атқарады. YOLO жүйесінің жылдамдығы мен дәлдігі, сонымен қатар оның онлайн режимде жұмыс істеу қабілеті, ым тілін тану процесін тиімді әрі нақты жасауға мүмкіндік береді. Нәтижесінде, дәлдік пен жылдамдықтың үйлесімі табиғи өзара әрекеттесуді қамтамасыз ететін адам-компьютер қарым-қатынасында жақсы нәтижелерге қол жеткізуге ықпал етеді.

Келесі зерттеушілердің [9–13] еңбектерінде MediaPipe кітапханасы қолданылады, ол қол қимылдарын және ым тілін тану жүйелерінде маңызды рөл атқарады. Бұл кітапхана бейнемазмұнды өңдеу мен қозғалыстарды талдау бойынша тиімді құралдар ұсына отырып, ым тілін тану технологияларының дамуына үлкен үлес қосады. Алдымен бірінші жұмыста MediaPipe кітапханасын пайдаланып, ASL ым тілін тану үшін CNN негізіндегі жүйе ұсынылды, ол 99,95% дәлдік көрсетті. Кейін екінші жұмыста MediaPipe-пен бірге компьютерлік көру әдістерін қолданып, ASL ым тілін дәлірек интерпретациялауға бағытталған жүйе әзірленді, мұнда OpenCV және MediaPipe кітапханаларының мүмкіндіктері пайдаланылды. Үшінші жұмыста MediaPipe негізінде RSL ым тілінің оқшауланған қол қимылдарын тану

жүйесі компоненттік тәсілмен ұсынылды, алайда динамикалық жесттерді тануда қосымша ерекшеліктер мен трансформациялар енгізу қажеттілігі анықталды. Төртінші жұмыста динамикалық қимылдарды тану үшін MediaPipe RNN модельдерімен біріктіріліп, жаңа DSL10 деректер жиынтығы жасалды және модельдің 99%-дан жоғары дәлдікке жеткені көрсетілді. Алайда, төрт жұмыста да динамикалық қимылдар мен статикалық қимылдар арасындағы нақты шекаралар мен деректер жиынтығының шектелуі, сондай-ақ нақты уақыт режиміндегі жүйелердің тиімділігі толыққанды зерттелмеген. MediaPipe кітапханасының қолданылуы ым тілін тану дәлдігін арттыруға ықпал етіп, әртүрлі дене кемістігі бар адамдар мен машиналар арасындағы өзара әрекеттестікті жақсартуға мүмкіндік берді.

Келесі зерттеуде ResNet-50 және VGG-19 негізіндегі ансамбль әдісі, ол қазақ ым тілінің 42 дактилін жіктейді. 57 мыңнан астам бейнелерден тұратын деректер жиыны қалыптастырылып, қол жеткізілген дәлдік 95,7%-ды құраған [14].

Сонымен қатар, қазақ мәтінін KSL глоссина аудару үшін трансформерлерді қолдануда деректердің параллель корпусы әзірленіп, машиналық аударма моделі жоғары дәлдікті көрсетіп, оқытылған [16]. Ал зерттеу [15] деректер сапасын жақсарту үшін аугментация әдістерін қолданудың маңыздылығын атап өткен. Негізгі мәселелерге аннотацияланған деректердің шектеулілігі, кластарды теңестіру қажеттілігі баяндалған.

Ым тілін тану технологиясын жетілдіру бағытындағы зерттеулер [17,18,19] ым қимылдарының динамикасын талдауға негізделген әдістерді ұсынады. Бұл әдістер уақыттық жинақтаушы сипаттамалар мен траекторияларды динамикалық түсіруді қолдану арқылы, сондай-ақ кросс-модальды ақпаратты біріктіретін DFCNet+ моделінің көмегімен жоғары дәлдікке қол жеткізеді. Мұндай тәсілдер ымдардың мағынасын контексте дәлірек анықтауға және түрлі деректер жиынтығында тиімділікті арттыруға мүмкіндік береді.

Бұл жұмыстарда қарастырылған зерттеулер мен әдістер ым тілін тану жүйелерінің тиімділігін арттыруға бағытталғанымен, динамикалық және статикалық қимылдар арасындағы шекараларды анықтау, нақты уақыт режимінде жүйелердің тиімділігін толық зерттеу және LSTM рекуррентті нейрондық желілерін қолдану арқылы қимылдар тізбегін өңдеудің жаңа әдістерін енгізу мәселелері әлі де толық зерттелмеген.

Осы мақалада авторлар бейне және қозғалыстарды дәл талдауға арналған MediaPipe платформасының мүмкіндіктерін қазақ ым тілінің динамикалық қозғалыстарын тану үшін LSTM рекуррентті нейрондық желісімен біріктіретін жаңа әдістемесін ұсынады. Бұл әдіс қимылдар тізбегін тиімді өңдеу және оларды нақты уақыт режимінде тану дәлдігін арттыру үшін қолданылатын шешім ретінде ұсынылады.

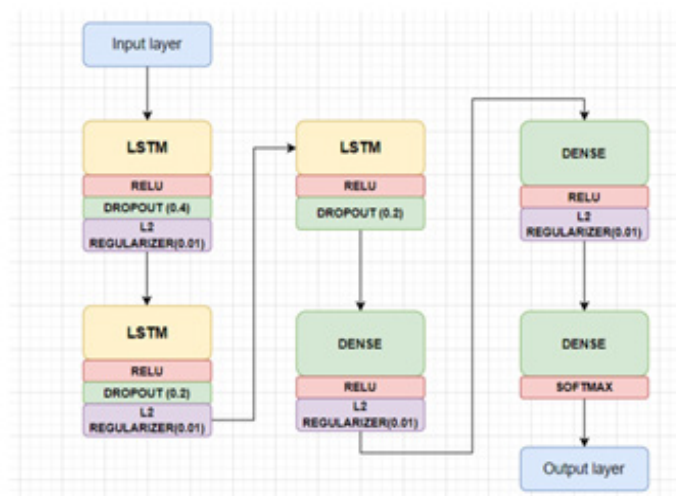
Материалдар мен әдістер

Зерттеу қазақ ым тіліне тән деректерді жинау және өңдеу әдістемесін ұсынады, ол тек статикалық емес, динамикалық қимылдарды да қамтиды. Бұл тәсіл ым тілінің басқа ым жүйелерінен ерекшеленетін бірегей ерекшеліктерін ескере алатын жүйелерді құру үшін маңызды.

Динамикалық қимылдарды тану үшін бірнеше LSTM қабатынан тұратын рекурренттік нейрондық желі құрылды [20]. Модельдің тиімділігіне айтарлықтай әсер ететін нейрондық желілерді құру кезінде қабаттардың саны мен әр қабаттағы нейрондардың саны. Бұл параметрлерді таңдау эксперименттік жолмен жүзеге асырылады, себебі олар оқыту мен тест үлгілеріндегі болжау дәлдігін тікелей анықтайды.

Жүйенің артық оқытылуы (overfitting) жиі кездеседі, егер жасырын қабаттарының саны тым көп болатын күрделі модельдер құрылса. Мұндай жағдайда модель тек оқыту деректерін жаттап алуға бейімделеді де, жаңа деректерді өңдеуде жинақтау қабілеті нашарлайды. Керісінше, қабаттар мен нейрондар саны тым аз болса, модель қажетті ақпаратты үйреніп үлгермей, жеткіліксіз оқытылған күйде қалады. Бұл баланста қол жеткізу нейрондық желінің функционалдық мүмкіндіктерін оңтайландырудың басты шарты.

Жоғарыда ұсынылған теориялық тұжырымдарды ескере отырып, 1-суретте жасанды нейрондық желінің архитектурасы көрсетілген.



Сурет 1 – Жасанды нейрондық желінің архитектурасы

Бірнеше маңызды құрамдас бөліктерден тұратын нейрондық желі архитектурасы 1-суретте сипатталған. Бұл модельде кіріс қабаты, үш LSTM қабаты, үш тығыз қабат және шығыс қабаты бар.

Жасанды нейрондық желінің кіріс деректері жүйеге әрбір кадрда бөлінген 1662 нүктенің орындарын көрсетеді. Бұл жағдайда өлшемі (30, 1662) болып табылады: мұндағы 30 – тізбектегі кадрлар саны, ал 1662 – MediaPipe кітапханасын қолдану нәтижесінде кадрды өңдеу арқылы алынған нүктелердің координаттарының саны.

Relu (Rectified Linear Unit) белсендіру функциясы градиенттің жоғалуын болдырмайды және кері тарату кезінде оңтайландыруды жеңілдететіндіктен, нейрондық желілердің өнімділігін арттыруға көмектеседі. Алғашқы екі LSTM қабатының әрқайсысы 128 нейроннан тұрады және осы қабаттарда Relu белсендіру функциясы қолданылған.

Желінің артық үйрену (overfitting) мәселесін шешуге көмектескен Dropout және $L2$ реттеуі сияқты оңтайландыру әдістері модельде қолданылды. Dropout әдісі кейбір нейрондарды кездейсоқ өшіру арқылы модельдің сенімділігін арттырды, ал $L2$ реттеуі нейрондардың салмағын шектеу арқылы модельдің күрделілігін төмендетіп, жалпы өнімділік пен үйлесімділікке оң әсер етті.

Dropout әдісі нейрондық желілерді оңтайландыру үшін артық үйренудің алдын алу мақсатында қолданылды. Бұл әдіс оқыту барысында әрбір epoch кезінде нейрондардың белгілі бір санын өшіруді қамтамасыз етеді. Нәтижесінде, нейрондық желі белгілі бір түйіндерге шамадан тыс тәуелді болмайды және барлық нейрондар тең дәрежеде оқытылады. Dropout әдісінде нейронның өшірілу ықтималдығы P $[0, 1]$ шегінде анықталады.

Нейрондық желі түріне байланысты ықтималдықты реттеуге арналған жалпы нұсқаулар:

- ♦ Толық байланысқан (Dense) қабаттар үшін Dropout дәрежесі әдетте 0,2 мен 0,5 аралығында орнатылады. Бұл қабаттар шу мен шамадан тыс оқуға (overfitting) бейім болғандықтан, нейрондардың жартысына дейін өшірілуі мүмкін;

- ♦ CNN (конволюциялық нейрондық желілер) үшін P мәндері әдетте 0,1 мен 0,3 арасында орнатылады. Деректердегі маңызды ерекшеліктерді тиімді тануға қабілетті CNN желілері кеңістік ерекшеліктерін сақтайтын құрылымның арқасында тиімділігін арттырады және шамадан тыс оқудан (overfitting) көп зардап шекпейді. Сол себепті, басқа оңтайландыру әдістері қолданылғаннан кейін көбінесе Dropout қабаты енгізіледі;

♦ Модельдің тиімділігін арттырып, артық үйренуді (overfitting) болдырмауға көмектесу үшін, LSTM және RNN модельдері бойынша кіріс пен негізгі қабаттарында Dropout мәндерін дәл орнатқан маңызды. Модельдің тиімді жұмысын қамтамасыз етіп, нейрондардың тек қажетті бөлігі ғана белсенді болып, жасырын қабаттарда Dropout мәні 0,2-ге дейін азайып, кіріс қабаттарында Dropout мәні 0,4 пен 0,8 арасында болуы ықтимал. Себебі RNN жоғары мәнді Dropout мәліметтердің берілуіне кедергі келтіріп, оқыту процесін тұрақсыз етуі мүмкін.

Жоғарыда аталған ұсыныстарға сүйене отырып, соңғы нейрондық желіде бірінші кіріс қабаты үшін Dropout мәні 0,4 деп белгіленді, ал келесі екі қабатта бұл мән екі есе төмендетілді.

Барлық төрт қабатында нейрондық желінің $L2$ сызықтық реттеуі қолданылды, модельге шектеулер қою оқыту процесінде жүзеге асырылып, бұл оңтайландыру техникасы оның жалпыға бірдей болу қабілетін жақсартуға ықпал етеді. Регуляризация – бұл оқу процесіне жаттығу үлгісінде әсер етпейтін, тест үлгісіндегі шығын функциясының мәнін төмендетуге бағытталған модельде енгізілетін кішігірім өзгеріс.

Белгілі бір міндеттерді шешуге бағытталған регуляризацияның екі негізгі түрі бар. Нақтырақ айтқанда, $L1$ реттеуі (Lasso реттеуі) модельде артық салмақтардың пайда болуын алдын алу үшін қолданылады. Бұл әдісте қателік функциясы барлық бар салмақтардың модульдерінің қосындысын белгілі бір коэффициентке көбейтіп қосу арқылы есептеледі. Формула (1) $L1$ реттеуін қолдана отырып, шығын функциясының мәні келесідей болады:

$$J(\theta) = MSE + \lambda \sum |\theta_i|, \quad (1)$$

мұндағы $J(\theta)$ – шығын функциясы және MSE – бұл орташа квадратталған қателік болса, сонда $\lambda \sum |\theta_i|$ – $L1$ реттеуінің коэффициенті. Қате функциясында қосымша терминнің қосылуы нәтижесінде модель тек кішірек салмақтармен жұмыс істеуге мүмкіндік алады. $L1$ реттеуінің ерекшелігі – бұл кейбір салмақтардың нөлге айналуына әкелуі мүмкін, яғни модельде босатылған салмақтар пайда болады. Ал $L2$ реттеу (жолды реттеу) салмақтар тым үлкен болса, үлгі жазаланады. Әр салмақтың квадраты алынады және модельде барлық салмақтардың қосындылары есептеледі (модуль алынғанымен, таңбалары әртүрлі салмақтардың мәндері жойылмайды). Осыдан кейін, (2) формуласында көрсетілгендей, осыдан кейін RMS қатесіне негізделген жоғалту функциясы төмендегідей болады:

$$J(\theta) = MSE + \sum \theta_i^2, \quad (2)$$

Мұнда егер ұсынылған $J(\theta)$ – қателіктің функциясы болса, ендеше MSE – бұл орташа квадратталған қателік, ал машиналық оқытудың моделінің салмақтары (weights) квадраттарын қосуды – $\sum \theta_i^2$ деп белгілейміз. Бұл $L1$ реттеуінде көрінбейтін, бірақ $L2$ реттеуінде нөлге тең болатын разрядталған салмақтар. Нөлдің орнына маңыздылығы аз салмақтар нөлге жақын шағын мәндерді қабылдайды, бірақ әлі де үлгіге қатысады. Тапсырмада әр қимыл 1662 координаттың 30 кадрының 30 тізбегін білдіретін 30 қимылды ажырату керек болғандықтан, жаттығу үшін салмақтар аз болады және оларды модельден толығымен алып тастаудың қажеті жоқ. Сондықтан $L2$ реттеуі таңдалды.

Нәтижелер мен талқылау

Тәжірибелік компонент. Ұсынылған модель сурдоаудармашыны пайдаланбай, ым тілін қолданатын адамдармен тікелей қарым-қатынас орнатуға арналған. Модельді қоғамдық ұйымдардың қызметіне, мысалы, Халыққа қызмет көрсету орталықтары, банктер, университеттер және басқа да мекемелерге интеграциялау жоспарлануда.

Жобаны әзірлеу барысында Python бағдарламалау тілі және оның бірнеше кітапханалары, атап айтқанда MediaPipe, Tensorflow, Keras, Numpy, Matplotlib, Sklearn және басқа да қажетті құралдар қолданылды. Модельді оқыту кезінде нейрондық желіні оңтайландыру әдістерінің

бірі ретінде жүйелеу (regularization) қолдану ұсынылады. Сонымен қатар, нейрондық желілердің бірнеше архитектурасын тексеріп, олардың ішінен ең тиімдісін таңдау керек. Нейрондық желіні оқыту барысында шығындар функциясы ретінде кросс-энтропия және Adam оптимизаторын пайдалану ұсынылады. Модельдің тиімділігін бағалау үшін Accuracy сияқты негізгі дәлдік көрсеткішін қолдануға болады.

Нейрондық желіні оқыту. Модель компьютеріміздің ресурстарын үнемдеу үшін Kaggle Kernels ортасында оқытылды. Модуль 1-кестеде келтірілген ым-ишараға үйретілуі керек болатын:

Кесте 1 – Оқытуға арналған модельдің қимылдары

Қимылдардың нөмірі	Қимылдардың қазақ тілінде (латын) ұсынылуы	Қимылдардың ағылшын тілінде ұсынылуы	Қимылдардың орыс тілінде ұсынылуы
I.	Сәлеметсіз бе (saleketsiz be)	Hello	Здравствуйте
II.	Мен (men)	I	Я
III.	Үйрену (uyrenu)	Learn	Учить
IV.	Ым (ym)	Gesture	Жест
V.	Бүгін (bugin)	Today	Сегодня
VI.	Ауа райы (aua rayi)	Weather	Погода
VII.	Жақсы (jaksy)	Good	Хорошо
VIII.	Мен сені сүйемін (men seni suyemin)	I love you	Я тебя люблю
IX.	Суық (suuq)	Cold	Холодно
X.	Ыстық (ystyq)	Hot	Жарко
XI.	Рақмет (rakhmet)	Thank you	Спасибо
XII.	Телефон (telefon)	Telephone	Телефон
XIII.	Тыныш (tynysh)	Silent	Тихо
XIV.	Билеу (bileu)	Dance	Танцевать
XV.	Кеуде (keude)	Chest	Грудь
XVI.	Ойнау (oynau)	Play	Играть
XVII.	Ақылды (akyldy)	Clever	Умный
XVIII.	Ұшақ (ushaq)	Plain	Самолет
XIX.	Бет (bet)	Face	Лицо
XX.	Адам (adam)	Person	Человек
XXI.	Әтеш (atesh)	Cock	Петух
XXII.	Такси (taxi)	Taxi	Такси
XXIII.	Дөрекі (doreki)	Rude	Грубый
XXIV.	Ит (it)	Dog	Собака
XXV.	Керемет (keremet)	Perfect	Отлично
XXVI.	Жаман (jaman)	Bad	Плохо
XXVII.	Су (su)	Water	Вода
XXVIII.	Сау болыңыз (sau bolynyz)	Good bye	До свидания
XXIX.	Қатты (katty)	Loud	Громко
XXX.	Үлкен (ulken)	Big	Большой

Kaggle платформасында толық деректер жиынтығы ретінде құрылып, ым тілінің қимылдары бір мұрағатта жинақталды. Қимылдарды өңдеу кезінде олардың толық құрылымын қалыптастыру мақсатында, бір қимылдың барлық тізбектерін бір-біріне қосу қажеттілігі туындады, өйткені әр қимыл үшін жеке каталогтар бар және әрбір каталогта кадрлардан тұратын реттіліктер орналасқан.

Келесі кезеңде sklearn кітапханасын пайдаланып, жалпы деректер жиынтығы 90/10 қатынасында оқу және тест үлгілеріне бөлінді. Зерттеу жұмысымызда негізгі мақсат – классификация тапсырмасын орындау болғандықтан, шығын функциясы ретінде көпкатегориялы кросс-энтропия пайдаланылды, ал оңтайландыру үшін модельде Adam алгоритмі қолдану таңдалды. Оқуды ерте тоқтату әдісі оңтайландыру процесінде қолданылады, бұл әдіс модельдің артық үйренуін тежеп, тиімді оқыту процесін қамтамасыз етуге мүмкіндік береді, бұл шамадан тыс оқытуды болдырмау және градиенттің жоғалуын алдын алу үшін көмектесті. Бұл жағдайда, егер қате көрсеткіші 10 қатарынан epoch бойы өзгермесе (patience = 10), онда оқыту ағымдағы epoch-тан кейін тоқтайды, тіпті оқу үшін әлі де epoch қалған болса да.

Модель 150 epoch бойы жаттықтырылды, ал негізгі оқу метрикасы ретінде көпкатегориялы дәлдік (categorical_accuracy) таңдалды (себебі бұл классификация тапсырмасы). Тест деректерінде нақты уақыт режимінде тестілеу жүргізгенде, LSTM және Dropout + L2 регуляризациясымен RNN моделі 92% жалпы дәлдік көрсеткені 2-кестеде ұсынылған.

Кесте 2 – LSTM және Dropout + L2 регуляризациясы бар RNN моделінің нәтижелері

Модель	Метрикасы	Мәні	Эксперименттік шарттар
RNN with LSTM and Dropout + L2 Regularization	Accuracy	92%	30 кадр әр ишараға, 30 класстан

LSTM және Dropout + L2 регуляризациясы бар RNN моделі жоғары дәлдік пен жылдам өңдеу уақытын көрсетеді, бұл оны YOLOv5 және SVM моделдерімен бәсекеге қабілетті етеді (3-кесте).

Кесте 3 – Ым тілі қимылдарын тану үшін модельдерді және олардың сипаттамаларын салыстыру

Модель	Деректер жиыны	Дәлдігі (%)	Кадрды өңдеу уақыты (мс).	Іске асырудың күрделілігі
YOLOv5	[8]	99,6	-	Жоғары
SVM	[13]	99	-	Жоғары
RNN with LSTM and Dropout + L2 Regularization	Біздің деректер жиынтығы	92	30	Орташа

Сынақ үлгісі 100 кадр тізбегінен тұрды және бір қимылда орта есеппен 3 кадр болды, сондықтан үлгіні теңдестірілген деп санауға болады. Оқу және бастапқы бағалау кезеңін аяқтағаннан кейін модельді нақты уақыт режимінде тексерілді. Бұл сценарий, деректерді белгілеуді жүзеге асырып, оларды бөлек функция арқылы өңдеп, модельге жіберу процесін автоматтандырады, алдымен камерада 30 кадрды оқытын нейрондық желі үшін дайын сынақ сценарийі іске қосылды, бұл модельдің кірісі үшін деректерді белгілеу мақсатында болатын. Тиімді оқыту процесі қамтамасыз етіледі, егер де осылайша жеке бір функция арқылы оқылып, модельге қосылып, дайын файл ашып, ол модельдің оқыту нәтижесінде алынған және бөлек сақталған оңтайланған салмақтары бар болуы қамтамасыз етілсе, бұл модельдің дәлдігін және өнімділігін арттыруға мүмкіндік береді.

Содан соң, сол функцияда кіріс ретінде алынған мәліметтер өңделіп, салмақтармен бірге нейрондық желінің кірісіне жіберіледі, нәтижесінде болжам жасалады. Функциядан алынған болжамдар класс белгісі түрінде қайтарылады, олар көрсетілген қимылды нақты бір кадрда анықтап, оқу процесінде қолданылып, қимылдардың дәл сәйкестігі анықталады.

Бірінші кезекте камера объектісі құрылып, оған (0) параметрлі Video Capture тағайындалады. Бұл әдетте жүйенің қателіктерге қарсы тұрақтылығын арттырады, бұл процесс автоматты түрде орындалуды тоқтатып, кадрды оқу барысында скриптта ақау байқалса, жүйенің дұрыс жұмыс істеуін қамтамасыз ету үшін destroyAll Windows () әдісі іске қосылады, кейін кадр арнайы функция арқылы өңделеді, алдымен қате коды мен кадр түрі бойынша оқылып, камерадан бейне ағыны 0 тағайындалады, өйткені көптеген ноутбуктерде камераның идентификаторы енгізу параметрі ретінде 0 болады.

MediaPipe кітапханасында кірістірілген process әдісін пайдаланып, беттегі және денедегі поза нүктелері анықталады. Алдымен draw_face() әдісі, содан кейін draw_hands(), кейін draw_pose() сияқты әртүрлі бөлек әдістер арқылы өңделіп, бұл әдістер нәтижесінде алынған ақпарат бейне ағынына қайтадан жолданады. Нәтижесінде, алынған нүктелер бейнеде бір-бірімен жалпылай сызықтармен байланысып, олар нақты нүктелер ретінде көрініс алады, осылайша бейнедегі адамның беті мен позасы нақты әрі айқын көрсетіледі.

Бұл дегеніміз, кадрда нүктелерді көрсету міндетті емес сияқты көрінуі мүмкін, себебі бізге тек болжанған қимылды сөз түрінде көру жеткілікті, ал кадрда нүктелерді артық көрсету қажет емес. Алайда тәжірибе көрсеткендей, модель екі қолды әрқашан анықтай алмайды, сондықтан бұл жерде нүктелерді көрсету қажеттілігі туындайды.

Модельді түзету және оңтайландыру арқылы ресурстарды жұмсау артық және болашақта кері әсерін тигізуі мүмкін, осындай жағдайда кадрда екі қолдың нүктелері көрсетілмесе, модель толық емес деректер негізінде қате болжамдар жасауы мүмкін немесе мүлдем болжам жасай алмауы ықтимал.

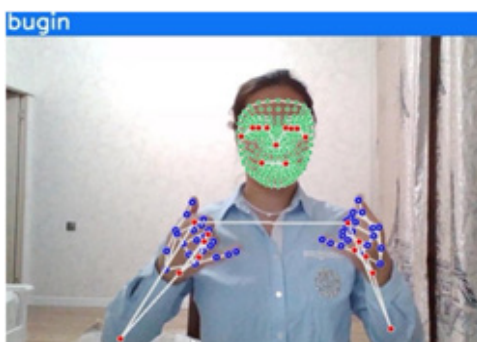
Осы фактіні ескере отырып, барлық нүктелер кадрға орналастырылады. Содан кейін модель алынған кадрларды өңдеуді бастайды. Алдымен скрипт 30 кадр жинауды қамтамасыз етеді.

Кейіннен 30 кадрдан тұратын видео ағысы жинақталып, одан кейін модельдегі predict функциясы шақырылады. Бұл функцияның нәтижесі answ айнымалысына енгізіліп, соңында болжам жасалады. Барлық болжанған сыныптар арасынан ең жоғары ықтимал мәні таңдалады. Егер бұл мән алдын ала белгіленген шектен асып кетсе, ол дұрыс нәтиже ретінде қабылданады.

Бұл модель нақты осы қимыл бейнеленгенін көрсетеді, өйткені ықтималдық мәндерінің ішінен ең жоғарысы argmax() әдісі арқылы таңдалады. Соңында 30 ықтимал мән қайтарылады, бұл жағдайда 30 қимыл = 30 класс құрайды. Массивтің ұзындығы жалпы сыныптардың санына тең болады және ықтималдық мәндерінің толық массиві беріледі. Мұнда модельдің жауабы тек бір нақты санды, сынып белгісін немесе жеке ықтималдық мәнін емес, барлық ықтималдық мәндерін қамтиды, себебі ол көпсыныпты классификация міндетін шешеді. Болжанған ым тілінің қимылы браузерде 2–3-суреттерде көрсетілгендей бейнеленеді:

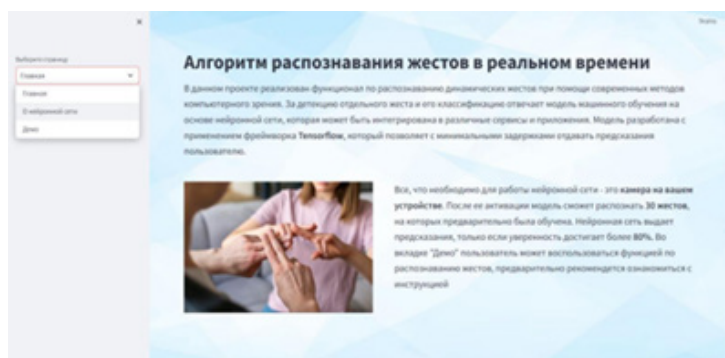


Сурет 2 – «Salemetsiz be!» деген қимылдың браузердегі бейнесі



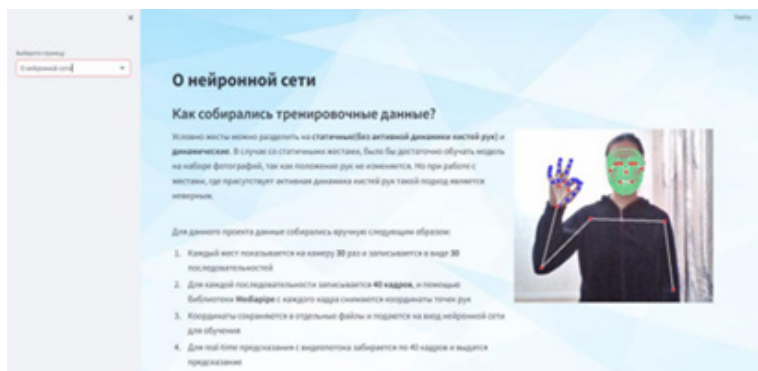
Сурет 3 – «Bugin» деген қимылдың браузердегі бейнесі

Веб-интерфейстің кідіріссіз жұмыс істеуі өте маңызды болды. Сондықтан оның таңдауының басты себебі – пайдаланушыларға модельдермен нақты уақытта жұмыс істеу мүмкіндігін беруі. Бұл құралдың негізгі артықшылығы – Streamlit ашық бастапқы кодты платформа ретінде Python негізінде машиналық оқыту жобаларына арналған дербес веб-қосымшаларды құруға мүмкіндік береді. Сынақтан өткен соң нейрондық желімен пайдаланушының ыңғайлы өзара әрекеттесуін қамтамасыз ету үшін арнайы қарапайым веб-қосымша интерфейсін құру қажеттілігі туындады. Веб-қосымшаның бөліктерін әзірлеу үшін Streamlit фреймворкі таңдалды. Қосымша үш беттен тұрады: алғашқысы – басталу парақшасы, онда жұмыстың басты мақсаты көрсетілген (4-сурет).



Сурет 4 – Браузердегі «Басталу» парақшасы

Қолданушы «Нейрондық желі туралы» тармағына басқанда, ол екінші бетке өтеді, онда нейрондық желінің архитектурасы және оқу деректерін жинаудың артықшылықтарына байланысты қысқаша ақпарат ұсынылады (5-сурет). Бұл мәліметтер қалқымалы навигация панеліндегі бүйірлік мәзірде орналасқан сол жақ бөлігі арқылы қолжетімді.



Сурет 5 – Браузердегі «Ақпарат алу» парақшасы



Сурет 6 – Браузердегі «Көрсету» парақшасы

Қолданушы қимылдарды көрсетуге және сол терезеде болжамдарды алуға мүмкіндік алады, түймені басқан сәтте камераның терезесі ашылып, «Көрсету» бетінде болжамдардың дәлдігін қамтамасыз ету үшін машиналық оқытудың үлгісін пайдалану бойынша нұсқаулықтар берілген (6-сурет). Кейін, «Көрсету» тармағын басу қолданушы ым тілі қимылдарын тану үшін тікелей үлгіні қолдана алатын негізгі бетке апарды.

Қорытынды

Бұл мақалада камерада 30 қазақ ым тілі қимылын нақты уақытта тануға және оларды мәтінге айналдыруға қабілетті LSTM нейрондық желісі әзірленіп, тестіленді. Сондай-ақ, нейрондық желі мен пайдаланушы арасындағы ыңғайлы өзара әрекеттесу үшін шағын веб-интерфейс жасалды. Жобаның негізгі артықшылықтарының бірі – деректерді жинаудағы стандартты емес, сонымен қатар шаблонсыз тәсіл, бұл бір ым тілі қимылын кадрлар тізбегі ретінде қабылдауға мүмкіндік береді, осылайша динамикалық қимылдарды дұрыс өңдеуге жол ашады. Барлық жинақталған деректер қазақ ым тілінің материалдарына жатады, бірақ болашақта орыс тілі сөздері мен алфавитін ескере отырып, оқу үлгісінің көлемін кеңейту жоспарланған. Қазақстандағы ең басты қарым-қатынас тіліне арналған жалғыз ым тілі қимылы базасының болуы нейрондық желіні жасаудың алғашқы қадамы болады, ол болашақта әртүрлі жүйелер мен қосымшаларға интеграцияланады. Нейрондық желінің болжамдау дәлдігі 80%-дан жоғары, бұл қазірдің өзінде жақсы көрсеткіш, бірақ болашақта архитектураны оңтайландыру арқылы дәлдік шегін 95%-ға дейін арттыру жоспарланған. Сондай-ақ, нейрондық желі моделі дискідегі бос орын тұрғысынан ықшам екенін атап өткен жөн, өйткені тек салмақтар жиынтығы сақталады, олар шығаруға оңай пішімде сақталған нақты сандар.

Осылайша, болжамдардың жоғары дәлдігі, кеңістіктің аздығы және архитектураның өнімді Python сценарийі түріндегі көрінісі нейрондық желіні кейіннен дайын шешімдермен біріктіруге ыңғайлы етеді. Есту және сөйлеу қабілеті нашар адамдарға қызмет көрсету жағдайында әлеуметтік қызметтерді көрсетудегі инклюзивтілік деңгейін арттырады. Нейрондық желіні әртүрлі қосымшалар мен қызметтерге жаппай енгізгеннен кейін, есту және сөйлеу қабілеті нашар адамдар сурдоаудармашы қызметтерін қолданбай-ақ мемлекеттік мекемелерге оңай бара алады.

Қаржыландыру туралы ақпарат. Бұл зерттеу Қазақстан Республикасының Ғылым және жоғары білім министрлігінің Ғылым комитеті тарапынан қаржыландырылды (Грант № BR24992875).

ӘДЕБИЕТТЕР

- 1 Sharma, S., Singh, S. Recognition of Indian sign language (ISL) using deep learning model. *Wireless personal communications*, 123 (1), 671–692 (2022). <https://doi.org/10.1007/s11277-021-09152-1>.
- 2 Mukhanov, S. et al. Gesture recognition of the Kazakh alphabet based on machine and deep learning models. *Procedia Computer Science*, 241, 458–463 (2024). <https://doi.org/10.1016/j.procs.2024.08.064>.
- 3 Wu, J. et al. Data glove-based gesture recognition using CNN-BiLSTM model with attention mechanism. *Plos one*, 18 (11), e0294174 (2023). <https://doi.org/10.1371/journal.pone.0294174>.
- 4 Mohamed, T. et al. Intelligent hand gesture recognition system empowered with CNN. 2022 International Conference on Cyber Resilience (ICCR). IEEE, 2022, pp. 1–8. <https://doi.org/10.1109/iccr56254.2022.9995760>.
- 5 Mesbahi, S.C. et al. Hand gesture recognition based on various deep learning YOLO models. *International Journal of Advanced Computer Science and Applications*, 14 (40) (2023). <https://doi.org/10.14569/ijacsa.2023.0140435>.
- 6 Doždor, Z. et al. TY-Net: Transforming YOLO for hand gesture recognition. IEEE access.2023, <https://doi.org/10.1109/access.2023.3341702>.
- 7 Zhang, Z., Wu, B., Jiang, Y. Gesture recognition system based on improved YOLO v3. 2022 7th International Conference on Intelligent Computing and Signal Processing (ICSP). IEEE, 2022, pp. 1540–1543. <https://doi.org/10.1109/icsp54964.2022.9778394>.
- 8 Ling, L., Tao, J., Wu, G. Research on gesture recognition based on YOLOv5. 2021 33rd Chinese Control and Decision Conference (CCDC). IEEE, 2021, pp. 801–806. <https://doi.org/10.1109/ccdc52312.2021.9602731>.
- 9 Kumar, R., Bajpai, A., Sinha, A. MediaPipe and cnns for real-time asl gesture recognition. arXiv preprint arXiv:2305.05296, 2023. <https://doi.org/10.48550/arxiv.2305.05296>.
- 10 Chong, K.S., Subaramaniam, K., Al-Hadi, I. A.A.Q. Developing a Prototype Hand Gesture Recognition System in Interpreting American Sign. *Interpretation*, 3, 1 (2022). <https://doi.org/10.5954/icarob.2024.gs7-4>.
- 11 Grif, M.G., Kondratenko, Y.K. Recognition of Isolated Gestures of the Russian Sign Language Based on the Component Approach //2023 IEEE XVI International Scientific and Technical Conference Actual Problems of Electronic Instrument Engineering (APEIE). IEEE, 2023, pp. 1510–1513. <https://doi.org/10.1109/apie59731.2023.10347694>.
- 12 Samaan, G.H. et al. MediaPipe's landmarks with rnn for dynamic sign language recognition. *Electronics*, 11 (19), 3228 (2022). <https://doi.org/10.3390/electronics11193228>.
- 13 Zholshiyeva, L. et al. A Real-Time Approach to Recognition of Kazakh Sign Language //2022 International Conference on Smart Information Systems and Technologies (SIST). IEEE, 2022, pp. 1–6. <https://doi.org/10.1109/sist54437.2022.9945799>.
- 14 Amirgaliyev, Y., Ataniyazova, A., Buribayev, Z., Zhassuzak, M., Urmashiev, B., & Cherikbayeva, L. Application of neural networks ensemble method for the Kazakh sign language recognition. *Bulletin of Electrical Engineering and Informatics*, 13(5), 3275–3287 (2024). <https://doi.org/10.11591/eei.v13i5.7803>.
- 15 Yerimbetova, A., Sakenov, B., Berzhanova, U., Mukazhanov, N., Daiyrbayeva, E., & Othman, M. Development of A Model of Kazakh Sign Language Recognition Based on Deep Learning Method. In 2024 9th International Conference on Computer Science and Engineering (UBMK) (pp. 822-827). IEEE, 2024, October. <https://doi.org/10.11591/eei.v13i5.780310.1109/UBMK63289.2024.10773578>.
- 16 Yerimbetova, A., Sakenov, B., Sambetbayeva, M., Daiyrbayeva, E., Berzhanova, U., & Othman, M. Creating a Parallel Corpus for the Kazakh Sign Language and Learning. *Applied Sciences*, 15(5), 2808 (2025). <https://doi.org/10.3390/app15052808>.
- 17 Vijitkunsawat, W., Anunvapong, P., & Chantngarm, P. Human Joint Coordinate Sequencing in Video-Based Thai Finger Spelling Recognition. In 2024 8th International Conference on Information Technology (InCIT) (pp. 681–686). IEEE, 2024, November. <https://doi.org/10.1109/InCIT63192.2024.10810526>.
- 18 Katti, R.K., Sujatha, C., Desai, P., & Shankar, G. Character and word level gesture recognition of Indian Sign language. In 2023 IEEE 8th International Conference for Convergence in Technology (I2CT) (pp. 1–6). IEEE, 2023, April. <https://doi.org/10.1109/I2CT57861.2023.10126314>.
- 19 Feng, Y., Chen, N., Wu, Y., Jiang, C., Liu, S., & Chen, S. DFCNet+: Cross-modal dynamic feature contrast net for continuous sign language recognition. *Image and Vision Computing*, 151, 105260 (2024). <https://doi.org/10.1016/j.imavis.2024.105260>.
- 20 Aleshin, N.A. Recurrent neural networks. *World science: problems and innovations*, pp. 10–12, 2021.

^{1,2,3}**Еримбетова А.,**

к.т.н., ассоциированный профессор, PhD, ORCID ID: 0000-0002-2013-1513,
e-mail: aigerian8888@gmail.com

^{2*}**Бержанова У.Г.,**

докторант, ORCID ID: 0009-0000-2467-5721,
*e-mail: berzhanovaulmekenn@gmail.com

^{1,4}**Дайырбаева Э.,**

преподаватель, ORCID ID: 0000-0002-4255-5456,
e-mail: nurbekkyzy_e@mail.ru

¹**Сакенов Б.,**

инженер-программист, ORCID ID: 0000-0002-9849-6176,
e-mail: sbakzhan22@gmail.com

⁵**Самбетбаева М.,**

PhD, ORCID ID: 0000-0001-9358-1614
e-mail: madina_jgtu@mail.ru

¹Институт информационных и вычислительных технологий КН МНВО РК,
г. Алматы, Казахстан

²Казахский национальный университет им. аль-Фараби, г. Алматы, Казахстан

³Евразийский технологический университет, г. Алматы, Казахстан

⁴Казахский национальный исследовательский технический университет им. К.И. Сатпаева,
г. Алматы, Казахстан

⁵Евразийский университет им. Л.Н. Гумилева, г. Астана, Казахстан

СОЗДАНИЕ МОДЕЛИ ДЛЯ РАСПОЗНАВАНИЯ КАЗАХСКОГО ЖЕСТОВОГО ЯЗЫКА В РЕЖИМЕ РЕАЛЬНОГО ВРЕМЕНИ С ПОМОЩЬЮ MEDIAPIPE И МЕТОДОВ ГЛУБОКОГО ОБУЧЕНИЯ

Аннотация

В данной статье рассматривается процесс разработки системы распознавания казахского жестового языка с использованием платформы MediaPipe. Платформа позволяет эффективно распознавать жесты в реальном времени. Основное внимание уделено созданию моделей для распознавания жестов, обучению нейронных сетей и интеграции с платформой MediaPipe. Один из важных аспектов – достижение высокой точности и скорости обработки жестов с использованием архитектуры нейронных сетей. Система была обучена на большом наборе аннотированных жестов, что значительно улучшило качество распознавания. Для распознавания жестов казахского языка использовалась нейронная сеть типа LSTM, так как она эффективно работает с временными рядами и последовательностями данных. Модель была обучена на 30 жестах казахского жестового языка, что позволяет преобразовывать жесты в текст в реальном времени. Этот подход значительно облегчает коммуникацию с людьми, имеющими проблемы со слухом и речью, и способствует повышению инклюзивности. Кроме того, был разработан удобный веб-интерфейс, который позволяет легко интегрировать нейронную сеть с приложениями для распознавания жестов. Одним из ключевых аспектов работы является улучшение методов аннотирования и обработки данных для повышения точности распознавания. Будущее развитие системы включает расширение базы данных жестов и интеграцию с веб-приложениями. Это позволит улучшить социальную интеграцию людей с нарушениями слуха и речи и создать широкую и доступную платформу.

Ключевые слова: казахский жестовый язык, MediaPipe, глубокое обучение, искусственная нейронная сеть, режим реального времени.

^{1,2,3}**Yerimbetova A.,**

Cand. Tech. Sc., Associate Professor, PhD, ORCID ID: 0000-0002-2013-1513,
e-mail: aigerian8888@gmail.com

^{2*}**Berzhanova U.G.,**

PhD student, ORCID ID: 0009-0000-2467-5721,
e-mail: berzhanovaulmekenn@gmail.com

^{1,4}**Daiyrbayeva E.,**

Senior Lecturer, ORCID ID: 0000-0002-4255-5456,
e-mail: nurbekkyzy_e@mail.ru

¹**Sakenov B.,**

Software engineer, ORCID ID: 0000-0002-9849-6176,
e-mail: sbakzhan22@gmail.com

⁵**Sambetbayeva M.,**

PhD, ORCID ID: 0000-0001-9358-1614,
e-mail: madina_jgtu@mail.ru

¹Institute of Information and Computational Technologies of the Committee
of Science of the Ministry of Science and Higher Education
of the Republic of Kazakhstan, Almaty, Kazakhstan

²Al-Farabi Kazakh National University, Almaty, Kazakhstan

³Eurasian Technological University, Almaty, Kazakhstan

⁴Satbayev University, Almaty, Kazakhstan

⁵L.N. Gumilyov Eurasian National University, Astana, Kazakhstan

DEVELOPMENT OF A MODEL FOR REAL-TIME RECOGNITION OF KAZAKH SIGN LANGUAGE USING MEDIAPIPE AND DEEP LEARNING METHODS

Abstract

This article discusses the process of developing a Kazakh sign language recognition system using the MediaPipe platform. The platform allows for efficient real-time gesture recognition. The main focus is on creating models for gesture recognition, training neural networks, and integrating with the MediaPipe platform. One of the key aspects is achieving high accuracy and speed in gesture processing by using neural network architecture. The system was trained on a large dataset of annotated gestures, which significantly improved the recognition quality. For recognizing Kazakh sign language gestures, an LSTM neural network was used because it effectively works with time series and data sequences. The model was trained on 30 Kazakh sign language gestures, enabling the conversion of gestures into text in real-time. This approach greatly facilitates communication with people who have hearing and speech impairments and contributes to increased inclusivity. Additionally, a user-friendly web interface was developed, allowing easy integration of the neural network with applications for gesture recognition. One of the key aspects of the work is improving data annotation and processing methods to enhance recognition accuracy. The future development of the system includes expanding the sign language gesture database and integration with web applications. This will improve social inclusion for people with hearing and speech impairments and create a broad, accessible platform..

Key words: Kazakh sign language, MediaPipe, deep learning, artificial neural network, in real-time.

Мақаланың редакцияға түскен күні: 13.05.2025

ӘОЖ 004.912; 004.93
ҒТАХР 28.23.37

<https://doi.org/10.55452/1998-6688-2025-22-4-23-30>

^{1*}**Оралбекова Д.,**

PhD, аға ғылыми қызметкер, ORCID ID: 0000-0003-4975-6493,

*e-mail: dinaoral@mail.ru

¹**Мамырбаев О.,**

PhD, профессор, бас ғылыми қызметкер, ORCID ID: 0000-0001-8318-3794,

e-mail: morkenj@mail.ru

¹**Еримбетова А.,**

PhD, т.ғ.к., қауымдастырылған профессор, ORCID ID: 0000-0002-2013-1513,

e-mail: aigerian8888@gmail.com

¹**Бекарыстанқызы А.,**

PhD, аға ғылыми қызметкер, ORCID ID: 0000-0003-3984-2718,

e-mail: akbayan.b@gmail.com

¹**Тұрдалыұлы М.,**

PhD, аға ғылыми қызметкер, ORCID ID: 0000-0002-1470-3706,

e-mail: m.turdalyuly@gmail.com

¹Ақпараттық және есептеуіш технологиялар институты, Алматы қ., Қазақстан

АВТОГРЕССИЯЛЫҚ ЕМЕС ДЕКОДТАУДЫҢ ҚАЗАҚ ТІЛІНДЕГІ СӨЙЛЕУДІ ТАНУДА ҚОЛДАНУ

Андатпа

Сөйлеуді тану саласында интегралдық модельдер дәстүрлі және гибриді тәсілдерді біртіндеп ығыстырып келеді. Олардың негізгі принципі – автогрессиялық декодтау, мұнда шығатын тізбек солдан оңға қарай қалыптасады. Алайда осы уақытқа дейін бұл әдіс дауысты мәтінге ауыстыруда ең жақсы нәтиже беретіні дәлелденген жоқ. Сонымен қатар, интегралдық модельдер тек алдыңғы контекстке сүйенеді, бұл анық емес немесе бұрмаланған дауыстарды өңдеуді қиындатады. Осыған байланысты Insertion әдісі ұсынылды, ол автогрессиялық декодтауды пайдаланбайды және шығатын деректерді кез келген ретпен генерациялайды. Осы жұмыста Insertion әдісі мен коннекциялық уақытша классификация (СТС) негізінде оқытылған қазақ тіліндегі сөйлеуді тану моделі қарастырылады. Жүргізілген эксперименттер бұл әдістің тану дәлдігін арттыруға мүмкіндік беретінін көрсетті. Автогрессиялық модельдерден айырмашылығы, Insertion әдісі тізбектерді өңдеуде үлкен икемділік береді, себебі ол шығатын деректерді қалыптастыруда қатаң тәртіпті қажет етпейді. Бұл декодтау кезіндегі кідірістерді азайтады және модельді анық айтылмаған сөздерге төзімді етеді. Сонымен қатар, Insertion әдісін СТС-мен үйлестіру аудио деректер мен мәтін транскрипциясының сәйкестігін жақсартуға мүмкіндік береді. Бұл агглютинативті тілдер, мысалы, қазақ тілі үшін өте маңызды. Эксперименттер нәтижесінде ұсынылған модельдің тану дәлдігі 10,2%-ға дейін жетіп, қазіргі таңда бәсекеге қабілетті көрсеткіштерге ие болды.

Тірек сөздер: коннекциялық уақытша классификация, интеграциялық модельдер, қазақ тіліндегі дыбысты тану, Insertion, Transformer.

Кіріспе

Сөйлеуді синтездеу және тану технологиялары роботтандырылған және автоматтандырылған жүйелердің ажырамас бөлігіне айналды. Олар call-орталықтарда операторлардың жүктемесін азайтуға, қоңырауларды мамандар арасында жылдам бөлуге, сондай-ақ банк саласында интеллектуалды дауыс көмекшілері мен басқа да сервистер түрінде кеңінен қолданылуда. Интегралдық (end-to-end, E2E) модельдер [1], оның ішінде коннекциялық уақытша классификация (СТС) [2], кодер-декодерлік архитектуралар мен назар аудару механизмдері [3],

сондай-ақ ағымдық дауысты өңдеуге арналған онлайн-модельдер [4, 5], дыбысты автоматты тану (ASR) жүйелерін құруды айтарлықтай жеңілдетті. E2E-модельдер дәстүрлі көпдеңгейлі жүйелерді бірегей терең нейрожелімен ауыстырып, оқыту және декодтау уақытын қысқартуға мүмкіндік береді. Сонымен қатар, олар тану процесін және оның кейінгі өңдеуін, мысалы табиғи тілдерді түсіну тапсырмалары үшін, оңтайландыруға жол ашады.

Соңғы жылдары E2E-модельдердің тиімділігін арттыру бойынша белсенді жұмыстар жүргізілуде, әсіресе CTC [6, 7] және назар аудару механизмдерін [8, 9] жетілдіруде үлкен жетістіктерге қол жеткізілді. Зерттеулер көрсеткендей, әртүрлі архитектураларды біріктіру арқылы интегралдық модельдерді қолдану дәстүрлі ASR жүйелерінен ғана емес, сондай-ақ жеке жүзеге асырылған E2E тәсілдерінен де жоғары нәтижелерге қол жеткізуге мүмкіндік береді. Осы саладағы айтарлықтай прогресс Transformer моделінің пайда болуы және оның CTC-мен біріктірілуі [10, 11] арқасында мүмкін болды, бұл дыбысты тану сапасын жаңа деңгейге көтерді.

Көптеген E2E-модельдер «тізбектен тізбекке» (seq2seq) тәсілімен жұмыс істейді, онда автогрессиялық декодтау қолданылады, мұнда шығатын тізбек солдан оңға қарай қалыптасады. Алайда бұл әдіс шектеулерге ие, ол тек алдыңғы контексті ескеріп, болашақ сөздер туралы ақпаратты елемейді. Бұл анық емес айтылған дыбысты өңдеуде қателіктерге әкелуі мүмкін. Мұндай жағдайда тек алдыңғы емес, сондай-ақ кейінгі сөздерді ескеру маңызды. Бұндай көзқарас модельдің фонетикалық бұрмалауларға төзімділігін арттырады.

Машиналық аударма тапсырмаларында белсенді түрде автогрессиялық емес трансформерлер [12] қолданылады, және осы тәсілді сөйлеуді тануға бейімдеу бойынша әрекеттер жасалды [13]. Бұл әдіс негізінде декодтау процесінде жасыратын токендерді болжау жатыр. Модель жасырынып қалған элементтерді назар аудару механизмі арқылы қайта құруға оқытылады, бұл оған автогрессиялық емес тәсілмен тізбектерді генерациялауға мүмкіндік береді. Бұл әдіс автогрессиялық модельдермен салыстырғанда жоғары тиімділік көрсетті, бірақ оны қолдану кезінде қиындық туындайды. Бұл – шығатын тізбектің ұзындығын алдын ала анықтау қажеттілігі.

Осы мәселені шешу үшін Insertion әдісі ұсынылды, ол шығатын деректерді кез-келген ретпен генерациялауға мүмкіндік береді, бұл тізбектің ұзындығын болжау үшін қосымша элементтерді қажет етпейді [14]. Бұл тәсіл декодтау қадамдарының санын азайтып, өңдеу уақытын қысқартуға мүмкіндік береді.

Біздің зерттеуіміз қазақ тілін ақпараттық жүйелерде автоматты тану технологияларын дамытуға және енгізуге бағытталған. Алдыңғы жұмыстарда [15,16] қазақ тілін машиналық аударма мен морфологиялық талдау мәселелері қарастырылған, сондай-ақ дауысты идентификациялау және верификациялау жүйелері әзірленген [17]. Бұған дейін біз агглютинативті тілдер, мысалы қазақ және әзірбайжан тілдерінде тану үшін назар аудару механизмі бар CTC-модельдерін [18], сондай-ақ қазақ тілінде ағымдық тану үшін RNN-T моделін [19] іске асырдық. Алайда, қолданыстағы шешімдер қазақ тілін мәтінге түрлендірудің дәлдігін арттыру үшін жетілдіруді қажет етеді.

Бұл жұмыстың негізгі мақсаты – Insertion әдісіне негізделген қазақ тілін интегралдық тану моделін әзірлеу және тестілеу, сондай-ақ оның тиімділігін қолданыстағы тәсілдермен салыстыру. Зерттеу барысында келесі міндеттер қойылды:

- ♦ Қазақ тілін автоматты тану үшін Insertion моделін әзірлеп, оқыту.
- ♦ Декодтау дәлдігін арттыру үшін Insertion әдісін CTC-мен біріктіру.
- ♦ Қазақ тілін тану бойынша эксперименттер жүргізіп, нәтижелерді дәстүрлі автогрессиялық модельдермен салыстыру.

♦ Декодтау қадамдарының санын азайту және модельдің жұмысын жеделдету үшін Insertion әдісінің әсерін бағалау.

Ұсынылған тәсіл қазақ тілін тану дәлдігін арттыруға және сөздер мен таңбаларды танудағы қателіктерді азайтуға бағытталған. Бұл агглютинативті морфологиясы бар тілдерді өңдеуге перспективті болатынын көрсетеді.

Материалдар мен әдістер

Ең ықтимал сөздер тізбегі W барлық мүмкін болатын V тізбектері арасында $P(W|A)$ ықтималдылығын максимизациялау арқылы анықталады (1). Бұл процесті келесі өрнекпен көрсетуге болады [20]:

$$W^* = \underset{W \in V^*}{\operatorname{argmax}} P(W|A) \quad (1)$$

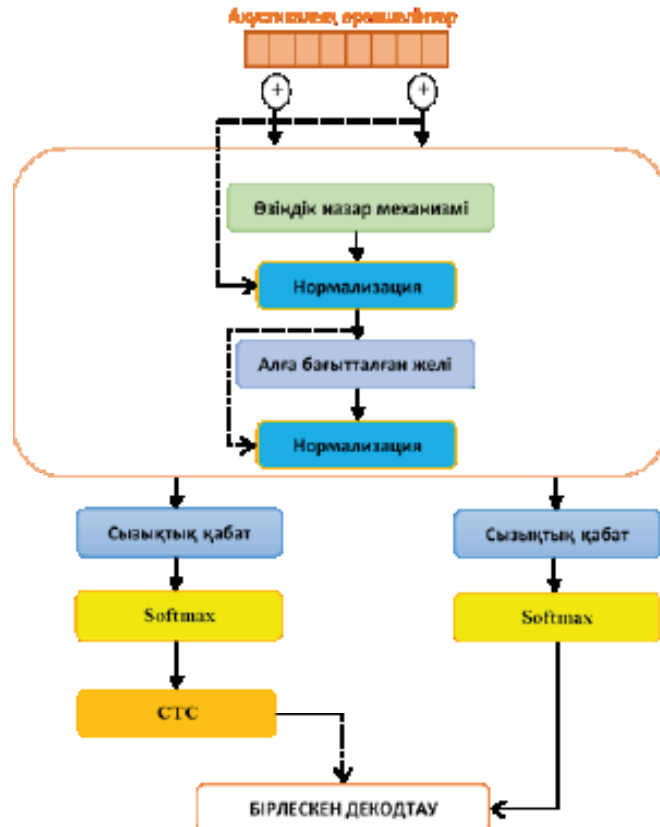
Бұл өрнекте акустикалық сипаттамалар A түріндегі ерекшелік векторларының тізбегі ретінде көрсетіледі, ал W – шығыс сөздер тізбегін білдіреді. Осылайша, интегралдық автоматты сөйлеуді тану жүйесінің негізгі міндеті – барлық ықтимал V^* тізбектері үшін $P(W | A)$ апостериорлы үлестірімін дәл бағалай алатын модель құру.

Insertion моделін іске асыру үшін тек сөздер тізбегін емес, олардың қосылу тәртібін де ескеру қажет. Order параметрі сөздердің декодтау процесінде нәтижелік тізбекке қандай тәртіппен енгізілетінін анықтайды. Мысалы, сөздерді бірінен соң бірін тізбектей қосудың орнына, модель алдымен негізгі сөздерді немесе ең ықтимал токендерді болжай алады, сосын оларды қалған элементтермен толықтырады. Бұл декодтауды икемді етеді, себебі модель әртүрлі жағдайларға бейімделіп, болашақ сөздер контексін пайдалана алады.

Order параметрін ескере отырып, (1) теңдеуі келесі түрде түрленеді:

$$P(W | A) = \sum_{\text{Order}} p(W, \text{Order} | A) = \sum_{\text{Order}} p(W_{\text{Order}} | A) p(\text{Order} | A) \quad (2)$$

Бұл тәсіл декодтау кезінде шығыс тізбегін құрудың бірнеше нұсқасын ескеруге мүмкіндік береді. Insertion әдісіне негізделген декодтауды іске асыру үшін Transformer нейрожелісі пайдаланылады, мұнда позициялық тәртіп механизмі қолданылады (1-сурет).



Сурет 1 – Insertion-CTC гибриді моделінің құрылымы

Әзірленген модель интегралдық әдіспен оқытылады, бұл E2E архитектураларының барлық артықшылықтарын сақтауға мүмкіндік береді. Бұл әдістің негізгі артықшылықтарының бірі – шығатын тізбектің ұзындығын алдын ала анықтаудың қажеттілігі жоқ, сонымен қатар бұл декодтау процесін икемді етеді.

(2) теңдеуде көрсетілгендей, сөздердің позицияларын болжау алдын ала жасалған сөздерге назар аудару механизмі арқылы енгізу әдісімен жүзеге асырылады. Позицияларды жаңарту бұрын есептелген ұсыныстарға әсер етпегендіктен, енгізу тәртібі туралы ақпарат кейінгі декодтау кезеңдерінде қайта қолданылуы мүмкін.

Шығыс тізбегін құру процесінде әр сөздің шартты ықтималдығы сөздердің және олардың позицияларының біріктірілген үлестірімін ескере отырып есептеледі. Бұл енгізу тәртібін оңтайландырады және декодтаудың дәлдігін арттырады.

Нәтижелер мен талқылау

Insertion әдісіне негізделген модельді оқыту үшін Ақпараттық және есептеуіш технологиялар институты жинаған көлемі 405 сағаттан астам тілдік корпус пайдаланылды. Бұл корпус әртүрлі жастағы және жыныстағы адамдардың қазақ тіліндегі жазбаларынан тұрады. Деректер құрамында телефон арқылы сөйлесу жазбалары мен транскрипциялары, сондай-ақ жаңалықтар көздерінен алынған аудиоматериалдар мен көркем аудиокітаптар бар.

Барлық дыбыс файлдары оқыту (85%) және тестілеу (15%) жиынтығына бөлінді, бұл модельді тиімді оқытуға және тексеруге мүмкіндік береді. Алынған корпус түрлі деректер жиынтығының тану сапасына әсерін зерттеуге, сондай-ақ модель параметрлерін талдап, олардың қазақ тілін декодтау дәлдігіне әсерін бағалауға мүмкіндік береді.

Аудиофайлдар .wav форматында ұсынылған, ал оларды өңдеуде жоғары дәлдікпен сандық деректерді алуға мүмкіндік беретін сызықтық импульс-кодтық модуляция (PCM) әдісі қолданылған. Жазбалар 44,1 кГц дискретизациясы және 16 биттік разрядтылыққа ие, бұл дауыс деректерін өңдеу үшін сапа стандарттарына сәйкес келеді.

Негізгі параметрлерді баптау

Модельді оқыту барысында кодер мен декодер үшін 10 блок таңдалды, бұндай таңдау тізбектерді өңдеуді жақсартып, есептеу күрделілігі мен болжау дәлдігі арасындағы балансты сақтауға мүмкіндік берді.

Декодтау сапасын жақсарту және жақындауды жылдамдату үшін Insertion моделіне коннекциялық уақытша классификация (CTC) интеграцияланды. Модельде әр қабатта 512 ұяшығы бар алты қабатты екі бағытты LSTM (BLSTM) қолданылды. Осының арқасында контексті өңдеу жақсартылып, ұзақ сөйлемдерді тану дәлдігі артты. Артық оқытуды болдырмау үшін dropout = 0,1 коэффициентпен шектелді.

Модель параметрлерін оңтайландыру үшін градиентті төмендету тұрақтылығын қамтамасыз ететін және салмақтардың жылдам бейімделуін қамтамасыз ететін Adam алгоритмі пайдаланылды. Декодтау кезінде CTC салмағы 0,4 деңгейінде орнатылды. Осындай шектеу модельге аудио сигналының мәтіндік транскрипциямен жақсы үйлестіруге мүмкіндік берді.

Қосымша beam search іздеу ені 15 пайдаланылды, себебі осындай шектеу күрделі тіркестер мен нақты қазақ сөз түрлерін тану сапасын жақсартуға ықпал етті. Іздеу енін ұлғайту модельге соңғы оптималды болжауды таңдамас бұрын көбірек гипотезаларды ескеруге мүмкіндік берді, бірақ бұл есептеу шығындарын аздап арттырды.

Оқыту 100 кезең бойы жүргізілді, ол модельдің тереңірек жақындауына және шығыс деректерінің жақсы генерациясына мүмкіндік берді. Сонымен қатар, алдын ала оқытылған эмбеддингтер қолданылды. Олар оқытудың бастапқы кезеңін жылдамдатып, жалпы дәлдікті арттыруға ықпал етті.

Декодтау кезеңінде сөйлеу тізбектерін үйлестіру, сирек сөздерді тануды жақсарту және модельдің соңғы шығысынан тиімді нәтиже алу үшін тілдік модель қолданылды.

Эксперимент жүргізу және нәтижелерді алу

Қазіргі уақытта қазақ тіліндегі сөйлеуді тану бойынша интегралдық модельдерге арналған зерттеулер әлі шектеулі. Сонымен қатар, қазақ сөйлеуін автоматты тануға арналған дәстүрлі және гибриді тәсілдерді қарастырған бірнеше ғылыми жұмыстар бар.

Біздің алдыңғы жұмысымызда [18] трансферлік оқытуға негізделген тәсіл ұсынылған, ол қазақ және әзірбайжан тілдерінде интегралдық сөйлеуді тану үшін қолданылды. Зерттеу аясында агглютинативті тілдерді тиімді өңдеуге мүмкіндік беретін CTC-назар аудару модельдері әзірленді. Оқыту екі тілдік корпус бойынша бір уақытта жүргізілді, оның ішінде 400 сағаттық қазақ тіліндегі аудиолар болды.

Басқа зерттеуде [19] қазақ тіліндегі дыбысты тану үшін бейімделген RNN-T моделінің нұсқасы іске асырылды. Бұл жұмыста сөйлеу болжамдарын жақсартуға оң әсерін тигізген тілдік модель (TM) қолданылды. Модель 300 сағаттық корпус бойынша оқытылды, бұл соңғы транскрипциялардың дәлдігін айтарлықтай арттырды.

Аталған зерттеулерде алынған дыбысты тану нәтижелерін салыстырмалы талдау және осы жұмыста ұсынылған Insertion моделінің тиімділігі 1-кестеде ұсынылған.

Кесте 1 – Дыбысты тану жүйелері үшін CER көрсеткіштері

Модель	CER (%)	Деректер көлемі, сағ
End-to-end трансфермен, TM-сіз	14.23	400
LSTM және BLSTM негізіндегі RNN-T, TM-мен	10.6	300
Ұсынылған модель Insertion-CTC	10.2	405

Insertion моделін CTC және тілдік модельмен үйлестіріп қолдану тану дәлдігін 4%-ға арттыруға мүмкіндік берді. Алайда, осы компоненттердің жоқ болуы Insertion моделінің тиімділігіне теріс әсер етіп, болжау сапасын төмендетті. Сонымен қатар, Insertion әдісінің ерекшеліктерінің арқасында декодтау қадамдарының саны азайтылды. Бұл басқа модельдермен салыстырғанда оқыту мен сөйлеуді өңдеу процесін жылдамдатуға ықпал етті.

Қорытынды

Бұл жұмыс қазақ тіліндегі сөйлеуді тану үшін CTC және Insertion бірлескен интегралдық модельдерінің әсерін зерттеуге арналады. Эксперименттер қазақ тілінің 405 сағаттық аралас сөйлеу корпусымен жүргізілді және нәтижелер көрсеткендей, жүйе RNN негізіндегі тілдік модельдер қолданылғанда жоғары нәтижелерге қол жеткізе алды. Бұл модельдер негізіндегі декодтау есептеу шығындарын арттырмайды, сондықтан декодтау жылдамдығы баяуламайды. Сонымен қатар, Insertion моделі іске асырылғанда декодтау кезінде итерациялар саны азаяды. Осылайша, CER көрсеткішінің ең жақсы нәтижесі 10.2%-ды құрады, бұл бүгінгі таңда бәсекеге қабілетті нәтиже болып табылады. Ұсынылған әдіс икемді және айнымалылардың шартты тәуелсіздігін талап етпейді. Сонымен қатар, бұл модельді түркі тілдері сияқты туыс тілдерде сөйлеуді тану үшін қолдануға болады деген қорытынды жасауға болады.

Келесі жұмыстарда біз агглютинативті тілдерді тану үшін басқа Transformer негізіндегі модельдерді зерттеуді жоспарлап отырмыз.

Қаржыландыру туралы ақпарат. Бұл зерттеу Қазақстан Республикасының Ғылым және жоғары білім министрлігі Ғылым комитеті тарапынан қаржыландырылды (Грант № BR24992875).

ӘДЕБИЕТТЕР

- 1 Mamyrbayev, O., Oralbekova, D. Modern trends in the development of speech recognition systems. *News of the National Academy of Sciences of the Republic of Kazakhstan*, 4 (332), 42–51 (2020).
- 2 Graves, A., Fernandez, S., Gomez, F., and Schmidhuber, J. Connectionist Temporal Classification: Labelling Unsegmented Sequence Data with Recurrent Neural Networks. *ICML*, Pittsburgh, USA, pp. 369–376 (2006).
- 3 Majhi, M.K., and Saha, S.K. An automatic speech recognition system in Odia language using attention mechanism and data augmentation. *International Journal of Speech Technology*, 27, 717–728 (2024). <https://doi.org/10.1007/s10772-024-10132-6>.
- 4 Cheng, L., Zhu, H., Hu, Z., and Luo, B. A Sequence-to-Sequence Model for Online Signal Detection and Format Recognition. *IEEE Signal Processing Letters*, 31, 994–998 (2024). <https://doi.org/10.1109/LSP.2024.3384015>.
- 5 Mundotiya, R.K., Mehta, A., Baruah, R., and Singh, A.K. Integration of morphological features and contextual weightage using monotonic chunk attention for part of speech tagging. *Journal of King Saud University – Computer and Information Sciences*, 34, 7324–7334 (2021).
- 6 Deng, K., Cao, S., Zhang, Y., Ma, L., Cheng, G., Xu, J., and Zhang, P. Improving CTC-based speech recognition via knowledge transferring from pre-trained language models (2022). <https://doi.org/10.48550/arXiv.2203.03582>.
- 7 Dingliwal, S., Sunkara, M., Ronanki, S., Farris, J., Kirchhoff, K., and Bodapati, S. Personalization of CTC Speech Recognition Models. *2022 IEEE Spoken Language Technology Workshop (SLT)*, Doha, Qatar, pp. 302–309 (2023). <https://doi.org/10.1109/SLT54892.2023.10022705>.
- 8 Zeyer, A., Kazuki, I., Ralf, S., and Hermann, N. Improved training of end-to-end attention models for speech recognition. *arXiv preprint arXiv:1805.03294* (2018).
- 9 Wang, Z., Tao, Z., Shao, Y., and Ding, B. LSTM-convolutional-BLSTM encoder-decoder network for minimum mean-square error approach to speech enhancement. *Applied Acoustics*, 172, 107647 (2021).
- 10 Huang, Z., Wang, P., Wang, J., Miao, H., Xu, J., and Zhang, P. Improving Transformer Based End-to-End Code-Switching Speech Recognition Using Language Identification. *Applied Sciences*, 11 (19), 9106 (2021). <https://doi.org/10.3390/app11199106>.
- 11 Miao, H., Cheng, G., Gao, C., Zhang, P., and Yan, Y. Transformer-Based Online CTC/Attention End-To-End Speech Recognition Architecture. *ICASSP 2020 – 2020 IEEE International Conference on Acoustics, Speech and Signal Processing*, pp. 6084–6088 (2020). <https://doi.org/10.1109/ICASSP40776.2020.9053165>.
- 12 Li, F., Chen, J., and Zhang, X. A Survey of Non-Autoregressive Neural Machine Translation. *Electronics*, 12 (13), 2980 (2023). <https://doi.org/10.3390/electronics12132980>.
- 13 Chen, N., Watanabe, S., Villalba, J., and Dehak, N. Listen and fill in the missing letters: Non-autoregressive transformer for speech recognition. *arXiv preprint arXiv:1911.04908* (2020).
- 14 Fujita, Y., Watanabe, S., Omachi, M., and Chan, X. Insertion-Based Modeling for End-to-End Automatic Speech Recognition. *INTERSPEECH 2020* (2020). <https://doi.org/10.48550/arXiv.2005.13211>.
- 15 Rakhimova, D., Sagat, K., Zhakypbaeva, K., and Zhunussova, A. Development and Study of a Post-editing Model for Russian-Kazakh and English-Kazakh Translation Based on Machine Learning. In: Wojtkiewicz K., Treur J., Pimenidis E., Maleszka M. (eds) *Advances in Computational Collective Intelligence. ICCCI 2021. Communications in Computer and Information Science*, vol. 1463. Springer, Cham (2021). https://doi.org/10.1007/978-3-030-88113-9_42.
- 16 Karyukin, V., Rakhimova, D., Karibayeva, A., Turganbayeva, A., and Turarbek, A. The neural machine translation models for the low-resource Kazakh–English language pair. *PeerJ Computer Science*, 9, e1224 (2023). <https://doi.org/10.7717/peerj-cs.1224>.
- 17 Kydyrbekova, A., and Oralbekova, D. Speaker identification using distribution-preserving x-vector generation. *News of the National Academy of Sciences of the Republic of Kazakhstan, Physico-mathematical series*, 4 (352), 152–162 (2024). <https://doi.org/10.32014/2024.2518-1726.314>.
- 18 Mamyrbayev, O., Alimhan, K., Oralbekova, D., Bekarystankyzy, A., and Zhumazhanov, B. Identifying the influence of transfer learning method in developing an end-to-end automatic speech recognition system with a low data level. *Eastern-European Journal of Enterprise Technologies*, 1 (9(115)), 84–92 (2022). <https://doi.org/10.15587/1729-4061.2022.252801>.

19 Mamyrbayev, O., Oralbekova, D., Kydyrbekova, A., Turdalykyzy, T., and Bekarystankyzy, A. End-to-End Model Based on RNN-T for Kazakh Speech Recognition. 2021 3rd International Conference on Computer Communication and the Internet (ICCCI), pp. 163–167 (2021). <https://doi.org/10.1109/ICCCI51764.2021.9486811>.

20 Soleymanpour, M., Johnson, M.T., Soleymanpour, R., and Berry, J. Synthesizing Dysarthric Speech Using Multi-Speaker TTS for Dysarthric Speech Recognition. ICASSP 2022 – 2022 IEEE International Conference on Acoustics, Speech and Signal Processing, Singapore, pp. 7382–7386 (2022). <https://doi.org/10.1109/ICASSP43922.2022.9746585>.

^{1*}Оралбекова Д.,

PhD, старший научный сотрудник, ORCID ID: 0000-0003-4975-6493,
e-mail: dinaoral@mail.ru

¹Мамырбаев О.,

PhD, профессор, главный научный сотрудник, ORCID ID: 0000-0001-8318-3794,
e-mail: morkenj@mail.ru

¹Еримбетова А.,

PhD, к.т.н., ассоц. профессор, ORCID ID: 0000-0002-2013-1513,
e-mail: aigerian8888@gmail.com

¹Бекарыстанқызы А.,

PhD, старший научный сотрудник, ORCID ID: 0000-0003-3984-2718,
e-mail: akbayan.b@gmail.com

¹Тұрдалыұлы М.,

PhD, старший научный сотрудник, ORCID ID: 0000-0002-1470-3706,
e-mail: m.turdalyuly@gmail.com

¹Институт информационных и вычислительных технологий КН МНВО РК,
г. Алматы, Казахстан

ПРИМЕНЕНИЕ НЕАВТОРЕГРЕССИОННОГО ДЕКОДИРОВАНИЯ ДЛЯ РАСПОЗНАВАНИЯ КАЗАХСКОЙ РЕЧИ

Аннотация

В области распознавания речи интегральные модели постепенно вытесняют традиционные и гибридные подходы. Основным их принципом является автогрессионное декодирование, при котором выходная последовательность формируется слева направо. Однако до сих пор не было доказано, что этот метод дает наилучшие результаты при преобразовании речи в текст. Кроме того, интегральные модели полагаются только на предыдущий контекст, что затрудняет обработку нечетких или искаженных звуков. В связи с этим был предложен метод Insertion, который не использует автогрессионное декодирование и генерирует выходные данные в произвольном порядке. В данной работе рассматривается модель распознавания казахской речи, обученная на основе метода Insertion и коннекционной временной классификации (СТС). Проведенные эксперименты показали, что этот метод позволяет повысить точность распознавания. В отличие от автогрессионных моделей, метод Insertion предоставляет большую гибкость в обработке последовательностей, поскольку не требует строгого порядка формирования выходных данных. Это снижает задержки при декодировании и делает модель более устойчивой к нечетко произнесенным словам. Кроме того, сочетание метода Insertion с СТС позволяет улучшить соответствие аудиоданных и текстовой транскрипции. Это особенно важно для агглютинативных языков, таких как казахский. По результатам экспериментов точность распознавания предложенной модели достигла 10,2%, что делает ее конкурентоспособной на сегодняшний день.

Ключевые слова: коннекционная временная классификация, интегральные модели, распознавание казахской речи, Insertion, Transformer.

^{1*}Oralbekova D.,

PhD, Senior Researcher, ORCID ID: 0000-0003-4975-6493,

e-mail: dinaoral@mail.ru

¹Mamyrbayev O.,

PhD, Professor, Chief Researcher, ORCID ID: 0000-0001-8318-3794,

e-mail: morkenj@mail.ru

¹Yerimbetova A.,

PhD, Cand.Tech.Sc., Associate Professor, ORCID ID: 0000-0002-2013-1513,

e-mail: aigerian8888@gmail.com

¹Bekarystankyzy A.,

PhD, Senior Researcher, ORCID ID: 0000-0003-3984-2718,

e-mail: akbayan.b@gmail.com

¹Turdalyuly M.,

PhD, Senior Researcher, ORCID ID: 0000-0002-1470-3706,

e-mail: m.turdalyuly@gmail.com

¹ Institute of Information and Computational technology, Almaty, Kazakhstan

APPLICATION OF NON-AUTOREGRESSIVE DECODING FOR KAZAKH SPEECH RECOGNITION

Abstract

In the field of speech recognition, end-to-end models are gradually replacing traditional and hybrid approaches. Their main principle is autoregressive decoding, where the output sequence is formed from left to right. However, it has not yet been proven that this method provides the best results in converting speech to text. Moreover, end-to-end models rely solely on the previous context, which complicates the processing of unclear or distorted sounds. In this regard, the insertion method was proposed, which does not use autoregressive decoding and generates output data in an arbitrary order. This paper examines a Kazakh speech recognition model trained using the insertion method and Connectionist Temporal Classification (CTC). The experiments conducted showed that this method improves recognition accuracy. Unlike autoregressive models, the Insertion method provides greater flexibility in processing sequences, as it does not require a strict order for generating output data. This reduces decoding delays and makes the model more robust to poorly pronounced words. Furthermore, combining the Insertion method with CTC improves the alignment of audio data and text transcription. This is especially important for agglutinative languages such as Kazakh. According to the experimental results, the recognition accuracy of the proposed model reached 10.2%, making it competitive today.

Keywords: connectionist temporal classification, end-to-end models, Kazakh speech recognition, insertion, transformer.

Мақаланың редакцияға түскен күні: 24.02.2025

МРНТИ 28.23.15
УДК 004.85

<https://doi.org/10.55452/1998-6688-2025-22-4-31-39>

¹Әбжанова А.Е.,
старший преподаватель, ORCID ID: 0009-0003-7796-1862,
e-mail: abdygappar29@gmail.com
¹Танирбергенов А.Ж.,
к.т.н., и.о. доцента, ORCID ID: 0009-0000-8401-5434,
e-mail: t.adilbek@mail.ru
²Тасуов Б.,
доцент, ORCID ID: 0000-0002-2000-6720,
e-mail: b.tasuov@dulaty.kz
²Тасжурекова Ж.К.,
и.о. доцента, ORCID ID: 0000-0002-8307-9417,
e-mail: tashjurekova@mail.ru
^{1*}Серикбаева С.К.,
PhD, и.о. доцента, ORCID ID: 0000-0002-3627-3321
*e-mail: inf_8585@mail.ru

¹Евразийский национальный университет им. Л.Н. Гумилева, г. Астана, Казахстан

²Таразский университет им. М.Х. Дулати, г. Тараз, Казахстан

ПРИМЕНЕНИЕ ГИБРИДНОЙ МОДЕЛИ МАШИННОГО ОБУЧЕНИЯ ДЛЯ КЛАССИФИКАЦИИ ТИПОВ ПОЧВ

Аннотация

В данной статье представлена гибридная модель машинного обучения, предназначенная для классификации типов почв на основе анализа их геофизических характеристик. Предложенная модель объединяет два алгоритма – RandomForestClassifier и MLPClassifier, что позволяет использовать преимущества ансамблевых методов, обеспечивающих высокую точность классификации, и нейронных сетей, способных выявлять сложные нелинейные зависимости между параметрами. В качестве исходных данных использовались показатели электропроводности, плотности, скорости распространения Р-волн и глубины залегания. Перед обучением модели была проведена предварительная обработка данных, включающая удаление выбросов, стандартизацию и кодирование категориальных признаков. Гибридная архитектура позволила объединить результаты двух моделей с различными весами, что обеспечило оптимизацию точности классификации. Проведен сравнительный анализ эффективности предложенного подхода с альтернативными алгоритмами, включая XGBoost и Keras, на основе метрик Accuracy, F1-score, Precision и Recall. Результаты показали, что гибридная модель достигает точности 96,07%, превосходя по качеству прогнозирования отдельные алгоритмы. Дополнительно выполнена визуализация матриц ошибок, что позволило выявить распределение классов и оценить устойчивость модели. Полученные результаты подтверждают, что комбинирование ансамблевых и нейросетевых методов обеспечивает более стабильные и надежные прогнозы при работе с геофизическими данными. Разработанная модель может быть использована для автоматизированной классификации почв в геотехнических исследованиях, строительстве, сельском хозяйстве и экологическом мониторинге, повышая эффективность анализа и снижая необходимость дорогостоящих лабораторных испытаний.

Ключевые слова: машинное обучение, классификация почв, гибридная модель, RandomForest, MLPClassifier, геофизические параметры, ансамблевые методы, нейросети.

Введение

В последние годы машинное обучение (ML) стало мощным инструментом в области геотехнических исследований, позволяя автоматизировать и значительно повысить точность ана-

лиза почв. Классификация типов грунтов играет ключевую роль в строительстве, сельском хозяйстве и экологии, обеспечивая надежные прогнозы для оптимального проектирования и эксплуатации земельных ресурсов.

Традиционные методы классификации почв, основанные на лабораторных исследованиях и эмпирических моделях, зачастую требуют значительных временных и финансовых затрат. В связи с этим возникает необходимость в разработке автоматизированных подходов, использующих современные алгоритмы машинного обучения для точного определения типа грунта на основе геофизических параметров.

В данном исследовании предлагается гибридная модель, объединяющая ансамблевый метод RandomForestClassifier и нейросетевой MLPClassifier. Данный подход позволяет объединить преимущества двух различных алгоритмов: высокая точность ансамблевых моделей и способность нейросетей выявлять сложные нелинейные зависимости в данных. В ходе работы использовались ключевые геофизические параметры: электропроводность, плотность, скорость распространения Р-волн и глубина залегания.

Результаты моделирования показали, что предложенная гибридная модель обеспечивает точность 96,07%, превосходя по этому показателю отдельные алгоритмы, такие как XGBoost и Keras. Проведенный сравнительный анализ подтвердил, что комбинирование методов машинного обучения позволяет добиться более устойчивых и надежных предсказаний, чем использование одиночных моделей.

В статье представлены принципы построения гибридной модели, методы предобработки данных, алгоритмы классификации и сравнительный анализ полученных результатов. Также рассматриваются ограничения модели и перспективы ее дальнейшего развития для повышения точности и обобщаемости предсказаний в различных географических регионах.

В данной работе [1] исследуется применение различных методов машинного обучения для классификации почв. Проведен сравнительный анализ алгоритмов и выявлены наиболее эффективные модели для прогнозирования типов грунтов. Основное внимание уделяется ансамблевым методам и нейросетям.

В статье [2] описан алгоритм случайных лесов (Random Forest), его принципы работы, преимущества и недостатки, а также области применения в задачах классификации и регрессии. Применение данного алгоритма позволило улучшить точность предсказаний по сравнению с традиционными методами.

Авторы [3] рассматривают применение методов глубокого обучения, включая сверточные и рекуррентные нейронные сети, для автоматической классификации почв на основе спектральных данных. Подчеркивается важность предварительной обработки данных для повышения точности моделей.

В данной работе [4] представлена глобальная база данных почвенных характеристик, созданная с использованием методов машинного обучения. Анализируются преимущества использования таких баз данных в геотехнических исследованиях.

В статье [5] исследуются различные алгоритмы машинного обучения, включая ансамблевые методы и нейросети, применяемые для прогнозирования урожайности и классификации почв. Приведен анализ точности разных моделей.

Рассматривается [6] применение случайных лесов в задачах дистанционного зондирования и классификации ландшафтов. Демонстрируется высокая точность метода по сравнению с традиционными методами классификации.

В статье [7] обсуждается использование случайных лесов для обработки данных дистанционного зондирования и решения задач классификации. Авторы анализируют точность модели и возможности ее применения.

Представлен [8] обзор методов глубокого обучения, их развитие, архитектуры и основные области применения. Особое внимание уделено применению нейросетевых моделей в геотехнических задачах.

В данной [9] книге представлена теория статистического обучения, обсуждаются основы машинного обучения и методы построения прогнозных моделей. Работа является фундаментальным источником для понимания основ ML.

В работе [10] рассматривается применение машинного обучения в геотехнических и геоэкологических исследованиях. Анализируются основные алгоритмы и их эффективность при обработке больших массивов данных.

Анализ приведенных работ показывает, что современные методы машинного обучения активно применяются для классификации почв. Ансамблевые модели, такие как Random Forest, позволяют достигать высокой точности предсказаний, а методы глубокого обучения обеспечивают выявление сложных закономерностей в данных. Наша работа продолжает эту тенденцию, объединяя ансамблевые и нейросетевые методы в единую гибридную модель, что позволяет повысить надежность и точность предсказаний. Таким образом, предложенный подход основан на предыдущих исследованиях, но в то же время предлагает новый уровень комбинированного машинного обучения для решения задач классификации почв.

Материалы и методы

В данном исследовании использовался набор данных, включающий информацию о физических характеристиках почвы, таких как электропроводность, плотность, скорость распространения Р-волн и глубина залегания. Данные были предварительно обработаны: выполнено удаление пропущенных значений, фильтрация выбросов и стандартизация признаков с использованием метода StandardScaler. Для преобразования категориальных переменных применялся LabelEncoder.

Для классификации типов почв были использованы два алгоритма: RandomForestClassifier и MLPClassifier. Метод случайного леса (Random Forest) строит ансамбль деревьев решений, объединяя их прогнозы для повышения точности. Нейросетевой классификатор (MLP) основан на многослойном перцептроне с тремя скрытыми слоями, функцией активации ReLU и оптимизатором Adam. Гибридная модель объединяла предсказания обоих алгоритмов с весами 0,8 и 0,2 соответственно.

Оценка точности модели проводилась с использованием метрики accuracy_score, а также показателей Precision, Recall и F1-score. Данные были разделены на обучающую и тестовую выборки в пропорции 80/20. Для проверки устойчивости модели был проведен кросс-валидационный анализ. Результаты сравнивались с альтернативными алгоритмами, такими как XGBoost и Keras, чтобы определить наиболее эффективный метод классификации почв.

Результаты и обсуждение

Создание, обучение и тестирование модели (create_model.py)

Разработана гибридная модель для прогнозирования типа почвы. Данные загружаются с помощью Pandas, а ключевые параметры (электропроводность, плотность, скорость Р-волны, глубина) используются для обучения. Тип почвы кодируется через LabelEncoder, а затем данные разделяются на 80% для обучения и 20% для тестирования (train_test_split).

Для стандартизации применяется StandardScaler. Обучение проводится на двух моделях:

RandomForestClassifier (случайный лес) – анализирует данные с помощью множества деревьев решений.

MLPClassifier (многослойный перцептрон) – использует нейросеть с тремя скрытыми слоями.

В гибридной модели результаты объединены: RandomForest получил вес 0.8, MLP – 0.2. Это позволило повысить точность прогнозов. Оценка проводилась с помощью accuracy_score, и итоговая точность составила 96.07%.

Модель сохраняется через joblib под названием «hybrid_model1.PKL», что позволяет ее повторно загружать и использовать. Высокая точность достигнута за счет комбинирования двух методов, что делает прогнозы более надежными.

Сравнение моделей (rivals_model.py)

Разработана модель машинного обучения для классификации типов почв на основе характеристик: электропроводность, плотность, скорость Р-волн и глубина.

Данные были загружены и предварительно обработаны через файл CSV.

Мы разделили данные на метки (X) и целевую переменную (y). Для целевой переменной использовался LabelEncoder для преобразования текстовых значений в числовой формат. Позже набор был разделен на обучающую и тестовую части в пропорции 80/20.

Данные были масштабированы с помощью StandardScaler для нормализации меток. Обучение проводилось с использованием трех моделей: гибридной модели (RandomForest + MLP), XGBoost и Keras. Гибридная модель была загружена из предварительно сохраненного файла и использовалась для прогнозирования, объединяя результаты RandomForest и многослойного перцептрона (MLP) с коэффициентами 0,6 и 0,4.

Модель XGBoost была обучена 5 деревьям решений, 5 уровням глубины и скорости чтения 0,001. Модель Keras состояла из трех слоев: входного (64 нейрона), скрытого (32 нейрона) и выходного (количество типов почвы). ReLU использовался в скрытых слоях и softmax в выходных слоях в качестве функции активации. Оптимизатор Adam использовался для обучения модели вместе с функцией затрат sparse_categorical_crossentropy. Процесс обучения модели Keras проводился серией из 10 эпох и 8 измерений.

Были получены прогнозы каждой модели и рассчитаны показатели точности (Accuracy, Precision, Recall, F1 Score). Также была создана матрица ошибок для каждой модели. Результаты сравнения по метрикам были визуализированы с помощью графика, который показывает точность и F1 Score для каждой модели. В результате визуализации было обнаружено, что гибридная модель RandomForest + MLP достигла максимальной точности (0,962), XGBoost показал 0,957, а Keras показал 0,938 (рисунок 1).

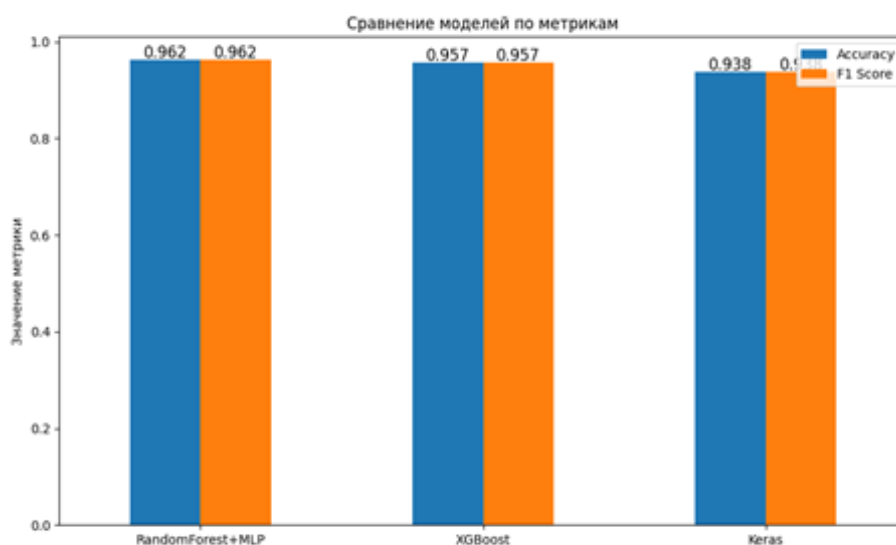


Рисунок 1 – Сравнение моделей по метрикам

Для каждой модели были составлены тепловые карты матриц ошибок, позволяющие оценить распределение прогнозов по отношению к истинным классам.

Рисунок 2 представляет собой три тепловые карты матриц ошибок для различных моделей машинного обучения: RandomForest + MLP (слева), XGBoost (в центре) и Keras (справа). По оси Y отложены истинные классы, по оси X – предсказанные модели классы. Значения внутри клеток показывают количество объектов, отнесенных моделью к определенному классу, а цветовая интенсивность отражает частоту встречаемости предсказаний – чем темнее клетка, тем больше значений приходится на данное соответствие. Эти тепловые карты помогают визуаль-но оценить точность моделей и выявить наиболее частые ошибки классификации.

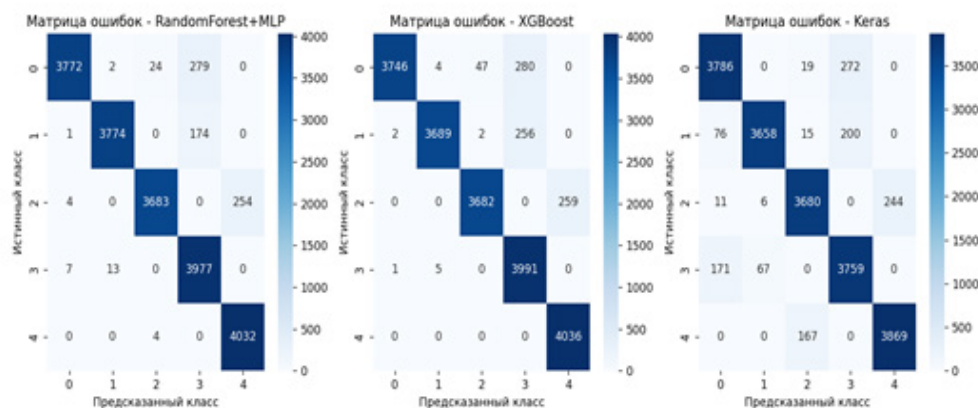


Рисунок 2 – Точность прогноза при классификации типов почв

Таким образом, проведенное исследование показало, что комбинирование нескольких алгоритмов позволяет повысить точность прогноза при классификации типов почв. Модель, созданная на основе комбинации RandomForest и многослойного перцептрона (MLP), сохраненная как «hybrid_model1.pkl», обеспечивала пользователям максимальную точность (0,962).

Данная модель продемонстрировала высокую точность в классификации типов почв за счет объединения алгоритмов RandomForest и MLP, что позволило эффективно использовать их сильные стороны. Полученные результаты подтверждают способность модели корректно обрабатывать разнородные данные и точно дифференцировать классы.

Рисунок 3 показывает распределение различных физических характеристик грунта в датасете. Электропроводность имеет асимметричное распределение с наибольшей частотой значений около 50 мСм/м, плотность сосредоточена в пределах 1.8–2.2 г/см³, скорость Р-волн демонстрирует двухпиковое распределение, что может свидетельствовать о наличии различных типов материалов. Глубина распределена неравномерно, с несколькими локальными максимумами, что может указывать на стратификацию грунта и наличие слоев с разными физическими свойствами.

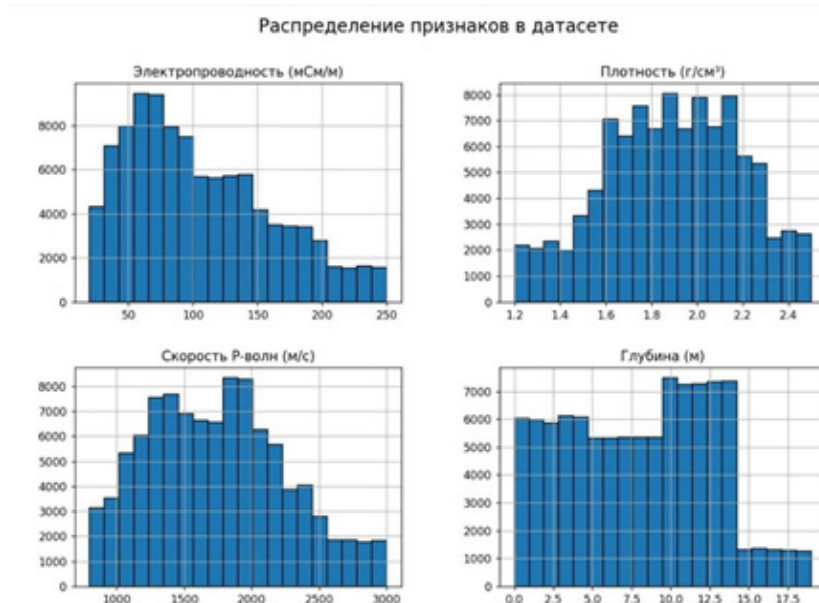


Рисунок 3 – Распределение признаков в датасете

Рисунок 4 представляет собой столбчатую диаграмму распределения классов грунта, включая торф, суглинок, песок, ил и глину. Каждый тип грунта встречается примерно одинаковое количество раз, что говорит о сбалансированности выборки. Это важно для построения точных моделей машинного обучения, так как равномерное представление классов снижает вероятность смещения результатов.

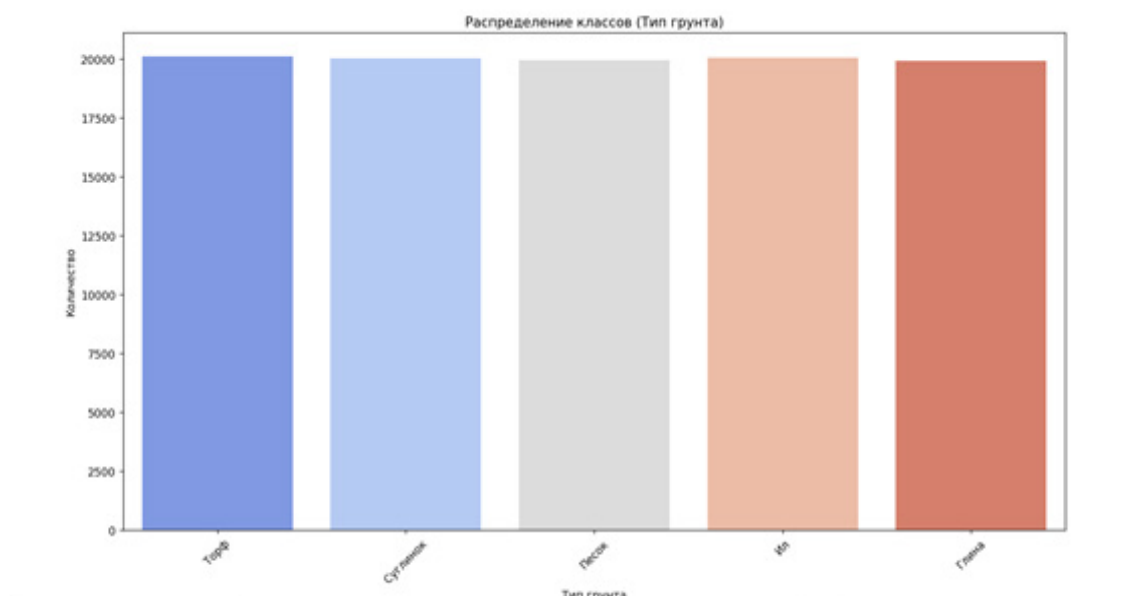


Рисунок 4 – Распределение классов грунта

Модель XGBoost отличается высокой гибкостью и хорошо работает с реальными данными, но требует тщательной настройки параметров. Она особенно эффективна при обработке табличных данных и способна выявлять сложные зависимости, но при этом может быть чувствительна к шуму и выбросам.

Нейронная сеть Keras, использующая методы глубокого обучения, способна изучать сложные модели и находить скрытые закономерности. Однако ее скорость обучения относительно низкая, особенно при работе с большими объемами данных, что требует значительных вычислительных ресурсов.

Эффективность гибридной модели была доказана, поскольку она обеспечивала высокую надежность и точность, давала стабильные результаты и успешно сочетала преимущества XGBoost и нейронных сетей Keras.

Результаты экспериментов показывают, что гибридная модель, комбинирующая RandomForest и MLPClassifier, значительно превосходит отдельные алгоритмы по точности классификации. Комбинированный подход позволяет учесть как линейные, так и нелинейные зависимости между характеристиками почвы, что приводит к снижению ошибок классификации. Проведенный анализ матрицы ошибок подтвердил, что предложенная модель обеспечивает более сбалансированное распределение предсказаний, что особенно важно при работе с разнородными данными.

Кроме того, исследование показало, что использование предварительной обработки данных, таких как нормализация и удаление выбросов, положительно сказывается на точности модели. Сравнение с XGBoost и Keras продемонстрировало, что гибридная модель обеспечивает не только высокую точность, но и устойчивость к изменению параметров данных. В будущем планируется исследование влияния дополнительных геофизических характеристик на качество предсказаний и тестирование модели на новых выборках.

Заклучение

В данной работе предложена гибридная модель машинного обучения, сочетающая алгоритмы случайного леса (Random Forest) и многослойного перцептрона (MLP) для классификации типов почв. Разработанный подход позволил достичь высокой точности прогнозирования (96,07%), что демонстрирует эффективность комбинации ансамблевых методов и нейросетей при анализе почвенных характеристик. В ходе исследования проведено сравнение с альтернативными методами, такими как XGBoost и Keras, подтверждающее преимущество предложенной модели в устойчивости и точности предсказаний.

Сравнительный анализ показал, что гибридный подход превосходит традиционные алгоритмы, обеспечивая более точные и надежные результаты. Применение метрик *assurasy_score*, *Precision*, *Recall* и *F1-score* подтвердило, что предложенный метод стабильно демонстрирует высокие показатели качества классификации.

В дальнейшем предполагается улучшение модели путем оптимизации гиперпараметров, расширения набора данных и тестирования на других географических регионах. Внедрение предложенного подхода может способствовать развитию автоматизированных систем мониторинга почв и повысить точность геотехнических прогнозов в различных сферах применения.

ЛИТЕРАТУРА

- 1 Aydın, A., Keskin, H., Öztürk, A. Use of Machine Learning Techniques in Soil Classification. *Sustainability*, 15 (3), 2374 (2023). <https://doi.org/10.3390/su15032374>.
- 2 Breiman, L. Random Forests. *Machine Learning*, 45(1), 5–32 (2001). <https://doi.org/10.1023/A:1010933404324>.
- 3 Zhang, W., Li, Y., Chen, X. Deep Learning-Based Soil Type Classification. *Computers and Electronics in Agriculture*, 169, 105205 (2020). <https://doi.org/10.1016/j.compag.2019.105205>.
- 4 Hengl, T., Heuvelink, G.B.M., Kempen, B. SoilGrids250m: Global Gridded Soil Information Based on Machine Learning/ *PLoS ONE*, 12 (2), e0169748 (2017). <https://doi.org/10.1371/journal.pone.0169748>.
- 5 Chlingaryan, A., Sukkarieh, S., Whelan, B. Machine Learning Approaches for Crop Yield Prediction and Soil Classification. *Agricultural Systems*, 167, 144–153 (2018). <https://doi.org/10.1016/j.agry.2018.09.012>.
- 6 Gislason, P.O., Benediktsson, J.A., Sveinsson, J.R. Random Forests for Land Cover Classification. *Remote Sensing of Environment*, 110 (4), 435–449 (2006). <https://doi.org/10.1016/j.rse.2007.03.012>.
- 7 Pal, M. Random Forest Classifier for Remote Sensing Classification. *International Journal of Remote Sensing*, 26 (1), 217–222 (2005). <https://doi.org/10.1080/01431160412331269698>.
- 8 Schmidhuber, J. Deep Learning in Neural Networks: An Overview. *Neural Networks*, 61, 85–117 (2015). <https://doi.org/10.1016/j.neunet.2014.09.003>.
- 9 Vapnik, V. *Statistical Learning Theory*. New York: John Wiley & Sons, 1998.
- 10 Gandomi, A.H., Alavi A.H. A Review of Machine Learning Applications in Geotechnical and Geoenvironmental Engineering. *Engineering Geology*, 152, 63–81 (2013). <https://doi.org/10.1016/j.enggeo.2013.05.016>.

¹Әбжанова А.Е.,

аға оқытушысы, ORCID ID: 0009-0003-7796-1862,

e-mail: abdygappar29@gmail.com

¹Танирбергенов А.Ж.,

т.ғ.к., доцент м.а., ORCID ID: 0009-0000-8401-5434,

e-mail: t.adilbek@mail.ru

²Тасуов Б.,

доцент, ORCID ID: 0000-0002-2000-6720,

e-mail: b.tasuov@dulaty.kz

²Тасжурекова Ж.К.,

доцент м.а., ORCID ID: 0000-0002-8307-9417,

e-mail: tashjurekova@mail.ru

¹Серикбаева С.К.,

PhD, доцент м.а., ORCID ID: <https://orcid.org/0000-0002-3627-3321>

*e-mail: inf_8585@mail.ru

¹Л.Н. Гумилев атындағы Еуразия ұлттық университеті, Астана қ., Қазақстан

²М.Х. Дулати атындағы Тараз университеті, Тараз қ., Қазақстан

ТОПЫРАҚ ТҮРЛЕРІН ЖІКТЕУ ҮШІН МАШИНАЛЫҚ ОҚЫТУДЫҢ ГИБРИДТІ МОДЕЛІН ҚОЛДАНУ

Аңдатпа

Бұл жұмыста RandomForestClassifier және MLPClassifier алгоритмдерін біріктіретін топырақ типтерін жіктеуге арналған машиналық оқытудың гибриді моделі ұсынылады. Өзірленген тәсіл ансамбльдік әдістердің жоғары дәлдігін нейрондық желілердің күрделі бейсызық тәуелділіктерді анықтау мүмкіндігімен үйлестіреді. Бастапқы деректер ретінде электр өткізгіштігі, тығыздығы, Р-толқындарының таралу жылдамдығы және жату тереңдігі көрсеткіштері пайдаланылды. Модельді оқытпас бұрын деректер алдын ала өңделіп, ауытқулар жойылды, стандарттау және категориялық белгілерді кодтау жүргізілді. Гибридтік архитектура екі модельдің нәтижелерін түрлі салмақтармен біріктіріп, жіктеу дәлдігін оңтайландырды. Ұсынылған тәсілдің тиімділігі XGBoost және Keras сияқты баламалы алгоритмдермен салыстырылды және Accuracy, F1-score, Precision, Recall метрикалары қолданылды. Зерттеу нәтижелері гибридік модельдің 96,07% дәлдікке жеткенін көрсетті. Алынған нәтижелер геофизикалық деректермен жұмыс істеу кезінде ансамбльдік және нейрондық әдістердің комбинациясы неғұрлым тұрақты және сенімді болжамдар беретінін дәлелдейді. Ұсынылған модельді геотехникалық зерттеулерде, құрылыста, ауыл шаруашылығында және экологиялық мониторингте қолдануға болады.

Тірек сөздер: машиналық оқыту, топырақ классификациясы, гибридік модель, RandomForest, MLPClassifier, геофизикалық параметрлер, ансамбльдік әдістер, нейрондық желілер.

¹**Abzhanova A.,**

Senior Lecturer, ORCID ID: 0009-0003-7796-1862,

e-mail: abdygappar29@gmail.com

¹**Tanirbergenov A.,**

PhD., acting Associate Professor, ORCID ID: 0009-0000-8401-5434,

e-mail: t.adilbek@mail.ru

²**Tassuov B.,**

Associate Professor, ORCID ID: 0000-0002-2000-6720,

e-mail: b.tasuov@dulaty.kz

²**Taszhurekova Zh.,**

acting Associate Professor, ORCID ID: 0000-0002-8307-9417,

e-mail: taszhurekova@mail.ru

^{1*}**Serikbayeva S.,**

PhD, acting Associate Professor, ORCID ID: 0000-0002-3627-3321,

*e-mail: inf_8585@mail.ru

¹L.N. Gumilyov Eurasian National University, Astana, Kazakhstan

²Taraz University named after M.KH. Dulaty, Taraz, Kazakhstan

APPLICATION OF A HYBRID MACHINE LEARNING MODEL FOR SOIL TYPE CLASSIFICATION

Abstract

This article presents a hybrid machine learning model designed for soil type classification based on the analysis of geophysical characteristics. The proposed model combines two algorithms – RandomForestClassifier and MLPClassifier – integrating the high accuracy of ensemble methods with the ability of neural networks to capture complex nonlinear dependencies between parameters. The input dataset included indicators such as electrical conductivity, density, P-wave propagation velocity, and burial depth. Prior to training, data preprocessing was performed, including outlier removal, standardization, and categorical feature encoding. The hybrid architecture allowed the integration of results from both models with different weights, optimizing classification accuracy. The effectiveness of the proposed approach was compared with alternative algorithms such as XGBoost and Keras using metrics including Accuracy, F1-score, Precision, and Recall. The hybrid model achieved an accuracy of 96.07%, outperforming individual algorithms. Visualization of confusion matrices provided insights into class distribution and model robustness. The results confirm that combining ensemble and neural methods ensures more stable and reliable predictions when working with geophysical data. The developed model can be effectively applied in geotechnical studies, construction, agriculture, and environmental monitoring, enhancing analytical efficiency and reducing the need for costly laboratory testing.

Keywords: machine learning, soil classification, hybrid model, RandomForest, MLPClassifier, geophysical parameters, ensemble methods, neural networks.

Дата поступления статьи в редакцию: 24.02.2025

УДК 004.725.4
МРНТИ 50.47.29

<https://doi.org/10.55452/1998-6688-2025-22-4-40-59>

^{1,2*}**Жукабаева Т.К.,**

PhD, профессор, ORCID ID: 0000-0001-6345-5211,

*e-mail: zhukabayeva tk@enu.kz

^{1,3}**Марденов Е.М.,**

магистр, ORCID ID: 0000-0001-9284-9797, e-mail: emardenov@gmail.com

²**Танирбергенов А.Ж.,**

к.т.н., и.о. доцента, ORCID ID: 0009-0000-8401-5434,

e-mail: t.adilbek@mail.ru

¹Международный научный комплекс «Астана», Казахстан, г. Астана

²Евразийский национальный университет им. Л.Н. Гумилева, Казахстан, г. Астана

³Astana International University, Казахстан, г. Астана

КОМПЛЕКСНАЯ МЕТОДИКА ВЫЯВЛЕНИЯ И АНАЛИЗА ИНЦИДЕНТОВ БЕЗОПАСНОСТИ В БЕСПРОВОДНЫХ СЕНСОРНЫХ СЕТЯХ

Аннотация

Статья посвящена вопросам киберфизической безопасности беспроводных сенсорных сетей (БСС). Современные БСС уязвимы к широкому классу атак, таких как sinkhole и wormhole-атаки, атаки «человек посередине», атаки подмены данных, универсальные сетевые атаки и др. На практике возможности по защите БСС от данного вида атак затруднены вследствие разнообразия возможных воздействий, узкоспециализированной направленности инфраструктуры и ограниченности ресурсов узлов сети. В настоящей статье предлагается комплексная методика выявления инцидентов безопасности в БСС для эффективного обнаружения атак и реагирования на инциденты, что позволит минимизировать потенциальный ущерб и обеспечить бесперебойную работу таких сетей. К элементам новизны методики относится ее комплексность, способность выявлять разнообразные киберфизические воздействия и обеспечивать высокую точность и полноту обнаружения инцидентов, учитывая распределенную структуру и динамику изменений в составе узлов БСС. Методика апробирована на фрагменте сети, функционирующей по протоколу ZigBee для контроля характеристик состояния атмосферного воздуха промышленного объекта, города. Разработанная методика будет способствовать повышению качества и своевременности обнаружения инцидентов безопасности в беспроводных сенсорных сетях, что позволит улучшить устойчивость сетей к внешним и внутренним злонамеренным воздействиям и предотвращать длительные перебои в работе инфраструктуры в случае успешных атак.

Ключевые слова: беспроводная сенсорная сеть, инцидент безопасности, выявление, моделирование, методика.

Введение

В настоящее время проблематика киберфизической безопасности беспроводных сенсорных сетей (БСС), и выявления и анализа инцидентов безопасности в частности, представляет крайне важный пласт актуальных научно-технических задач, требующих эффективных подходов к их решению. В первую очередь это связано с все более широким распространением данного вида беспроводных сетей. БСС начинают использоваться практически повсеместно, начиная от мониторинга окружающей среды и заканчивая управлением промышленными процессами и медицинскими приложениями в рамках индустриального интернета вещей, что делает БСС привлекательными объектами для проведения атакующих воздействий со стороны злоумышленников.

По сравнению с проводными и беспроводными сетями общего назначения, работающими по протоколам TCP/IP, специализированные беспроводные сенсорные сети, функционирую-

щие по протоколам LoRa, LPWAN, ZigBee и др., подвержены более значительному числу и большому разнообразию угроз киберфизической безопасности, включающим атаки, связанные с нарушением маршрутизации в сети, такими как sinkhole и wormhole-атаки, не характерные для традиционных сетей общего назначения [1]. При этом, как правило, БСС также остаются подверженными таким видам атак, как атаки типа «человек посередине», спуфинг-атаки, атаки отказа в обслуживании, характерные для различных видов сетей.

Ограниченность ресурсов устройств узлов БСС, в том числе энергоресурсов, ресурса оперативной памяти, вычислительных мощностей, затрудняет внедрение сложных криптографических алгоритмов и методов защиты, что увеличивает уязвимость БСС перед атаками [2]. Вместе с тем с увеличением числа экземпляров узлов в сети управление безопасностью, как правило, становится сложнее. Поэтому необходимо разрабатывать эффективные способы обнаружения атак и реагирования на инциденты для того, чтобы минимизировать ущерб от таких атак и обеспечить бесперебойную и киберустойчивую работу беспроводной сети.

Беспроводные сенсорные сети применяются в современных критически важных областях, таких как телемедицинские системы, транспорт, электроэнергетика и физическая безопасность [3]. Нарушение работы таких сетей может привести к серьезным последствиям, в том числе к угрозам нарушений требований экологической, медицинской, транспортной и промышленной безопасности. При этом в условиях стремительного развития современных технологий, таких как машинное обучение, искусственный интеллект, большие данные, интернет вещей, возникает необходимость оптимизировать и адаптировать существующие методы анализа инцидентов и разработки новых подходов, учитывающих особенности современных технологических решений. Помимо этого, различная информация, передаваемая в БСС, зачастую содержит конфиденциальные данные, как, например, журналы событий безопасности БСС. Поэтому важно предотвращать утечки и несанкционированный доступ к таким данным, одновременно обеспечивая их целостность и конфиденциальность.

Отметим, что потребность в комплексности необходимых средств выявления и анализа инцидентов безопасности БСС обуславливается разнообразием факторов, которые необходимо учитывать в процессе обнаружения инцидентов [4]. В частности, при выявлении инцидентов безопасности необходимо рассматривать все основные известные виды угроз, которым подвержена целевая беспроводная сеть. К таким угрозам относятся угрозы нарушения конфиденциальности, включающие раскрытие секретной информации, например перехват данных. Также должны учитываться угрозы целостности, в том числе связанные с изменением или разрушением данных, к примеру, подделка данных, а также угрозы доступности, включающие ограничение доступа к сервисам или ресурсам, например, универсальные сетевые атаки, такие как DoS-атаки [5]. Кроме того, необходимо учитывать различные производные угрозы, объединяющие одновременно нарушение нескольких свойств безопасности.

Отметим также, что процесс выявления инцидентов должен охватывать как активные атаки, которые вмешиваются в работу сети, в том числе внедрение вредоносного программного обеспечения, так и пассивные атаки, включающие наблюдение за сетевой активностью без какого-либо значимого вмешательства в процесс функционирования БСС и отдельных ее узлов [6]. Вместе с тем обнаружение атакующих воздействий должно охватывать различные атаки, отличающиеся масштабом распространения, начиная от локализованных атак, затрагивающих отдельные узлы или сегменты сети, и заканчивая более масштабными воздействиями, охватывающими в том числе всю беспроводную сенсорную сеть или значимую ее часть. По части объектов приложения воздействий выделяются атаки, направленные на конкретные устройства сети, сенсоры, исполнительные механизмы узлов, атаки на каналы связи, нацеленные на нарушение коммуникации между узлами, а также воздействия на инфраструктуру, включающую серверы, системы управления базами данных и другие сущности.

По методам реализации атак выделяются, во-первых, атаки, включающие в первую очередь аппаратные, физические воздействия на оборудование, к примеру, кражу устройства узла,

повреждение сенсора [7]. Во-вторых, учитывается возможность воздействий на программное обеспечение, эксплуатирующих известные виды уязвимостей в программном обеспечении, вирусы, эксплойты [8]. В-третьих, учитываются также различные комбинированные, в том числе многошаговые атаки, объединяющие воздействия аппаратно-физического и программно-информационного характера [9].

Комплексность процессов обнаружения атак в БСС связана также с потребностью в дифференциации атак в зависимости от уровня возможных рисков и предполагаемого ущерба от них [10]. Могут рассматриваться незначительные последствия, такие как отключение единичного узла, временная потеря доступа к определенному сервису. Пример последствий среднего масштаба – временная потеря доступности к сегменту сети, тогда как значительные последствия, например, включают перехват управления и внедрение вредоносного кода в программное обеспечение большого числа узлов сети с потерей работоспособности всей сети.

По источнику воплощаемой угрозы необходимо учитывать как атаки, осуществляемые извне, так и воздействия, проводимые непосредственно со скомпрометированных узлов БСС. При этом в соответствии с этапами жизненного цикла БСС выделяются проявления атакующих воздействий на этапах разработки, развертывания и проектирования БСС, на этапе эксплуатации БСС, а также при обновлении программных модулей и настроек сети и технической поддержки функционирования. Учитываются также степень скрытности атаки и ее проявления, которые зависят в том числе от различных шаблонов и статистических распределений признаков функционирования при нормальных и аномальных операционных сценариях устройств беспроводной сенсорной сети. К значимым факторам, учет которых целесообразен для комплексного обнаружения инцидентов безопасности в БСС относятся также организационная и техническая сложность [11] обнаружения атак и связанная с этим степень автоматизации выявления инцидентов безопасности, определяющая объемы вовлечения человеческого ресурса оператора сети и/или участия эксперта по киберфизической безопасности [12].

Помимо этого, обнаружение инцидентов безопасности должно учитывать специфику сценария функционирования сети, условия и ограничения по ее расширяемости [13], допустимость изменений пространственного местоположения узлов сети, изменений сетевой топологии БСС [14–15], наличие энергоресурсов и особенности их расходования в процессе нормального функционирования [16].

В настоящее время опубликован ряд научно-исследовательских работ, фокусирующихся в первую очередь на вопросах обнаружения аномалий в функционировании беспроводных сенсорных сетей [17–19], на вопросах выявления универсальных сетевых атак на узлы БСС, таких как различные flooding-атаки и атаки сканирования [20–21], атаки нарушения процессов идентификации узлов и маршрутизации в БСС, такие как wormhole-атаки, sinkhole-атаки и Sybil-атаки [22–24], атаки истощения энергоресурсов узлов БСС и другие. При этом детальный анализ литературы в данной предметной области выявил существенную нехватку подходов и конкретных методик по комплексному выявлению инцидентов безопасности в таких сетях, которые могли бы применяться на практике для повышения осведомленности об уровне безопасности БСС в применении к широкому спектру областей приложений.

Основным вкладом настоящей работы является предложенная методика выявления и анализа инцидентов безопасности в беспроводных сенсорных сетях, которая апробирована на фрагменте БСС, функционирующей по протоколу ZigBee для контроля характеристик состояния атмосферного воздуха промышленного объекта, города.

К элементам новизны данной методики можно отнести ее комплексность и ориентированность на выявление широкого класса киберфизических воздействий в условиях достижения высоких значений показателей точности и полноты обнаружения инцидентов с учетом распределенного характера анализируемой БСС и динамического изменения состава ее узлов. Помимо повышения уровня киберфизической безопасности на практике следование данной методике будет также способствовать снижению рисков возникновения инцидентов безопас-

ности и оптимизации стратегии защиты, соответствующей условиям и требованиям безопасности конкретной беспроводной инфраструктуры.

Оставшаяся часть статьи организована следующим образом. Раздел 1 посвящен анализу основных видов инцидентов киберфизической безопасности БСС. В разделе 2 раскрывается сущность методики, в том числе основные предъявляемые к ней требования и стадии методики. В разделе 3 описывается использования методики на примере БСС для контроля характеристик состояния атмосферного воздуха промышленного объекта, города с использованием микроконтроллеров, датчиков и другого электронного оборудования. В заключение подведены основные итоги и сформулированы дальнейшие шаги по данному направлению исследований.

Анализ инцидентов безопасности в беспроводных сенсорных сетях

Проведенный анализ актуальных материалов в предметной области исследования позволил выявить следующие основные виды инцидентов киберфизической безопасности в беспроводных сенсорных сетях, охватывающие широкий спектр потенциальных угроз, направленных в том числе на нарушение конфиденциальности, целостности и доступности данных в таких сетях (таблица 1). Отметим, что конкретный инцидент в БСС может одновременно иметь признаки нескольких приведенных в таблице 1 видов инцидентов. Вместе с тем перечисление указанных видов инцидентов не является полным, а отображает лишь основные, наиболее актуальные виды проявлений.

Таблица 1 – Основные виды инцидентов киберфизической безопасности в БСС

№	Вид инцидента	Описание	Последствия
1	Утечка данных с узла сенсорной сети	Неконтролируемое распространение чувствительной информации о состоянии узла сети	Утечка конфиденциальных данных может приводить к финансовым потерям, репутационным издержкам или другим негативным последствиям
2	Нарушение целостности данных	Несанкционированное изменение данных, передающихся между узлами сети	Снижение достоверности данных в БСС, неправильные и неоптимальные решения, принимаемые в процессе реагирования на события и инциденты безопасности
3	Отказ в обслуживании	Продолжающаяся во времени перегрузка коммуникационных каналов сети, приводящая к недоступности узлов сети и предоставляемых сервисов	Нарушение нормального функционирования БСС, внесение задержек в процесс коммуникации, потеря фрагментов данных, циркулирующих по сети
4	Несанкционированный доступ	Нелегитимное проникновение атакующего в сеть с целью получения доступа к ресурсам узлов и их использования, в том числе для последующих атак	Потеря доступа и контроля над узлами БСС и их ресурсами
5	Несанкционированная модификация данных	Передача модифицированных и скомпрометированных данных от злоумышленника	Принятие ошибочных решений на основе некорректной информации
6	Физическое вмешательство в работу узлов сети	Физические инциденты, такие как кража, повреждение или модификация, злонамеренные изменения конфигурации оборудования	Полная или частичная приостановка работы БСС, потеря данных, необходимость ремонта или замены оборудования

Продолжение таблицы 1

7	Несанкционированный перехват и анализ трафика	Прослушивание и хранение циркулирующих в сети данных для последующего их анализа	Раскрытие конфиденциальных данных, определение топологии БСС, нахождение открытых уязвимостей в программном обеспечении узлов сети
8	Проявление вредоносного программного обеспечения на узлах сети (вирусы, черви)	Заражение сети вредоносным ПО, направленным на модификацию данных, блокировку работы узлов БСС, хищение информации	Необходимость восстановления сети, включая замену отдельных программно-аппаратных модулей и конфигурационных настроек, финансовые потери, репутационные издержки
9	Подделка данных учетных записей	Компрометация учетных записей пользователей и администраторов сети и сетевых сервисов	Несанкционированный доступ на устройства сети
10	Атаки типа Man-in-the-Middle (человек посередине)	Вмешательство в коммуникационные каналы между узлами БСС с целью перехвата, модификации данных	Подмена передаваемых данных, утечки данных
11	Некорректная конфигурация БСС	Ошибки в настройке беспроводной сети, приводящие в большей уязвимости БСС	Повышение вероятности успешных атак и снижение уровня защищенности БСС

Отметим, что разнообразие видов инцидентов безопасности в беспроводных сенсорных сетях обуславливает необходимость и важность построения комплексной методики для обеспечения безопасности БСС, снижения числа успешно выполненных атакующих воздействий и снижения предполагаемого ущерба от них.

Материалы и методы

Методика выявления и анализа инцидентов безопасности в беспроводных сенсорных сетях

Предлагаемая комплексная методика выявления и анализа инцидентов безопасности в беспроводных сенсорных сетях предназначена для осуществления на этапе функционирования БСС обслуживающим персоналом инфраструктуры сети. Конечной целью методики является повышение уровня защищенности БСС, что выражается в следующих целях:

- ♦ максимизация доли успешно выявленных инцидентов безопасности, произошедших за определенный промежуток времени Δt по отношению ко всему объему инцидентов за данный промежуток времени

$$p_{\Delta t}(I, m) \xrightarrow{m \in M} \max,$$

где p – обобщенный показатель успешно выявленных инцидентов I с использованием методики m , подвергающийся максимизации;

- ♦ минимизации числа выявленных ложно положительных инцидентов безопасности за промежуток времени Δt :

$$f_{\Delta t}(I, m) \xrightarrow{m \in M} \min;$$

♦ снижения затрат коммуникационно-вычислительных ресурсов узлов БСС, выделяемых на функционирование распределенного механизма выявления и анализа инцидентов безопасности, а также достижения цели.

В общем случае одновременное достижение поставленных целей представляет собой многокритериальную оптимизационную задачу, решение которой производится с использованием следующих методов. На первом шаге производится сужение пространства возможных решений на основе метода оптимизации по Парето, при котором находится граница Парето для заданной конфигурации беспроводной сенсорной сети – набор неуплощааемых решений, оптимальных по Парето. К альтернативному варианту можно отнести использование DEA-оптимизации, который позволяет получить набор граничных значений, принимаемых за условно оптимальные [25]. На втором шаге применяется метод свертки с определением весовых коэффициентов на основе важности отдельных формулируемых оптимизационных критериев.

Расчет минимизации коммуникационно-вычислительных ресурсов узлов БСС в условиях дискретных моментов времени в рамках промежутка Δt производится следующим образом:

$$\max_{\Delta t \in T} \left\{ \frac{\max_{t \in \Delta t} (|r^{(p)}|) - \max_{t \in \Delta t} (|r|)}{\max_{t \in \Delta t} (|r|)} \right\} \xrightarrow{m \in M} \min ,$$

где $r = (r_n, r_c, r_e)$ – вектор показателей ресурсопотребления с компонентом r_n для определения затрачиваемых коммуникационных ресурсов, r_c – затрачиваемых вычислительных ресурсов и r_e – энергоресурсов, затрачиваемых в процессе нормального функционирования сети. Аналогичным образом задается векторный показатель $r^{(p)} = (r_n^{(p)}, r_c^{(p)}, r_e^{(p)})$, специфицирующий каждый вид ресурсов, затрачиваемых в условиях функционирования БСС, включающей работу программных модулей выявления и анализа инцидентов безопасности. Величина T задает множество равных по продолжительности промежутков времени Δt , каждый из которых формирует конечную последовательность дискретных моментов времени t для измерения вычислимых показателей.

Отметим, что значения величин g и $g^{(p)}$ представляют собой нормированные средневзвешенные вектор-значения с использованием коэффициентов важности, задаваемых с учетом специфики конкретного сценария функционирования БСС. Нормировка компонентов величин g и $g^{(p)}$ производится на интервале $(0,1)$, где 1 обозначает максимальное потребление соответствующего ресурса. Интегральное значение ресурсопотребления вычисляется следующим образом

$$|r| = \sqrt{(r_n^{(i+1)})^2 + (r_c^{(i+1)})^2 + (r_e^{(i+1)})^2} ,$$

$$imp : r^{(i)} \rightarrow r^{(i+1)} ,$$

$$r_n^{(i+1)} = \frac{k_n \cdot r_n^{(i)}}{k_n + k_c + k_e}$$

$$r_c^{(i+1)} = \frac{k_c \cdot r_c^{(i)}}{k_n + k_c + k_e}$$

$$r_e^{(i+1)} = \frac{k_e \cdot r_e^{(i)}}{k_n + k_c + k_e} , \text{ где}$$

величины k_n , k_c и k_e – коэффициенты важности, в общем случае определяемые рекуррентно, исходя из апостериорно установленных равновесных значений в виде $(0,33; 0,33; 0,33)$,

после чего на основе отображения imp с учетом внешних знаний преобразуются в требуемый вектор для учета критичности используемых ресурсов БСС. При этом алгоритм преобразования коэффициентов производится с учетом знаний об известных уязвимостях программно-аппаратного обеспечения беспроводной сети, извлекаемых из открытых баз данных, таких как CPE для определения характеристик программно-аппаратного обеспечения, CVE – как средства представления уязвимостей и шаблонов атак CAPEC [26].

В частности, из базы данных CAPEC для каждой найденной уязвимости, соответствующей характеристикам CPE, получаются следующие значения:

- ♦ *Likelihood_Of_Attack* – задающее вероятность эксплуатации данной уязвимости; используются три категориальных значения High, Medium и Low, преобразуемые в числовые значения 0,33, 0,66 и 1 соответственно;

- ♦ *Typical_Severity* – определяющее уровень критичности уязвимости, также задаваемое в рамках настоящей методики на основе трех категориальных значений High, Medium и Low, преобразуемые в числовые – 0,33, 0,66 и 1 соответственно.

Отметим, что в рамках текущей реализации методики рекуррентный характер вычисления коэффициентов важности ограничивается одной итерацией применения отображения $imp: r^{(0)} \rightarrow r^{(1)}$, в то время как в общем случае большее число итераций позволит более комплексно учесть разноплановость актуальных видов уязвимостей и дифференциацию по характеристикам программно-аппаратного обеспечения в соответствии с конкретными используемыми идентификаторами из базы CPE и их значений.

При этом для расчета коэффициентов важности для коммуникационного ресурса r_n используются шаблоны атак, принадлежащие категориям «CAPEC-216: Communication Channel Manipulation», «CAPEC-192: Protocol Analysis», «CAPEC-117: Interception» и др., для вычислительного ресурса r_c используются категории «CAPEC-20: Encryption Brute Forcing», «CAPEC-49: Password Brute Forcing», «CAPEC-97: Cryptanalysis» и др., тогда как энерго-ресурса r_e используются, в частности, такие категории, как «CAPEC-604: Wi-Fi Jamming», «CAPEC-124: Shared Resource Manipulation» и «CAPEC-233: Privilege Escalation». Таким образом, коэффициенты важности вычисляются по следующим формулам:

$$k_n = \underset{v \in V_n}{count(v, r_n)} \cdot \underset{v \in V_n}{mean(Likelihood\_Of\_Attack(v))} \cdot \underset{v \in V_n}{mean(Typical\_Severity(v))} ,$$

$$k_c = \underset{v \in V_c}{count(v, r_c)} \cdot \underset{v \in V_c}{mean(Likelihood\_Of\_Attack(v))} \cdot \underset{v \in V_c}{mean(Typical\_Severity(v))} ,$$

$$k_e = \underset{v \in V_e}{count(v, r_e)} \cdot \underset{v \in V_e}{mean(Likelihood\_Of\_Attack(v))} \cdot \underset{v \in V_e}{mean(Typical\_Severity(v))} ,$$

где функция $count$ возвращает общее число уязвимостей, релевантных рассматриваемому сценарию БСС в контексте потребления ресурсов r_n , r_c , r_e соответственно; функция $mean$ возвращает усредненные значения вероятности эксплуатации атаки v и уровень ее критичности. При этом соответствующие множества V_n , V_c и V_e – подмножества уязвимостей, относящихся к расходованию рассматриваемых ресурсов и специфицируемых экспертным путем.

Сформулируем основные группы требований, предъявляемых к методике и определяющих основные критерии проверки ее выполнимости (таблица 2).

Отметим, что выполнение этих требований будет способствовать построению надежной и более гибкой методики выявления и анализа инцидентов безопасности в беспроводных сенсорных сетях, способной адаптироваться к меняющимся с течением времени условиям и угрозам в зависимости от особенностей функционирования конкретной БСС. Рассмотрим основные стадии предлагаемой методики (рисунок 1).

Таблица 2 – Основные виды требований, предъявляемых к методике

№	Группы требований	Требования
1	Универсальность и масштабируемость	- Ориентированность на разные типы БСС, отличающиеся протоколы связи, состав и количество узлов. - Масштабируемость для увеличения числа узлов и датчиков без значительного снижения функциональности, нарушений доступа и задержек, а также повышения интенсивности потоков данных.
2	Качество обнаружения инцидентов	- Высокие точность и полнота выявляемых инцидентов безопасности с возможностью настройки гиперпараметров моделей обнаружения. - Снижение размерности и объемов анализируемых данных за счет удаления случайных выбросов и незначимых событий и их детализации. - Проверка качества работы механизма обнаружения инцидентов на независимых выборках репрезентативных данных достаточных объемов.
3	Оперативность выявления инцидентов безопасности	- Должна обеспечиваться оперативность выявления всех значимых инцидентов безопасности с учетом сценарных ограничений. - Отслеживание инцидентов должно проводиться непрерывно с учетом текущего состояния БСС и ее доступных устройств. - Должны также определяться статистически значимые отклонения от нормального функционирования работы сети, и эти данные необходимо использовать для уточнения и расширения массива анализируемых данных. - Должно соблюдаться ограничение на максимально допустимую задержку между моментом фактического возникновения инцидента и его выявлением.
4	Комплексность методики	- Модульность архитектуры и поддержка возможности интеграции с существующими стандартами и протоколами киберфизической безопасности.
5	Автоматизация методики	- Необходимость автоматизировать наиболее трудоемкие и повторяемые шаги методики по заданному набору правил, в том числе шаги по автоматическому сбору и обработке данных с минимальным участием оператора сети.
6	Защищенность процессов выявления и анализа инцидентов безопасности	- Защищенные хранение и передача исходных и промежуточных данных, а также результатов обнаружения и анализа инцидентов. - Наличие средств резервирования критичных модулей сети и возможности оперативного восстановления работоспособности в случае отказов.
7	Распределенный характер методики	- Необходимость распределения модулей по обнаружению и анализу инцидентов между узлами и устройствами БСС с возможностью применения принципов федеративных вычислений.
8	Эффективное ресурсопотребление	- Необходимость оптимизации затрат на эксплуатацию и обслуживание инфраструктуры БСС и процессов обнаружения и анализа инцидентов безопасности.



Рисунок 1 – Основные стадии методики выявления и анализа инцидентов безопасности в БСС

На стадии 1 методики производится сбор исходных данных о беспроводной сенсорной сети, в том числе:

- ♦ статические данные, включающие состав и топологию сети, спецификацию элементов программно-аппаратного обеспечения узлов с использованием идентификаторов CPE, используемый протокол сетевого взаимодействия, применяемые политики безопасности на основе предикативных правил [27], спецификацию инфраструктуры сети;
- ♦ динамические данные, включающие данные трафика БСС и журналы событий на узлах, в том числе пользовательские команды ввода/вывода на узлах, показания сенсоров и состояния исполнительных механизмов узлов.

Исходные данные собираются в потоковом режиме, обеспечивая выборки необходимых для анализа данных на протяжении функционирования сети.

Стадия 2 включает предобработку собранных на стадии 1 данных, в том числе:

- ♦ унификацию данных – их приведение к единому формату;
- ♦ устранение и исправление ошибочных и пропущенных значений;
- ♦ фильтрацию данных – удаление лишней, ненужной для последующего анализа информации (в том числе шума, случайных выбросов, незначимых аномалий, дублирующихся данных);
- ♦ объединение данных с различных узлов – построение и наполнение единого централизованного хранилища данных с привязкой к узлу-координатору с использованием реляционных и NoSQL-средств хранения. Отметим, что в общем случае такое хранилище может быть децентрализованным – фактически данные могут распределяться в модулях хранения нескольких узлов, но за их управление отвечает узел-координатор;
- ♦ снижение размерности данных, которое может осуществляться с использованием метода главных компонент [28] и автоэнкодеров [29];
- ♦ нормализацию/стандартизацию данных – приведение к универсальным диапазонам значений, что, в частности, позволяет повысить показатели качества обнаружения атак с использованием ряда обучающих моделей, как, например, KNN.

Стадия 3 представляет собой анализ собранных и предобработанных данных. При этом сбор данных и частично их предобработка могут производиться на уровне конечных узлов БСС с возможностью отбора и накопления фрагментов данных для их периодической от-

правки на сторону узла-координатора сети для централизованной обработки и анализа. Анализ данных включает, в частности, возможность балансировки данных по классам атак для пополнения представителей миноритарных классов. Помимо конкретных методов анализа, включающих метод опорных векторов, сверточные нейронные сети и методы на основе деревьев на данной стадии применяются методы ансамблирования классификаторов, включающие бэггинг и стекинг, бустинг, взвешенное и мажоритарное голосования и др. [30]. При этом такое комбинирование направлено в том числе на снижение влияния эффекта переобучения. Таким образом, для каждого актуального вида атак, который требуется иметь возможность обнаруживать эффективным образом, данная стадия включает выполнение следующих шагов:

- ♦ извлечение признаков из предобработанных выборок данных;
- ♦ формирование частной гипотезы относительно видов данных, являющихся первостепенными признаками для обнаружения данного вида атак;
- ♦ конструирование признакового пространства в соответствии с заданной гипотезой;
- ♦ выделение обучающих и валидационных выборок;
- ♦ определение требуемых значений показателей качества обнаружения, включающих точность, полноту, f1-меру, а также показателей оперативности и ресурсопотребления;
- ♦ выбор моделей и методов обучения для бинарной классификации;
- ♦ определение значений гиперпараметров для выбранных обучающих моделей и методов;
- ♦ определение разметки данных;
- ♦ обучение с применением методов классификации, кластеризации, регрессии;
- ♦ комбинирование бинарных классификаторов в рамках мультиклассовой классификации;
- ♦ проверка показателей путем тестирования на валидационных выборках. При недостижении требуемых значений показателей возврат на шаг формирования уточнения гипотезы.

Стадия 4 включает визуальную аналитику, выполняемую параллельно методам машинного обучения в целях визуального контроля анализируемых данных, в том числе результатов машинного анализа со стороны оператора целевой системы. Кроме того, визуальный анализ может быть применен для уточнения отдельных шагов стадии 3, в том числе для уточнения и корректировки сформированной гипотезы для проведения машинного обучения и формирования признакового пространства. Стадия 4 включает следующие шаги:

- ♦ подготовку источников и выборок данных для визуального анализа;
- ♦ подбор визуальных моделей под различные виды атакующих воздействий, а также параметров этих моделей, таких как число отображаемых слоев, осей, параметров группировки, типов используемых графовых структур и пр.;
- ♦ проверку и экспертную оценку построенных моделей на тестовых наборах данных, в том числе формирование и уточнение инцидентов безопасности, включающих информацию о типах атак, приведших к выявленному инциденту, устройствах – источниках атак и устройствах, вовлеченных в атаку, временные метки старта, завершения вредоносной активности и другую сопутствующую информацию.

На стадии 5 методики осуществляется поддержка принятия решений по реагированию на выявленные инциденты безопасности, которая включает следующие действия:

- ♦ принятие решений по реагированию на обнаруженные инциденты безопасности в БСС. Ключевыми факторами здесь являются категории активов, узлов БСС данных, которые непосредственно связаны с выявленным инцидентом, критичность этих активов и возможные последствия данного инцидента. Основными показателями, на основе которых осуществляется поддержка принятия решения являются величина ущерба от инцидента и риск возможного продолжения действия инцидента в случае продолжающейся во времени атаки;
- ♦ выработка мер по реагированию на выявленный инцидент, включающая принятие мер для каждого актива БСС, вовлеченного в инцидент с учетом целевых функций и ограничений

узлов, сценариев функционирования БСС. В частности, к подобным контрмерам относятся повышение отказоустойчивости узлов и предоставляемых сетью информационных сервисов, резервирование узлов, каналов связи и данных, необходимых в процессе работы БСС;

- ♦ уточнение и корректировка политик безопасности, определяющих, в частности, правила контроля доступа в сети, фильтрации трафика и проверки служебных команд, пересылаемых по сети;

- ♦ расследование, включающее выяснение причин возникновения выявленного инцидента и выяснение сопутствующей информации вектора атаки, включающей время, локации возникновения атак, вовлеченных пользователей и использованные данные и команды. В частности, сбор следов атаки включает сбор и сохранение данных, связанных с инцидентом, системных журналов, трафик и дампы памяти узлов. В случае фрагментарного понимания процесса атаки применяется аналитическое моделирование атаки, позволяющее воспроизвести и оценить ее возможные сценарии с точки зрения их выполнимости, организационно-технической сложности, оперативности и ресурсопотребления;

- ♦ непрерывный мониторинг состояния БСС, отдельных узлов и компонентов с целью устранения последствий атаки и предотвращения ее последующих итераций.

Результаты и обсуждение

Предложенная методика апробирована на фрагменте разработанного программно-аппаратного стенда киберфизической системы для контроля характеристик состояния атмосферного воздуха промышленного объекта, города с использованием микроконтроллеров, датчиков и другого электронного оборудования. Реализованный фрагмент БСС включает в свой состав узлы на основе микроконтроллеров Arduino Uno, беспроводные коммуникационные интерфейсы XBee, датчики температуры и влажности.

На данном фрагменте БСС был промоделирован инцидент киберфизической безопасности, представляющий собой следующую комбинированную атаку, осуществляемую с использованием натурного моделирования. Моделируемая атака включает, во-первых, воздействия типа wormhole [31–36], направленные на нарушение процесса корректной маршрутизации пакетов данных в ZigBee-сети и, во-вторых, spoofing-воздействие для подмены значений показаний части сенсоров БСС. На рисунке 2 показан фрагмент интерфейса моделирования сценариев работы БСС с заданием характеристик состояния атмосферного воздуха, снимаемых на узлах сети.

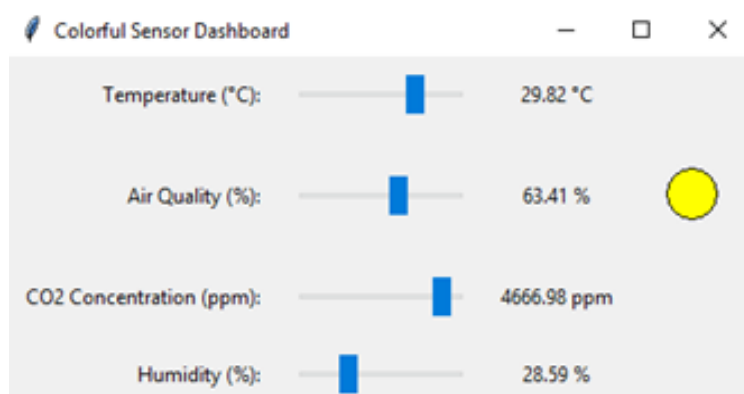


Рисунок 2 – Интерфейс моделирования сценариев работы БСС

В таблицах 3–7 приведены сведения о выполнении 5 стадий предложенной методики в рамках данного демонстрационного сценария.

Таблица 3 – Выполнение стадии сбора исходных данных о БСС

Стадия методики	Действия	Описание
Стадия 1. Сбор исходных данных о БСС	Сбор статических данных БСС	Топология сети в составе 6 узлов с ролью роутера ($R_0, R_1, R_2, R_3, R_4, R_5, R_6$), отвечающих за обеспечение связности и осуществляющих маршрутизацию в БСС, а также узел с ролью координатора C, относящиеся к уникальному идентификатору сети PAN ID. Топология сети задается в соответствии со схемой, приведенной на рисунке 2, с использованием пар вида (R_i, R_j), где $i, j = 0..6$ и $i \neq j$. Используемое программно-аппаратное обеспечение специфицируется использованием идентификаторов CPE, как, например, с помощью «cpe:/h:raspberrypi:raspberrypi_3_model_b%2b:-» задается используемый на узле C одноплатный компьютер Raspberry Pi 3B+, где параметр /h обозначает указание конкретного физического устройства, работающего по протоколу ZigBee с использованием модулей XBee series 2. Используемые политики фильтрации задаются с использованием правил вида $N_k \rightarrow N_j$, где k и j задают номера узлов, между которыми коммуникация на прикладном уровне сетевого взаимодействия разрешена. Узлы задаются при помощи уникальных 8-байтовых адресов, представленных в шестнадцатеричной форме.
	Сбор динамических данных БСС	Считываемые в процессе функционирования данные БСС включают: – исходящие и входящие на узел пакеты данных, каждый из которых включает стартовый байт, длину и тип фрейма, 64-битный адрес получателя, уровень приема сигнала RSSI, возможные дополнительные опции фрейма, полезную нагрузку и контрольную сумму, используемую для проверки целостности. – события $\{e_t(n_i)\}_{n_i \in N, t \in \Delta t}$, собираемые в журналах на узлах n_i в пределах промежутка Δt. Пример такого события – AT-команды ND, направляемая узлу-получателю в виде широковещательного запроса для сканирования ZigBee-сети для получения сведений об актуальных узлах-соседях данного узла. – показания цифровых и аналоговых сенсоров $\{(s_i, p_i)\}$, которые могут быть подключены непосредственно к цифровым или аналоговым пинам микроконтроллера, одноплатного компьютера или беспроводного интерфейса XBee. В рамках разработанного стенда на каждом узле $R_0 - R_6$ осуществляется периодическое снятие показаний температуры, концентрации CO_2 и влажности в качестве ключевых характеристик состояния атмосферного воздуха. Вместе с тем специализированных исполнительных механизмов в рамках стенда не предусмотрено.

Таблица 4 – Выполнение стадии предобработки данных

Стадия методики	Действия	Описание
Стадия 2. Предобработка данных	Унификация данных	На данной стадии осуществляется приведение разнородных данных, получаемых из различных источников, к единому формату представления данных. Формируемые события системы мониторинга состояния атмосферного воздуха записываются в JSON-формате. Ниже приведен пример события получения значения датчика температуры с узла R_2 : <pre>{ "event_id": "123", "timestamp": "2025-03-05T10:30:00Z", "node_id": "R02", "location": { "lat": 51.166667, "lon": 71.433333 }, "sensor_data": { "sensor_type": "temperature", "value": 19.3, "status": "in_progress" }, "alerts": [{ "type": "info", "message": "Inside threshold bounds." }] }</pre>
	Устранение и исправление ошибочных и пропущенных значений	– фильтрация событий с неполными данными. – фильтрация событий с данными вне заданных диапазонов на основе ограничений вида $s_i(p_i) \in [s_i^{(min)}, s_i^{(max)}]$ на заданные сенсоры температуры, влажности, углекислого газа и пр. для каждого из узлов.
	Фильтрация данных	– удаление лишней, ненужной для последующего анализа информации, в том числе удаление показаний дублирующих сенсоров или сенсоров, значений которых не требуется сохранять в соответствии с бизнес-логикой и настройками оператора БСС
	Объединение данных с различных узлов	– построение и наполнение единого централизованного хранилища данных, использованием NoSQL базы MongoDB.
	Снижение размерности данных	В целевой БСС снижение размерности направлено в первую очередь для оптимизации объемов размеров хранилища данных в условиях ограничений узлов БСС, осуществляемое на узле-координаторе в рамках Raspberry Pi с использованием библиотеки Keras <pre>autoencoder = Model(input_layer, decoded) autoencoder.compile(optimizer='adam', loss='binary_crossentropy') autoencoder.fit(X_train, X_train, epochs=50, batch_size=256, shuffle=True, validation_split=0.2)</pre>
	Нормализация данных	– в целях обработки данных осуществляется min-max-нормализация с использованием библиотеки Sklearn <pre>scaler = MinMaxScaler(feature_range=(1, 2)) event = scaler.fit_transform(data)</pre>

Таблица 5 – Выполнение стадии анализа данных

Стадия методики	Действия	Описание
Стадия 3. Анализ данных	Извлечение признаков из предобработанных выборок данных	В части собираемых на узлах БСС исходящих и входящих пакетов беспроводного трафика БСС извлекаемые признаки характеризуют конкретные экземпляры пакетов данных, причем каждый пакет включает следующие признаки: номер последовательности пакета, тип пакета, идентификатор PAN ID сети, MAC-адрес источника, MAC-адрес получателя, длину полезной нагрузки, метку времени фактической доставки пакета на узел и некоторые другие поля. В части событий журналов узлов БСС анализируемые образцы данных включают идентификатор команды управления (Command ID), тип команды, основные параметры (название параметра и значение), Cluster ID для указания типа взаимодействия. В части цифровых и аналоговых сенсоров в первую очередь извлекаются тройки (n_i, s_j, p_j, ts) , специфицирующие узел n_i , на котором располагается сенсор s_j , показание p_j с меткой времени ts .
	Формирование гипотезы для классификации	Формирование экспертным путем гипотезы о предполагаемой корреляции отдельных видов данных и целевой переменной классификации инцидентов безопасности.
	Конструирование признакового пространства	Следующие производные признаки используются в зависимости от вида инцидентов, которые необходимо выявлять: – статистики по узлам-источникам и получателям пакетов данных; – объемы полезной нагрузки пакетов данных; – статистические величины интервалов времени между пакетами данных; – дискретные значения плотности распределения пакетов данных по их типам; – различные статистические величины интенсивности трафика, в том числе максимальные, минимальные и усредненные значения задержек передачи данных, скорости поступления пакетов и байт в секунду; – доля потерянных и/или неправильно доставленных пакетов данных; – величины энергопотребления узлов БСС; – местоположения узлов БСС (с использованием GPS-координат или иного вида позиционирования в пространстве); – временные статистики доступности и активности узлов БСС и др.
	Выделения обучающих и валидационных выборок	Для обучения классификаторов применяется типовое разбиение на обучающую и тестовую выборки в соотношении 0,8 к 0,2 с учетом суммарного объема выборок и сбалансированности данных по классам.
	Определение требуемых значений показателей качества обнаружения	В качестве референсных значений по умолчанию выбраны значения точности, полноты и f1-меры, а также показателя ROC-AUC в размере 0,95.
	Выбор моделей и методов обучения	В качестве обучающих моделей используются следующие модели, осуществляющие классификацию с учителем на основе деревьев (Random Forest и AdaBoost), сверточная нейронная сеть на основе матриц изображений (CNN) Кроме того, применяется автоэнкодер для снижения размерности данных.
	Разметка данных и обучение с применением методов классификации	Разметка данных на их принадлежность нормальному и аномальному функционированию осуществляется экспертно с применением автоматизирующего скрипта, написанного на Python. Разметка конкретных образцов данных (событий, трафика, показаний сенсоров) проводится в первую очередь исходя из меток времени смоделированных инцидентов безопасности, а также идентификаторов узлов, инициаторов атаки и вовлеченных в нее.
	Комбинирование классификаторов	При недостижении установленных значений показателей качества классификации, а также в целях мультиклассификации одним из путей их повышения – дополнительное использование ансамблирования, включая стекинг и мажоритарное голосование.
	Тестирование классификаторов	Типовое тестирование построенных классификаторов осуществляется на установленной тестовой выборке данных.

Таблица 6 – Выполнение стадии визуальной аналитики

Стадия методики	Действия	Описание
Стадия 4. Визуальная аналитика	Подготовка источников и выборка данных для визуального анализа	Анализ источников данных об узлах БСС и их функционировании, в том числе журналы событий, сетевой трафик и показания сенсоров узлов.
	Подбор визуальных моделей	Используются следующие три основные визуальные модели: – график временного ряда для выявления и анализа аномальной сетевой активности при взаимодействии нескольких узлов БСС; – двумерные и трехмерные диаграммы рассеяния для определения отклонений в ключевых характеристиках событий, пакетов, показаний сенсоров, например для учета поддельных данных характеристик сенсоров температуры, влажности, энергопотребления, свидетельствующих о злонамеренной подделке данных мониторинга БСС. Кроме того, в случае контроля характеристик состояния атмосферного воздуха промышленного объекта или города такая модель визуализации используется также для определения кластеров нормального функционирования системы, что, в свою очередь, повысит осведомленность оператора БСС и повысит качество выявления аномалий; – тепловая карта, которая накладывается на фактическую схему топологии БСС для отображения интенсивности сетевой активности, энергопотребления, маршрутизации пакетов и других характеристик, изменения в которых являются признаками атак на БСС.
	Проверка и экспертная оценка визуальных моделей	Оценивание моделей на основе результатов их применения в случае конкретного фактического или моделируемого инцидента безопасности включает: – формирование машиночитаемых информационных признаков, подаваемых в качестве дополнительной информации на вход интеллектуальным моделям анализа данных в рамках стадии 3 методики; – оценку корректности и качества первичных данных, получаемых от БСС, в том числе данных нормального функционирования сети; – выявление дрейфа в данных с течением времени, что необходимо для своевременной корректировки и переобучения интеллектуальных моделей для выявления инцидентов безопасности; – для возможности повторного использования положительного опыта машинного обучения для выявления инцидентов на сходных кейсах (т.н. transfer learning); – для документирования и повышения человекочитаемости результатов интеллектуального анализа и обнаружения инцидентов в целях повышения эффективности обмена экспертными данными и расследования уже выявленных инцидентов.

Таблица 7 – Выполнение стадии поддержки принятия решений

Стадия методики	Действия	Описание
Стадия 5. Поддержка принятия решений	Принятие решений по реагированию на обнаруженные инциденты безопасности в БСС и выработка мер по реагированию на выявленный инцидент	Для БСС для контроля характеристик состояния атмосферного воздуха промышленного объекта города на данной стадии выполняются следующие действия: – оценка величины ущерба сервисам мониторинга состояния атмосферного воздуха, включая последствия от неработоспособности БСС и финансовые затраты на ремонт оборудования инфраструктуры сети; – анализ возможных последствий дальнейшего продолжения и распространения анализируемой атаки; – принятие решения о дальнейших контрмерах, в том числе продолжение функционирования в ограниченном, автономном режиме; – документирование инцидентов в целях обеспечения расследования инцидентов.
	Уточнение и корректировка политик безопасности	Производится уточнение и корректировка политик безопасности для повышения устойчивости сети к будущим атакам, в том числе для: – дифференциации функций и минимизации привилегий узлов; – повышения защищенности БСС на аппаратном уровне; – повышения эффективности процессов аварийного восстановления после инцидентов; – резервирования узлов, мощностей, компонентов защищенного хранения данных.

Результаты классификации на основе методов Random Forest, AdaBoost и сверточной нейронной сети (CNN), выполненных на наборе предобработанных данных, включающих смоделированную комбинированную атаку с использованием wormhole и spoofing-воздействий приведены в таблице 8. В качестве наиболее важного показателя для оценки качества работы классификаторов и их сравнения был выбран показатель f1-меры, представляющий баланс между точностью выявления атаки и полнотой. Показатель приведен с округлением до 2 знаков после запятой.

Таблица 8 – Результаты классификации

Классификатор	F1-мера
Random Forest	0,97
AdaBoost	0,98
CNN	0,67

Таким образом, методы Random Forest и AdaBoost позволяют осуществить классификацию с интегральным показателем качества превышающем референсное значение 0,95, что, в свою очередь, подтверждает выполнимость предлагаемой методики на примере практического сценария работы БСС в рамках построенного стенда киберфизической системы для контроля характеристик состояния атмосферного воздуха промышленного объекта/города. Также отметим, что относительная схожесть результатов классификации для методов Random Forest и AdaBoost обуславливается тем, что оба из них представляют ансамблированные методы, построенные на схожих принципах и основанные на деревьях решений. При этом лучшие результаты классификации у метода AdaBoost объясняются взаимосвязанностью отдельных

деревьев решений в процессе итерационного перебора, что более существенно для набора анализируемых данных.

Последующее совершенствование моделей классификации включает также отработку моделей при увеличении числа тестовых сценариев, а также расширение экспериментов для охвата других разновидностей комбинированных многошаговых воздействий на узлы БСС. В том числе предполагается анализ данных функционирования БСС под атакой, в рамках которой нарушитель использует поддельные идентификаторы узлов и манипулирует ими для перехвата данных и нарушений маршрутизации в сети.

Отметим, что в рамках предлагаемой методики расчет коэффициентов важности производится в целях оптимизации затрат аппаратных, коммуникационно-вычислительных ресурсов БСС, выделяемых на функционирование механизма обнаружения атак. Учитываются три вида ресурсов – вычислительные, коммуникационные и энергоресурсы. В зависимости от конкретного прикладного сценария коэффициенты позволяют выбирать те классификаторы из имеющихся, затраты на функционирование которых сбалансированы и в наибольшей степени близки к ресурсным требованиям рассматриваемого сценария. Например, в случае мобильных сетей мониторинга местности, узлы которых монтируются на беспилотные летательные аппараты, характеризующиеся строго ограниченным максимальным временем полета, наиболее важным представляется энергоресурс, тогда как вычислительные и коммуникационные ресурсы оказывают меньшее влияние на достижение целей миссии. И поэтому вычислительными и коммуникационными ресурсами в данном вопросе, вероятно, можно пренебречь.

Отметим, что в общем случае разработанная методика позволяет обеспечить значительное улучшение качества и скорости обнаружения инцидентов безопасности в беспроводных сенсорных сетях, а также оперативно выявлять как внешние, так и внутренние угрозы. Это способствует повышению устойчивости таких сетей к различного вида атакам, таким как DDoS, подмена данных, манипуляции с маршрутизацией и другие типы злонамеренных воздействий. Кроме того, это будет способствовать снижению рисков возникновения длительных перебоев в работе критически важных объектов, производственных комплексов, инфраструктур медицинских учреждений и городских служб мониторинга окружающей среды. В результате внедрения методики персонал сможет быстрее реагировать на инциденты, локализовать их влияние и минимизировать возможный ущерб, что приведет к поддержанию стабильности и надежности всей инфраструктуры. Также накопление данных о выявленных инцидентах даст возможность прогнозировать и предотвращать будущие атаки, что дополнительно повышает уровень киберфизической безопасности БСС.

Заключение

В статье исследуются вопросы киберфизической безопасности беспроводных сенсорных сетей и особенности обнаружения атак в таких сетях с учетом их архитектурных и ресурсных ограничений. Предложена комплексная методика для выявления инцидентов, способствующая более быстрому обнаружению широкого класса атак и минимизации ущерба от них. В качестве направлений дальнейшей работы по данной тематике предполагаются исследования по применению алгоритмов и средств обнаружения состязательных атак (adversarial attacks) на механизмы киберфизической безопасности и мониторинга уровня защищенности БСС.

Информация о финансировании. Работа проводится в рамках АР23489127 «Модели и алгоритмы повышения защищенности киберфизических систем промышленного интернета вещей с использованием граничных вычислений».

ЛИТЕРАТУРА

- 1 Alansari, Z., Anuar, N.B., Kamsin, A., Belgaum, M.R. A systematic review of routing attacks detection in wireless sensor networks. *PeerJ Computer Science*, 8, e1135 (2022).
- 2 Kaushal, K., Kaur, T. A survey on attacks of WSN and their security mechanisms. *International Journal of Computer Applications*, 118 (18) (2015).
- 3 Ali, A., Ming, Y., Chakraborty, S., Iram, S. A comprehensive survey on real-time applications of WSN. *Future Internet*, 9 (4), 77 (2017).
- 4 Kamaruzzaman, M., Chandra, A. Integration of wireless sensor network in robotics. *Machine Learning for Robotics Applications*, 71–84 (2021).
- 5 Elsadig, M.A. Detection of denial-of-service attack in wireless sensor networks: A lightweight machine learning approach. *IEEE Access*, 11, 83537–83552 (2023).
- 6 Ismail, S., El Mrabet Z., Reza, H. An ensemble-based machine learning approach for cyber-attacks detection in wireless sensor networks. *Applied Sciences*, 13 (1), 30 (2022).
- 7 Sharma, N., Kaushik, I., Agarwal, V.K., Bhushan, B., Khamparia, A. Attacks and security measures in wireless sensor network. *Intelligent Data Analytics for Terror Threat Prediction: Architectures, Methodologies, Techniques and Applications*, 237–268 (2021).
- 8 Oztoprak, A., Hassanpour, R., Ozkan, A., Oztoprak, K. Security challenges, mitigation strategies, and future trends in wireless sensor networks: A review. *ACM Computing Survey*, 57 (4), 1–29 (2024).
- 9 Monjur, M.M.R., Heacock, J., Calzadillas, J., Mahmud, M.S., Roth, J., Mankodiya, K., Yu, Q. Hardware security in sensor and its networks. *Frontiers in Sensors*, 3, 850056 (2022).
- 10 Chen, Y.Y., Xu, B., Long, J. Information security assessment of wireless sensor networks based on bayesian attack graphs. *Journal of Intelligent & Fuzzy Systems*, 41 (3), 4511–4517 (2021).
- 11 Subasini, C.A., Karuppiah, S.P., Sheeba, A., Padmakala, S. Developing an attack detection framework for wireless sensor network- based healthcare applications using hybrid convolutional neural network. *Transactions on Emerging Telecommunications Technologies*, 32 (11), e4336 (2021).
- 12 Delwar, T.S., Aras, U., Mukhopadhyay, S., Kumar, A., Kshirsagar, U., Lee, Y., et al. The intersection of machine learning and wireless sensor network security for cyber-attack detection: a detailed analysis. *Sensors*, 24 (19), 6377 (2024).
- 13 Premkumar, M., Ashokkumar, S.R., Jeevanantham, V., Mohanbabu, G., AnuPallavi, S. Scalable and energy efficient cluster based anomaly detection against denial of service attacks in wireless sensor networks. *Wireless Personal Communications*, 129 (4), 2669–2691 (2023).
- 14 Chen, N., Qiu, T., Daneshmand, M., Wu, D.O. Robust networking: Dynamic topology evolution learning for Internet of Things. *ACM Transactions on Sensor Networks (TOSN)*, 17 (3), 1–23 (2021).
- 15 Duan, G., Lv H., Wang H., Feng G., Li X. Practical cyber attack detection with continuous temporal graph in dynamic network system. *IEEE Transactions on Information Forensics and Security* (2024).
- 16 Nguyen, V.L., Lin, P.C., Hwang, R.H. Energy depletion attacks in low power wireless networks. *IEEE Access*, 7, 51915–51932 (2019).
- 17 Poornima, I.G.A., Paramasivan, B. Anomaly detection in wireless sensor network using machine learning algorithm. *Computer Communications*, 151, 331–337 (2020).
- 18 Ayadi, A., Ghorbel, O., Obeid, A.M., Abid, M. Outlier detection approaches wireless sensor networks: A survey. *Computer Networks*, 129, 319–333 (2017).
- 19 Rajasegarar, S., Leckie, C., Palaniswami, M. Anomaly detection in wireless sensor networks. *IEEE Wireless Communications*, 15 (4), 34–40 (2008).
- 20 Lakshmi, H.N., Anand, S., Sinha, S. Flooding attack in wireless sensor network-analysis and prevention. *International Journal of Engineering and Advanced Technology*, 8 (5), 1792–1796 (2019).
- 21 Dubey, A., Meena, D., Gaur, S. A survey in hello flood attack in wireless sensor networks. *International Journal of Engineering Research and Technology*, 3, 1882–1887 (2014).
- 22 Hu, Y.C., Perrig, A., Johnson, D.B. Wormhole attacks in wireless networks // *IEEE Journal on Selected Areas in Communications*, 24 (2), 370–380 (2006).
- 23 Rehman, A.U., Rehman, S.U., Raheem, H. Sinkhole attacks in wireless sensor networks: A survey. *Wireless Personal Communications*, 106, 2291–2313 (2019).
- 24 Xiao, L., Greenstein, L.J., Mandayam, N.B., Trappe, W. Channel-based detection of sybil attacks in wireless networks. *IEEE Transactions on Information Forensics and Security*, 4 (3), 492–503 (2009).

25 de Oliveira, M.S., Steffen, V., de Francisco, A.C., Trojan, F. Integrated data envelopment analysis, multi-criteria decision making, and cluster analysis methods: Trends and perspectives. *Decision Analytics Journal*, 8, 100271 (2023).

26 Jiang, Y., Atif, Y., Ding, J. Cyber-physical systems security based on a cross-linked and correlated vulnerability database. *International Conference on Critical Information Infrastructures Security* (Cham: Springer International Publishing, 2019), pp. 71–82.

27 Marsh, D.W., Baldwin, R.O., Mullins, B.E., Mills, R.F., Grimaila, M.R. A security policy language for wireless sensor networks. *Journal of Systems and Software*, 82 (1), 101–111 (2009).

28 Livani, M.A., Abadi, M. A PCA-based distributed approach for intrusion detection in wireless sensor networks. *2011 International Symposium on Computer Networks and Distributed Systems (CNDs)* (IEEE, 2011), pp. 55–60.

29 Luo, T., Nagarajan, S.G. Distributed anomaly detection using autoencoder neural networks in WSN for IoT. *2018 IEEE International Conference on Communications (ICC)* (IEEE, 2018), pp. 1–6.

30 John, A., Isnin, I.F.B., Madni, S.H.H., Faheem, M. Cluster-based wireless sensor network framework for denial-of-service attack detection based on variable selection ensemble machine learning algorithms. *Intelligent Systems with Applications*, 22, 200381 (2024).

31 Zahra, F., Jhanjhi, N.Z., Brohi, S.N., Khan, N.A., Masud, M., AlZain, M.A. Rank and wormhole attack detection model for RPL-based internet of things using machine learning. *Sensors*, 22 (18), 6765 (2022).

32 Alghamdi, R., Bellaiche, M. A cascaded federated deep learning based framework for detecting wormhole attacks in IoT networks. *Computers & Security*, 125, 103014 (2023).

33 Zhukabayeva, T., et al. A traffic analysis and node categorization-aware machine learning-integrated framework for cybersecurity intrusion detection and prevention of WSNs in smart grids. *IEEE Access*, 2024.

34 Adamova, A., Zhukabayeva, T., Mardenov, Y. Machine learning in action: An analysis of its application for fault detection in wireless sensor networks. *2023 IEEE International Conference on Smart Information Systems and Technologies (SIST)* (IEEE, 2023), pp. 506–511.

35 Mardenov, Y., Adamova, A., Zhukabayeva, T., Othman, M. Enhancing fault detection in wireless sensor networks through support vector machines: A comprehensive study. *Journal of Robotics and Control (JRC)*, 4 (6), 868–877 (2023).

36 Zhukabayeva, T., Adamova, A., Karabayev, N., Mardenov, Y., Satybaldina, D. Comprehensive vulnerability analysis and penetration testing approaches in smart city ecosystems. *2024 8th International Symposium on Innovative Approaches in Smart Technologies (ISAS)* (IEEE, 2024), pp. 1–6.

^{1,2*}Жукабаева Т.К.,

PhD, профессор, ORCID ID: 0000-0001-6345-5211,

*e-mail: zhukabayeva tk@enu.kz

^{1,3}Марденов Е.М.,

магистр, ORCID ID: 0000-0001-9284-9797,

e-mail: emardenov@gmail.com

²Танирбергенев А.Ж.,

т.ф.к., доцент м.а., ORCID ID: 0009-0000-8401-5434,

e-mail: t.adilbek@mail.ru

¹«Астана» халықаралық ғылыми кешені, Астана қ., Қазақстан

²Л.Н. Гумилев атындағы Еуразия ұлттық университеті, Астана қ., Қазақстан

³Astana International University, Астана қ., Қазақстан

СЫМСЫЗ СЕНСОРЛЫҚ ЖЕЛІЛЕРДЕГІ ҚАУІПСІЗДІК ИНЦИДЕНТТЕРІН АНЫҚТАУ ЖӘНЕ ТАЛДАУҒА АРНАЛҒАН КЕШЕНДІ ӘДІСТЕМЕ

Андатпа

Мақалада сымсыз сенсорлық желілердің (ССЖ) киберфизикалық қауіпсіздігі мәселелері қарастырылады. Қазіргі заманғы ССЖ sinkhole және wormhole шабуылдары, «ортадағы адам» типіндегі шабуылдар, деректерді алмастыру, әмбебап желілік шабуылдар және т.б. сияқты кең ауқымды қауіптерге осал. Тәжі-

рибеде мұндай шабуылдардан қорғануға ықпал ететін факторлардың әртүрлілігі, инфрақұрылымның тар бағытталған ерекшелігі және желі түйіндерінің шектеулі ресурстары кедергі келтіреді. Бұл мақалада ССЖ-дегі қауіпсіздік инциденттерін анықтауға арналған кешенді әдістеме ұсынылады. Ол шабуылдарды тиімді анықтауға және инциденттерге уақытылы әрекет етуге мүмкіндік береді, бұл өз кезегінде ықтимал зиянды азайтып, желінің үздіксіз жұмысын қамтамасыз етеді. Ұсынылған әдістеменің жаңалығы оның кешенділігінде, әртүрлі киберфизикалық қауіп-қатерлерді анықтау қабілетінде және ССЖ түйіндерінің таралған құрылымы мен құрамындағы өзгерістер динамикасын ескере отырып, инциденттерді жоғары дәлдікпен және толықтықпен айқындауында жатыр. Әдістеме ZigBee протоколы негізінде жұмыс істейтін және өндірістік нысанның немесе қаланың атмосфералық ауасының сипаттамаларын бақылауға арналған ССЖ фрагментінде сынақтан өткізілді. Дамытылған әдістеме сымсыз сенсорлық желілердегі қауіпсіздік инциденттерін анықтау сапасы мен жеделдігін арттырып, желілердің ішкі және сыртқы зиянды әсерлерге төзімділігін жоғарылатады және шабуыл сәтті болған жағдайда инфрақұрылым жұмысының ұзақ уақытқа тоқтап қалуын болдырмайды.

Тірек сөздер: сымсыз сенсорлық желі, қауіпсіздік инциденті, анықтау, модельдеу, әдістеме.

^{1,2}*Zhukabayeva T.K.,

PhD, Professor, ORCID ID: 0000-0001-6345-5211,

*e-mail: zhukabayeva tk@enu.kz

^{1,3}Mardenov E.M.,

MSc, ORCID ID: 0000-0001-9284-9797,

e-mail: emardenov@gmail.com

²Tanirbergenov A.,

PhD, acting Associate Professor, ORCID ID: 0009-0000-8401-5434

e-mail: t.adilbek@mail.ru

¹International Scientific Complex «Astana», Astana, Kazakhstan

²L.N. Gumilyov Eurasian National University, Astana, Kazakhstan

³Astana International University, Astana, Kazakhstan

COMPLEX TECHNIQUE FOR DETECTION AND ANALYSIS OF SECURITY INCIDENTS IN WIRELESS SENSOR NETWORKS

Abstract

The article comprises issues of cyber-physical security of wireless sensor networks (WSN). Modern WSNs are vulnerable to a wide class of attacks, such as sinkhole and wormhole attacks, man-in-the-middle attacks, data substitution attacks, universal network attacks, etc. In practice, the ability to protect WSNs from this type of attacks is hampered by the variety of possible impacts, highly specialized focus of the infrastructure and limited resources of network nodes. This article proposes a comprehensive technique for identifying security incidents in WSNs for effective attack detection and incident response, which will minimize potential damage and ensure uninterrupted network operation. The novelty of the technique includes its complexity, the ability to identify various cyber-physical threats and ensure high accuracy and completeness of incident detection, taking into account the distributed structure and dynamics of changes in the composition of WSN nodes. The technique has been tested on a WSN fragment operating on the ZigBee protocol to monitor the characteristics of the atmospheric air of an industrial facility or a city. The developed technique will help improve the quality and timeliness of detecting security incidents in wireless sensor networks, which will enhance the resilience of networks to external and internal malicious influences and prevent long-term interruptions in the operation of the infrastructure in the event of successful attacks.

Keywords: wireless sensor network, security incident, detection, modeling, technique.

Дата поступления статьи в редакцию: 04.05.2025

UDC 004.896
IRSTI 28.23.27

<https://doi.org/10.55452/1998-6688-2025-22-4-60-78>

¹**Samigulina Z.I.,**

PhD, ORCID ID: 0000-0002-5862-6415, e-mail:

z.samigulina@kbtu.kz

^{1*}**Shegentay I.T.,**

Master's student, ORCID ID: 0009-0008-8595-7069,

e-mail: i.shegentay@gmail.com

¹Kazakh-British Technical University, Almaty, Kazakhstan

DEVELOPMENT OF AUTOMATION AND CONTROL SYSTEM OF THE TECHNOLOGICAL PROCESS FOR SULFUR PRODUCTION

Abstract

Given the sophisticated technologies that modern industrial organizations are equipped with, monitoring and diagnostics of equipment condition are critical tasks. The current study aims to develop an improved diagnostic system for industrial equipment in the oil and gas industry using Schneider Electric M241 and M340 programmable logic controllers (PLCs). The first step in this process is to analyze the faults that occur during equipment operation, as well as to study the signal processing methods used in the oil and gas industry. The second step is to use PLCs for automated data collection, parameter monitoring and diagnostics of equipment condition. This approach allows for real-time control of key technological processes, reducing the probability of failures and increasing the reliability of production equipment. The study examined the impact of various data processing strategies on the efficiency of industrial equipment diagnostics. PLC data collection and analysis methods were considered, including continuous parameter monitoring, threshold control and trigger events. Based on these methods, diagnostic algorithms were developed and implemented in the EcoStruxure Machine Expert and EcoStruxure Control Expert, which provide automatic fault detection and alarm generation.

Key words: Industrial equipment diagnostics, programmable logic controller, sulfur production, PID control, Modicon M340, Modicon M241.

Introduction

Sulfur and its derivatives takes a huge applicational meaning in a wide range of industrial systems. In the realm of chemical production, approximately half of the global sulfur output is dedicated to the synthesis of sulfuric acid, while another quarter contributes to the creation of sulfites. Also, sulfur serves as a crucial component in the manufacturing of insecticides, accounting for up to 15% in the production of pesticides for grapes, cotton, and other crops. Also, sulfur is indispensable in the production of polymers and synthetic fibers, as well as in the formulation of pyrotechnical and explosive mixtures such as gunpowder and match head compositions. The paper industry also relies on sulfur for processing and subsequent production. Those factors make sulfur extremely useful material for humanity.

The production of certain steels with specialized properties is impossible without the involvement of sulfur. Additionally, sulfur plays a vital role in disinfecting agricultural facilities like vegetable storage warehouses, poultry houses, and basements. In the process of wine-making, the preservation of vegetables and fruits necessitates the involvement of sulfur. Sulfur also finds application in the creation of modified asphalt pavements and diverse types of concrete possessing specific properties, as well as in organic synthesis and the production of pharmaceuticals designed for addressing skin conditions.

Materials and methods

Claus process

The Claus method is still remains as a fundamental technique in gas desulfurization, primarily which is used for extracting elemental sulfur from hydrogen sulfide-rich gas streams. A method was introduced in the late 19th century by Carl Friedrich Claus and then the process has since evolved into a multi-stage industrial standard. Its widespread adoption is largely due to its ability to process waste gases from refining and natural gas treatment plants, converting them into usable sulfur through a combination of thermal and catalytic reactions. A hydrogen sulfide, which is produced, for example, during of the process of hydrodesulfurization in refinery naphthas and some other crude oil products, is being transformed into sulfur in a Claus reactors [1].

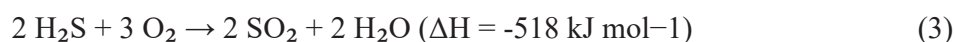
The Claus process involves two main stages:



Elemental sulfur serves as a valuable component in both fertilizer and pesticides. The Claus process can be divided into two stages: thermal and catalytic.

Thermal stage

In thermal stage of the Claus process, the hydrogen sulfide-rich gas is partially combusted under the substoichiometric conditions at the high temperatures which are above 850°C. This controlled reaction initiates the conversion of H₂S into sulfur dioxide (SO₂), releasing significant heat in the process. As the hot gas mixture cools, elemental sulfur begins to condense and is subsequently extracted in a gas cooling unit. This step not only drives the initial transformation of sulfur compounds but also ensures thermal decomposition of unwanted byproducts such as ammonia and hydrocarbons, preventing catalyst fouling in the subsequent stage. Claus gases are acidic gases and with no additional combustible substances aside from the hydrogen sulfide, are being ignited in a lances arranged around a central muffle through the whole following chemical reaction:



This is a highly exothermic process of complete combustion of hydrogen sulfide, resulting in the formation of sulfur dioxide, which undergoes further reactions. Of particular significance is the Claus reaction:



The overall equation looks like this:



The temperature in the Claus furnace exceeds 1050°C. This guarantees the destruction of BTEX compounds like benzene, toluene, ethylbenzene, and xylene, which would otherwise go to the downstream of Claus catalyst. The gases that contains ammonia, like those from the sour water stripper at the refinery, or hydrocarbons, are transformed inside of the burner muffler. A sufficient amount of air is introduced into the muffler to ensure the complete combustion of all the hydrocarbons and ammonia present. The air-to-acid gas ratio is carefully regulated so that one-third of the total hydrogen sulfide (H₂S) present is converted into SO₂. Usually, from 60 to 70% of the total amount of all elemental sulfur produced in the process is obtained in the thermal process stage. The bulk of the

heated gas emerging from the combustion chamber traverses the conduit of the process gas chiller, undergoing a cooling process that allows the sulfur generated during the reaction to condense. The sulfur forms in the thermal phase as highly reactive S_2 diradicals which combine exclusively to the S_8 allotrope:

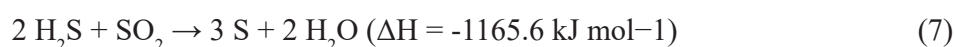


The thermal energy expelled by the process gas, coupled with the latent heat associated with condensation, is harnessed to generate medium- to low-pressure steam. The congealed sulfur is extracted at the liquid discharge segment of the process gas chiller.

Catalytic stage

Following the thermal step, the Claus process enters its catalytic phase, where the remaining hydrogen sulfide reacts with sulfur dioxide in the presence of a solid catalyst – commonly based on activated alumina or titanium oxide. This stage significantly increases the overall sulfur recovery rate by facilitating the main Claus reaction under milder temperature conditions, typically ranging from 200 °C to 330 °C. To prevent sulfur condensation and maintain catalyst activity, the process gas is reheated before entering each catalytic reactor. As the reaction progresses through multiple beds, additional sulfur is formed and then condensed downstream. Effective temperature control and catalyst management at this stage are essential for sustaining efficiency and minimizing byproduct formation.

During the Claus process, a significant amount of hydrogen sulfide (H_2S) reacts with sulfur dioxide (SO_2), which is produced during combustion, in the reaction chamber. This reaction leads to the formation of gaseous elemental sulfur:



The tail gas from the Claus process still containing combustible components and sulfur compounds (H_2S , H_2 and CO) is either burned in an incineration unit or further desulfurized in a downstream tail gas treatment unit.

Process Block Diagram

The first step of the process includes the preliminary treatment of the feed gas that generally consists of hydrogen sulfide (H_2S) as well as other constituents. The feed gas is then heated at a higher temperature inside a furnace, where the conversion of the hydrogen sulfide into sulfur dioxide (SO_2) occurs. It's referred to as thermal decomposition (Refer to Figure 1).

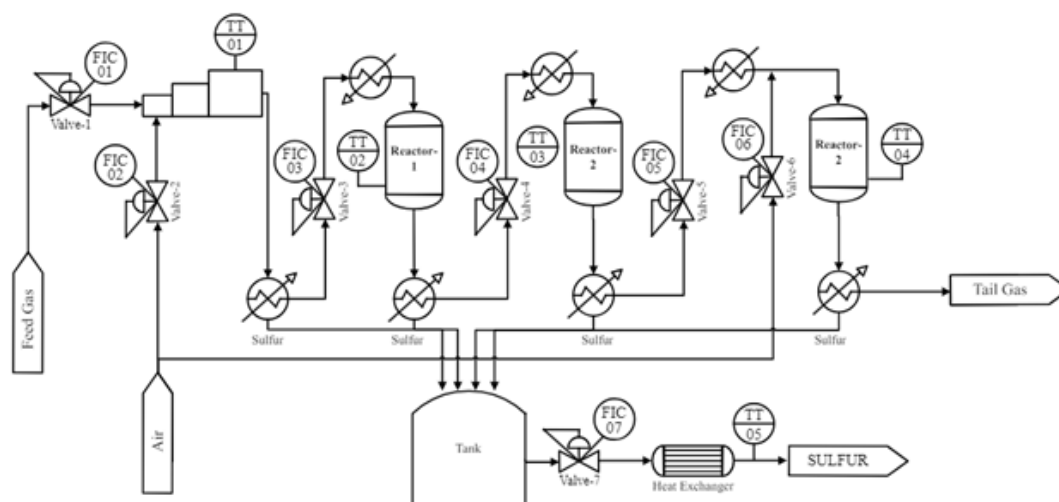


Figure 1 – Claus process, thermal stage

The remaining gas blend is then sent to a set of scrubber units such as the waste heat boiler and then the LP steam generating the sulfur condenser, where they remove any leftover sulfur constituents like carbonyl sulfide (COS) and the disulfide of carbon (CS₂). Sulfur gets condensed and extracted in the condenser of the sulfur.

Description of the general parameters

The Claus process is a tried and tested procedure utilized within the petroleum industry for the production of sulfur from sour gases. It has been used for many decades and its effectiveness relies on a number of crucial parameters such as temperature, pressure, residence time, and the feed gas composition. The raw material, which consists mainly of hydrogen sulfide, is fed into a steam boiler provided with gas tubes where it will go through a process of thermal conversion. In the combustion chamber of the boiler, temperatures of about 1100 °C are reached, progressive cooling to 350 °C as the gases make their way through the bundle of pipes. At this stage, water vapor at a pressure varying between 2.0 and 2.5 MPa is created.

Table 1 indicates general characteristics while Table 2 indicates general parameters.

Table 1 – Characteristics of purified VD/SD/ND gas and acid gas

Parameter	Purified VD gas, mol %	Purified SD gas, mol %	Purified ND gas, mol %	Acid gas, mol %
Temperature, °C	≤ 50	≤ 62	58	≤ 50
Pressure, bars	≤ 65	≤ 26	4.5	≤ 1.1
H ₂ S content, ppm about	≤ 15	≤ 35	≤ 15	not standardized

The reaction vessel residence time is a key variable for the Claus process as a measure of the exposure of the gases to the reaction conditions. Short residence times may lead to a lack of complete recovery of the sulfur present, while longer residence times may cause equipment fouling.

Table 2 – General Parameters on Installations

Parameter	Unit of measurem-ent	Accura-cy %	Acceptable range of technologi-cal parameters, des.	Acceptable range of technolo-gical parameter-rs, real	Recommend-ed optimal value
The water level in the steam collector	%	2.52	0–100	40–70	60–65
Steam pressure	Bars	0.82	33	20–30	27–29.5
...	...	...	...	...	...
Catalyst temperature	°C	1	400	260–370	310–320

The ratio of the composition of the feed gas, including hydrogen sulfide (H₂S) and carbon dioxide (CO₂), is another significant factor of the process. The source of most of the sulfur added by the feed comes from H₂S and depends on its concentration and the level of impurities present. The process requires high recovery rates of the sulfur, and these are obtained by careful management of residence time and carefully optimized composition of the feed. The operators are able to improve the efficiency and sustainability of the Claus process by recognizing and controlling these factors. The gas from the top of the contactor is sent directly to the afterburning furnace where the remaining sulfur compounds are oxidized into SO₂. The level of SO₂ present within the flue gas is then detected using the flow analyser. The temperature of the furnace is kept at around 800°C with the assistance

of the control of the temperature indicator. The flue gas fuel comes from two sources: the separator where the gas comes from the two separators and the flare, as well as the separator where the recycle streams of the top products of the fuel gas contactor and the phlegm of discharge from the diethanization tower. [16].

Description of the Control Object:

A heat exchanger is a device used for the exchange of heat energy between two or more fluids without physical mixing. The functioning of a heat exchanger is regulated by a host of parameters like the temperature difference between the fluids, the surface area exposed for the exchange of heat, the flow of fluids, and the respective modes of movement of the fluids. Heat exchangers are utilized on a large scale in multiple industries such as petroleum refining, food industry, manufacturing of petrochemicals, power production, nuclear power production, and even spacecraft system designing. There exist various types of heat exchangers consisting of [17]:

- ♦ Shell-and-tube heat exchangers (Figure 2): Such heat exchangers are made of a number of tubes having one type of fluids flowing inside the tubes and another flowing externally over the tubes within a shell. Shell-and-tube heat exchangers are commonly used in petroleum engineering because they are very effective and have the ability to withstand very high pressures. late heat exchangers: These types of units use thin plates as the separating medium between fluids and are generally smaller and more efficient for some uses, although they might be less long-lived than shell-and-tube types.

- ♦ Plate heat exchangers: These types of units use thin plates as the separating medium between fluids and are generally smaller and more efficient for some uses, although they might be less long-lived than shell-and-tube types.

- ♦ Air-cooled heat exchangers: These use air rather than water or other liquids as the cooling agent and thus are ideal for places where water supply is scarce.

Heat exchangers have a key role to play within petroleum engineering as they are used for numerous processes including:

- ♦ Heating crude oil or cooling it for the purposes of processing.
- ♦ Cooling water from the production of oil to avoid damage to the equipment and ensure ideal working conditions.
- ♦ Harnessing industrial process waste heat to make electricity or steam.
- ♦ Air-cooled heat exchangers: These use air rather than water or other liquids as the cooling agent and thus are ideal for places where water supply is scarce.

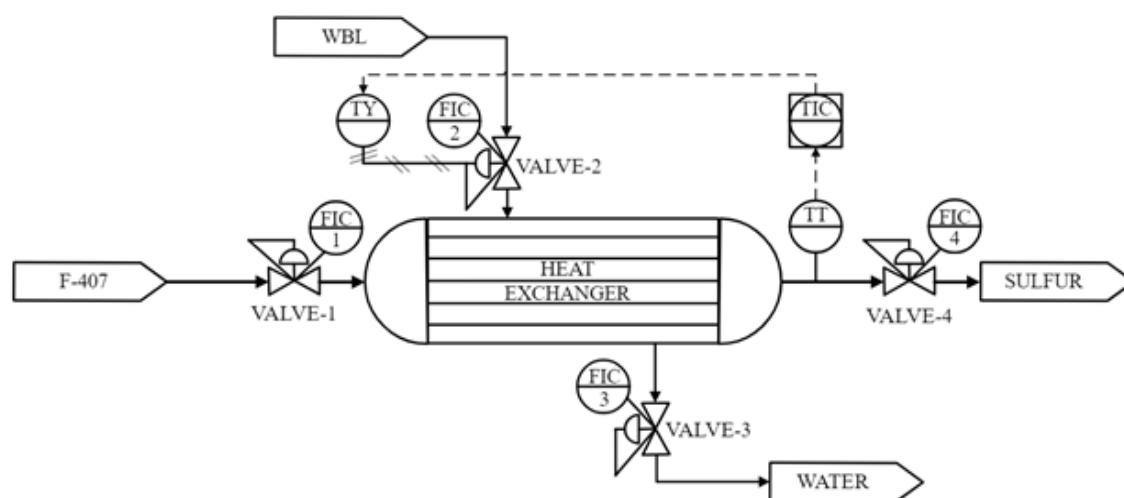


Figure 2 – Heat Exchanger's structural scheme

Specification and calculation of the general parameters of the mathematical model

By given parameters like the temperature of inlet and outlet sulfur it's possible to calculate the mathematical model for heat exchanger. Determining the thermal energy balance:

$$Q = G_s C_s (T_{s,in} - T_{s,out}) = G_w C_w (T_{w,in} - T_{w,out}) \quad (8)$$

The temperature of inlet sulfur is 220 C and the temperature of outlet is 130 C.

$$Q = G_s C_s (220 - 130) \quad (9)$$

Mass consumption of sulfur will be around 5 kg/s.

$$Q = 5 \times 0.73 \times (220 - 130) = 328.5 \quad (10)$$

The transfer function of the heat exchanger is described by an aperiodic link of the first order with a precise delay:

$$W(p) = \frac{K_t}{T_t p + 1} e^{-p\tau_t} \quad (11)$$

$$K_t = \frac{Q}{G_s C_s (T_{s,in} - T_{s,out})} \quad (12)$$

$$K_t = \frac{328.5}{5 \times 0.73 \times (220 - 130)} = 1 \quad (13)$$

After finding the transfer coefficient and time it's possible to find the transfer function of heat exchanger.

$$T_t = \frac{\rho_s V_s}{G_s} \quad (14)$$

$$T_t = \frac{1810 \times 0.5}{5} = 181 \quad (15)$$

A transfer function of a heat exchanger with precise delay:

$$W(p) = \frac{1}{181p + 1} e^{-p\tau_t} \quad (16)$$

For the convenience of analysis, we use the first-order Pade approximation.:

$$e^{-p\tau_t} = \frac{1 - \frac{p\tau_t}{2}}{1 + \frac{p\tau_t}{2}} \quad (17)$$

$$\tau_t = 10 \text{ s.}$$

$$W(p) = \frac{1}{181p + 1} \times \frac{1 - 5p}{1 + 5p} \quad (18)$$

$$W(p) = \frac{1 - 5p}{(181p + 1)(5p + 1)} \quad (19)$$

The final transfer function looks like:

$$W(p) = \frac{1 - 5p}{905p^2 + 186p + 1} \quad (20)$$

After determining the final transfer function it's possible to use it for further analysis. The transfer function is needed to mathematically describe a dynamic system [2]. Some purposes of using the transfer function. Determination of the dynamic properties of the system. The transfer function of a system determines all its dynamic properties, so the primary task of calculating a control system is to determine its transfer function. The transfer function describes the system's response to a change

in the input parameter, including the reaction delay and the rate of change in the output value [3]. Restoring the output signal. If the transfer function and the input signal of the system are known, then the output signal can be restored. In Figure 3 is shown a system of heat exchange with control valve and temperature indicator in transfer functions that was designed in a Matlab Software in Simulink tool. The system has 3 different PID regulators with different coefficients and creation of a regulator based on fuzzy logic principles was considered [4]. The need for a mathematical model of the selected control object lies in the ability to predict and evaluate the behavior of the process [5]. It is a differential operator that expresses the relationship between the output and input of a linear stationary system [6].

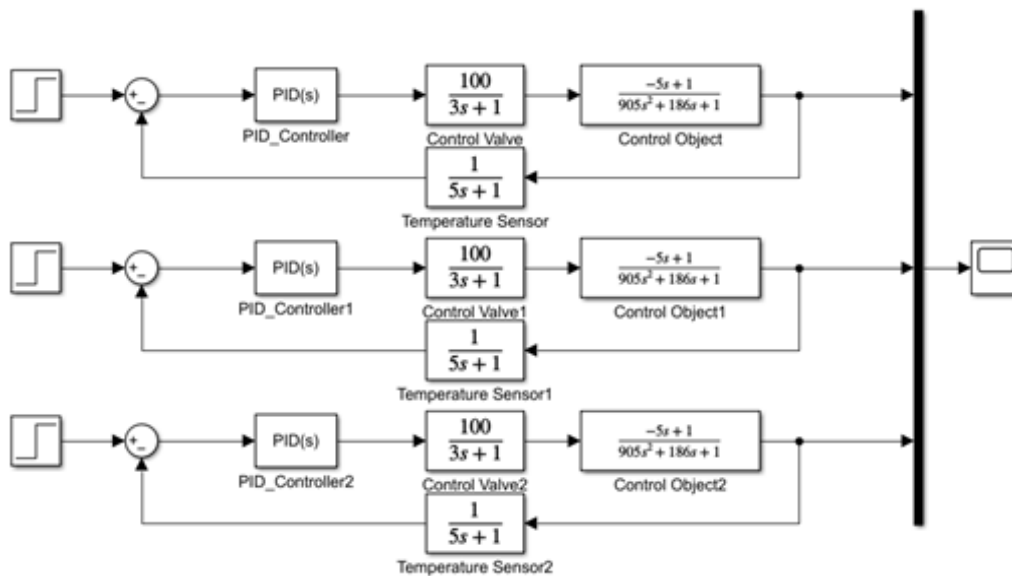


Figure 3 – Simulink heat exchange system with PID regulators

In Figure 4 a graph of the scope of each system with different PID coefficients is shown and Table 3 shows them.

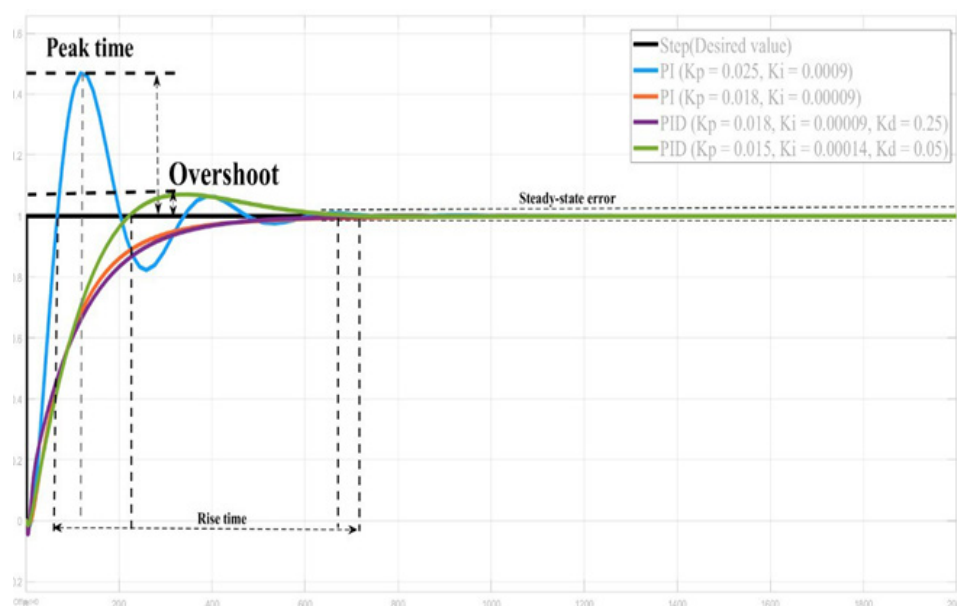


Figure 4 – Scope of each system with different PID coefficients

Table 3 shows the different PID coefficients for each system with the same transfer function.

Table 3 – Different PID coefficients

Color	Kp	Ki	Kd
Yellow	0.0187	0.00009	0.13
Blue	0.0150	0.00014	0.12
Red	0.1000	0.00015	0.15

Development of automation and control system on the basis of PLC

Programmable logic controllers (PLCs) are the basis of industrial automation, as it is a hardware and software component that directly interacts with sensors, actuators, and other field devices. A PLC provides access to control and monitoring of many machines connected to it and the software system integrated with them. PLCs can effectively work with various communication protocols, analyze data, perform conversions, interact with I/O modules, process signals, etc. PLCs are fully programmable by the user using the development environment and standard IEC 61131-3 programming languages, such as ST, FBD, SFC, LD, IL. EcoStruxure Control Expert is a software platform for developing, configuring and maintaining control systems in various industries. Its necessity is due to the ability to simplify and improve control processes, increase operational efficiency and ensure smooth integration of various automation components. EcoStruxure Machine Expert is a software from Schneider Electric that is used for the design, programming and maintenance of machines and equipment in industry. Configuration software EcoStruxure Operator Terminal Exper screen designed to create and save applications for automating control systems for Mageli GTU (Premium and Open units)/ STO Color/GTO terminals and Magelis industrial PCs (elpan and unit) [7]. In Figure 5 is shown «Industrial automation Lab» (Schneider Electric).



Figure 5 – Sulfur production control system realization in «Industrial automation Lab» (Schneider Electric)

«Industrial automation Lab» (Schneider Electric) was used for achieving the goal of a construction of a system for sulfur production. In Figure 6 a connection of controllers of 2 Masters M340 with 3 Slaves M241 each connected by TCP Modbus protocol and by one controller connected on HMI Harmony Panel with Ethernet network. If one of the connections with the controller broke the indicator will show exactly what controller connection were broken is shown from the laboratory.

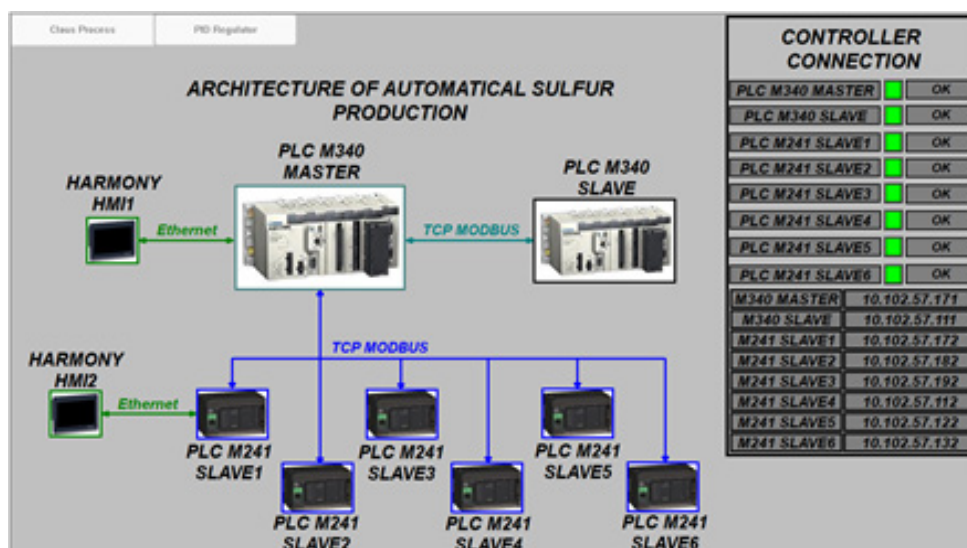


Figure 6 – Architecture of controller connection

Figure 7 represents us a control plant of sulfur production. Having reaction furnace which is RF – 100, 4 condensers C-XXX, 3 reheaters, 1 huge tank – T-100 and heat exchanger HE-100.

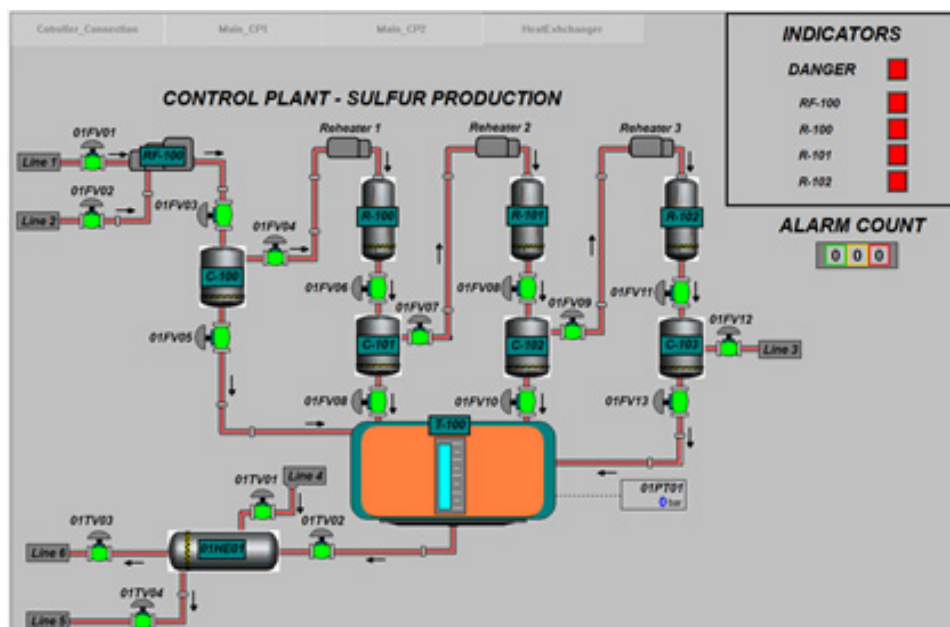


Figure 7 – Development of the operator screen for sulfur production in EcoStructure Control Expert

In Figure 8 all the equipment of a Stand from the laboratory is shown that it were used for the construction of the system.

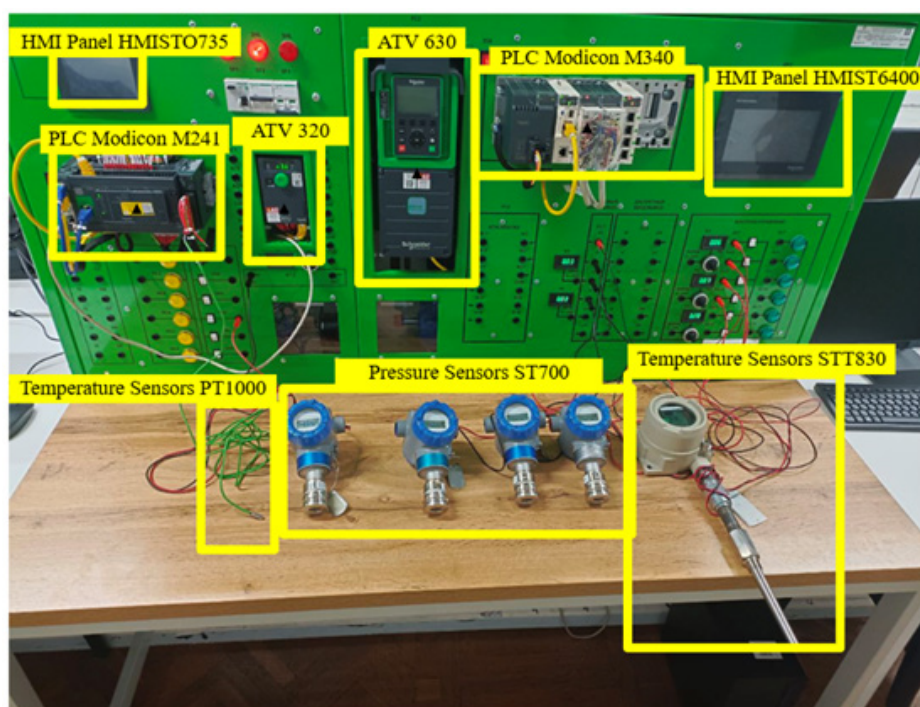


Figure 8 – Realization on stand

Figure 9 represents the view of an HMI panel for monitoring the control object.

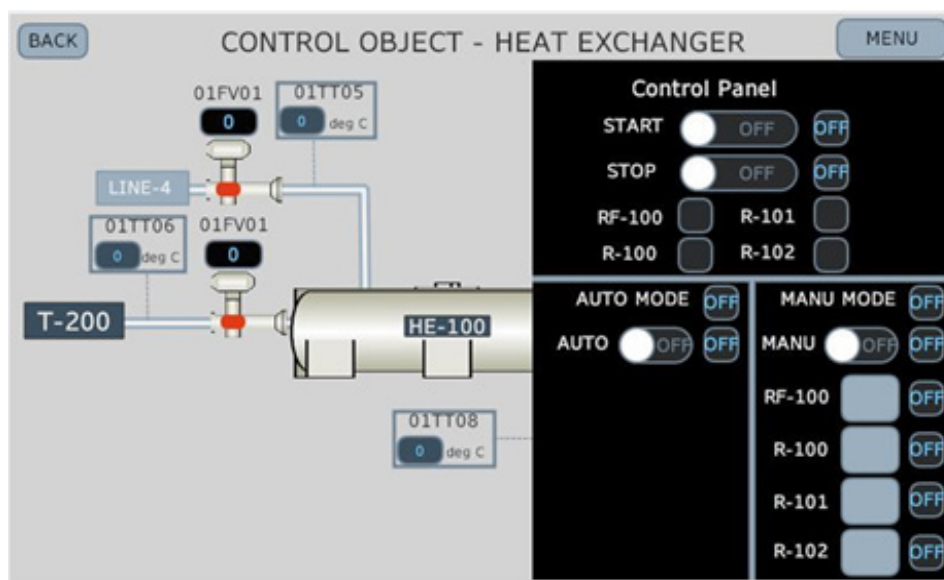


Figure 9 – Development HMI display on the basis of HMIST6400 Panel with EcoStructure Operator Terminal Expert

Figure 10 shows software for Panels EcoStruxure Operator Terminal Expert.

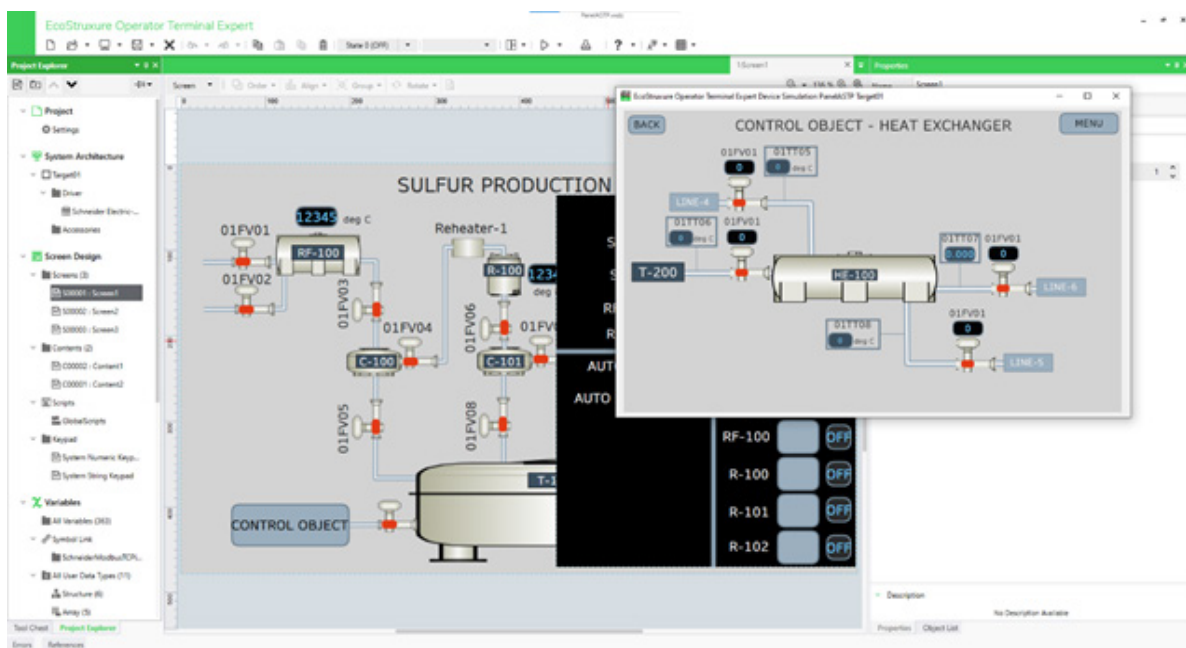


Figure 10 – EcoStruxure Operator Terminal Expert

The PLC connections are based on a sending pulse signal from Slave to Master controller for indication of controller connection. In Figure 11 an HMI panel with an opened PID graph which is realized with program on M340 controller by using the special software EcoStruxure Operator Terminal Expert.



Figure 11 – PID graph on HMIST6400

In Figure 12 an HMI panel with an opened PID graph which is realized with program on M241 controller by using the special software EcoStruxure Operator Terminal Expert.

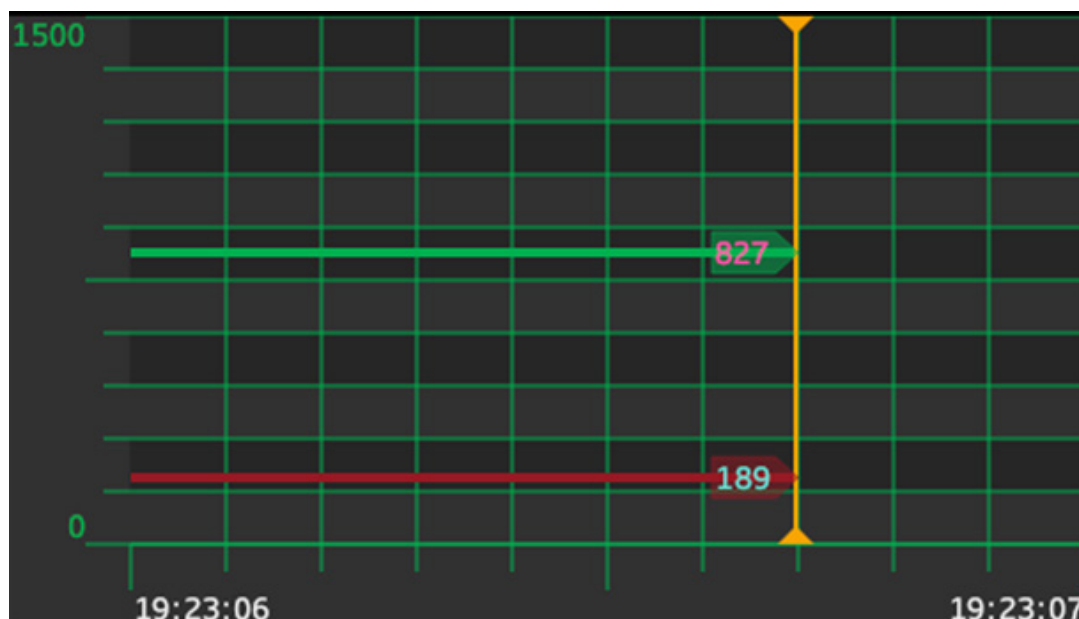


Figure 12 – PID graph on HMISTO735

By using the WebVisu from EcoStruxure Machine Expert it's possible to share the visualization that was realized in a Software to other devices that connected to the same network as PLC by using it's address with WebVisu extension. In Figure 13 is shown the exact phone connection to PLC for Emergency Shut down in case if operator would have troubles with control point/room.



Figure 13 – Mobile PLC Modicon M241 control with Webvisu and EcoStruxture Machine Expert

In Figure 14 is shown another software EcoStruxure Machine Expert with visualization and it's WebVisu configuration for generating a website based on controller's IP address.

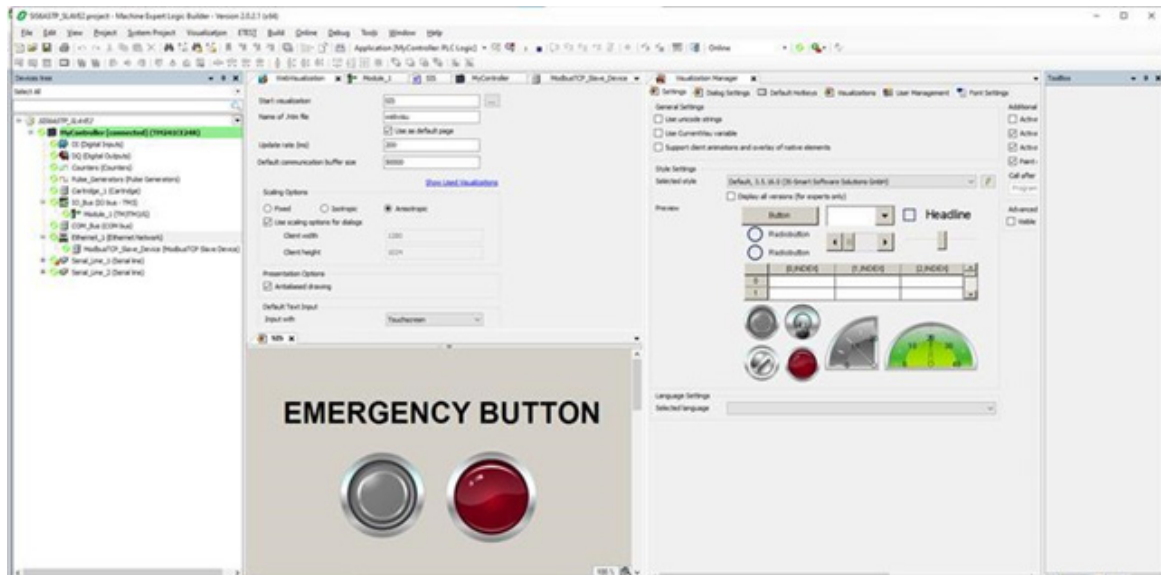


Figure 14 – EcoStruxure Machine Expert

Figure 15 represents HMI panel with opening menu and starting the Manu mode of the Automated Sulfur production system by using the special software EcoStruxure Operator Terminal Expert.

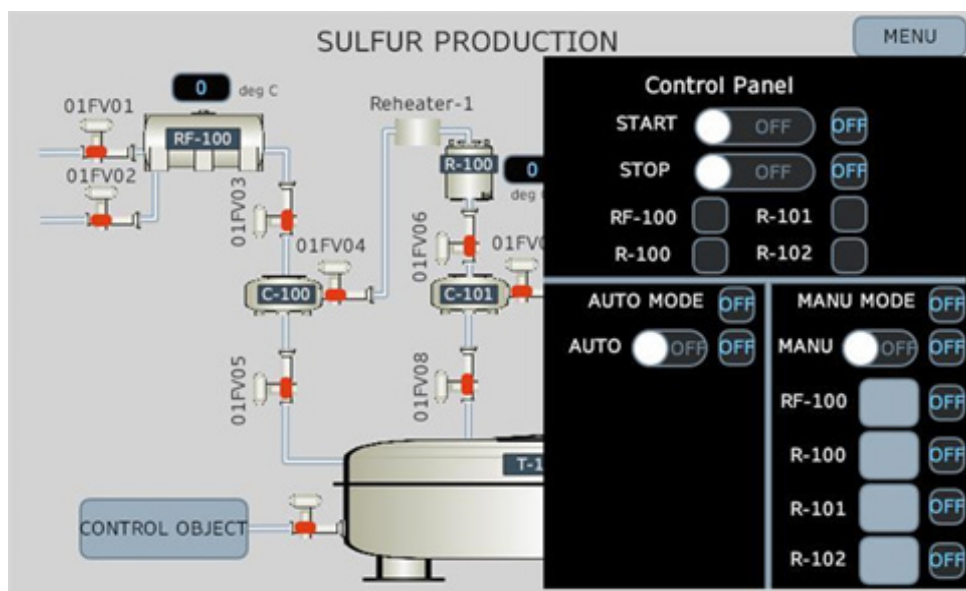


Figure 15 – Menu and Manu mode from HMI Panel

As the representation of control objects were chosen ATV630 (Figure 16a) as Reaction Furnace and ATV320 (Figure 16b) for valve opening control with 3 PID controllers – 1 for M340 and 2 for M241. Summarizing this part of the project, the strengths of the program realization is the

implementation of overall 8 PLC controllers with 17 sensors including 10 pressure sensors and 7 temperature sensors. The system has an HMI control with Auto and Manu modes and control objects as frequency converters like ATV320 and ATV630. System also has an emergency shut down by phone manipulation for remote control of the safety in case of trouble in operator's room or system.



a) ATV630 frequency converter



b) ATV320 frequency converter

Figure 16 – Description of the frequency converter in the system

The PLC connection is based on a TCP Modbus connection with Ethernet connection to HMI Panels. On stand realized the indicators based on a simple lamp that uses the Discrete output from the modules of PLCs

AI IMPLEMENTATION IN PLC SYSTEM

By using a software Modbus Poll it's possible to collect the data from sensors using Modbus addresses. In Figure 17 is shown an Excel table that was collected from Pressure sensors ST700 and Temperature sensor PT 1000.

The temperature at the inlet and outlet is an important parameter for diagnosis, as temperature changes may indicate a decrease in heat exchange efficiency. Inlet and outlet pressure – pressure drop or sudden fluctuations may also indicate clogging or other problems in the system.

This data can be used to create a model that will track anomalies and warn of possible accidents. For example, to predict clogging, you can build a model that will take into account changes in temperature and pressure as signs of deterioration of the heat exchanger.

Availability of accident data on previous accidents can be very useful for training an AI model. Typically, in industrial systems such as heat exchangers, data on accidents and failures is stored in an event log or in a maintenance database (CMMS - Maintenance management system). This data can be useful for training a machine learning model that will look for patterns in the behavior of the heat exchanger before an accident.

As a little experiment a simple code was written in a Google Colaboratory using a Python programming language to analyze and detect anomaly from collected data from sensors.

In Figure 18 is shown an Excel file that was gained from the result of Google Colaboratory. It used the Isolation Forest algorithm to detect the anomaly score from sensors. Isolation Forest is an

algorithm for detecting anomalies. He does not build a density model but tries to isolate outliers. A tree is built where each node randomly selects one feature and a random split value. The data is divided recursively until each point is in its own separate sheet. A point is considered an anomaly based on how quickly it was isolated (the shorter the path, the more anomalous). If an object separates in just a couple of steps, it means that it is very out of the picture

	A	B	C	D	E	F
4	2025-04-17 11:54:14.976	2,5	0,4	0,9	24,5	1,2
5	2025-04-17 11:54:15.093	2,5	0,4	0,9	24,5	1,2
6	2025-04-17 11:54:16.865	2,5	0,4	0,9	24,5	1,2
7	2025-04-17 11:54:17.861	2,5	0,4	0,9	24,5	1,2
8	2025-04-17 11:54:18.898	2,5	0,4	0,9	24,5	1,2
9	2025-04-17 11:54:19.895	2,5	0,4	0,9	24,8	1,2
10	2025-04-17 11:54:20.892	3,6	0,4	0,9	25	1,2
11	2025-04-17 11:54:21.904	5,5	0,4	0,9	25,2	1,2
12	2025-04-17 11:54:22.918	6,4	0,5	0,9	25,4	1,2
13	2025-04-17 11:54:23.959	6,8	1	0,9	25,7	1,2
14	2025-04-17 11:54:24.955	7,3	1,4	0,9	25,9	1,2
15	2025-04-17 11:54:25.960	7,5	1,5	0,9	26,1	1,2
16	2025-04-17 11:54:26.988	6,5	1,1	0,9	26,2	1,2
17	2025-04-17 11:54:28.003	5,7	0,6	0,9	26,4	1,2
18	2025-04-17 11:54:29.008	5,8	0,5	2,6	26,5	1,2
19	2025-04-17 11:54:30.023	6,3	0,4	5	26,6	1,2
20	2025-04-17 11:54:31.028	6,6	0,4	6,1	26,7	1,2
21	2025-04-17 11:54:32.054	6,6	0,4	6,5	26,7	1,2
22	2025-04-17 11:54:33.071	6,4	0,4	5,8	26,8	1,2
23	2025-04-17 11:54:34.081	4	0,4	2,7	26,8	1,2
24	2025-04-17 11:54:35.086	2,9	0,5	1,4	26,8	1,2
25	2025-04-17 11:54:36.095	2,6	1	1,1	26,9	4,2
26	2025-04-17 11:54:37.124	2,5	1,3	1	26,9	6,8
27	2025-04-17 11:54:38.136	2,5	1,5	0,9	26,9	7,4
28	2025-04-17 11:54:39.162	2,5	1,5	0,9	26,9	7,5
29	2025-04-17 11:54:40.175	2,5	1,4	0,9	26,9	5,9

Figure 17 – Collected data

Type log name here	Unnamed: 1	Unnamed: 2	Unnamed: 3	Unnamed: 4	Unnamed: 5	Anomaly_Score
Poll definition: ID = 1, Function = 03, Address = 0, ScanRate = 1000						
	Pres_Sen_1	Pres_Sen_2	Pres_Sen_3	Temp_Sen_1	Pres_Sen_4	
2025-04-17 11:54:14.976	2,5	0,400000006	0,899999976	24,5	1,200000048	0,304535275
2025-04-17 11:54:15.093	2,5	0,400000006	0,899999976	24,5	1,200000048	0,304535275
2025-04-17 11:54:16.865	2,5	0,400000006	0,899999976	24,5	1,200000048	0,304535275
2025-04-17 11:54:17.861	2,5	0,400000006	0,899999976	24,5	1,200000048	0,304535275
2025-04-17 11:54:18.898	2,5	0,400000006	0,899999976	24,5	1,200000048	0,304535275
2025-04-17 11:54:19.895	2,5	0,400000006	0,899999976	24,799999924	1,200000048	0,304535275
2025-04-17 11:54:20.892	3,599999905	0,400000006	0,899999976	25	1,200000048	0,304535275
2025-04-17 11:54:21.904	5,5	0,400000006	0,899999976	25,200000076	1,200000048	0,304535275
2025-04-17 11:54:22.918	6,400000095	0,5	0,899999976	25,399999962	1,200000048	0,304535275
2025-04-17 11:54:23.959	6,800000191	1	0,899999976	25,700000076	1,200000048	0,304535275
2025-04-17 11:54:24.955	7,300000191	1,399999976	0,899999976	25,899999962	1,200000048	0,304535275
2025-04-17 11:54:25.960	7,5	1,5	0,899999976	26,100000038	1,200000048	0,304535275
2025-04-17 11:54:26.988	6,5	1,100000024	0,899999976	26,200000076	1,200000048	0,304535275
2025-04-17 11:54:28.003	5,699999809	0,600000024	0,899999976	26,399999962	1,200000048	0,304535275
2025-04-17 11:54:29.008	5,800000191	0,5	2,599999905	26,5	1,200000048	0,304535275
2025-04-17 11:54:30.023	6,300000191	0,400000006	5	26,600000038	1,200000048	0,304535275
2025-04-17 11:54:31.028	6,599999905	0,400000006	6,099999905	26,700000076	1,200000048	0,304535275
2025-04-17 11:54:32.054	6,599999905	0,400000006	6,5	26,700000076	1,200000048	0,304535275
2025-04-17 11:54:33.071	6,400000095	0,400000006	5,800000191	26,799999924	1,200000048	0,304535275
2025-04-17 11:54:34.081	4	0,400000006	2,700000048	26,799999924	1,200000048	0,304535275
2025-04-17 11:54:35.086	2,900000095	0,5	1,399999976	26,799999924	1,200000048	0,304535275
2025-04-17 11:54:36.095	2,599999905	1	1,100000024	26,899999962	4,199999809	0,304535275
2025-04-17 11:54:37.124	2,5	1,299999952	1	26,899999962	6,800000191	0,304535275
2025-04-17 11:54:38.136	2,5	1,5	0,899999976	26,899999962	7,400000095	0,304535275

Figure 18 – Anomaly detection result

A code for building the trees is shown in Figure 19. The total number of trees that were built was 50. By using those trees code detects points and then calculates the anomaly score for the analysis.

```
[20] class InternalNode:
    def __init__(self, split_feature, split_value, left_subtree, right_subtree):
        self.split_feature = split_feature
        self.split_value = split_value
        self.left = left_subtree
        self.right = right_subtree

    def IsolationTree(data, current_depth, max_depth):
        if current_depth >= max_depth or len(data) <= 1:
            return ExternalNode(size=len(data))

        split_feature = random.choice(data.columns)
        values = data[split_feature].values
        split_value = random.uniform(min(values), max(values))

        left_data = data[data[split_feature] < split_value]
        right_data = data[data[split_feature] >= split_value]

        left_subtree = IsolationTree(left_data, current_depth + 1, max_depth)
        right_subtree = IsolationTree(right_data, current_depth + 1, max_depth)

        return InternalNode(split_feature, split_value, left_subtree, right_subtree)

    def IsolationForest(data, num_trees=100, max_depth=10):
        forest = []
        for _ in range(num_trees):
            sample_data = data.sample(frac=1, replace=True).reset_index(drop=True)
            tree = IsolationTree(sample_data, 0, max_depth)
            forest.append(tree)
        return forest

numeric_data = data.select_dtypes(include=[np.number])
numeric_data = numeric_data.fillna(numeric_data.mean())
forest = IsolationForest(numeric_data, num_trees=50, max_depth=8)
print("Num of trees: (len(forest))")
```

Num of trees: 50

Figure 19 – Isolation Tree code

By using matplotlib.pyplot library it was possible to construct the plot to visually see the anomaly detection of collected data (Figure 20). The closer coefficient gets to 1 the more anomaly appears, and it means a very short path in the trees. As it's possible to see from a plot a collected data is closer to 0 which means that it's closer to a normal. The Isolation Forest algorithm has shown its effectiveness in spotting anomalies in a collected data set. By using the same principle it's possible to build a Model which can detect and predict the accidents and prevent them if it has access to database of sensors in a real time.

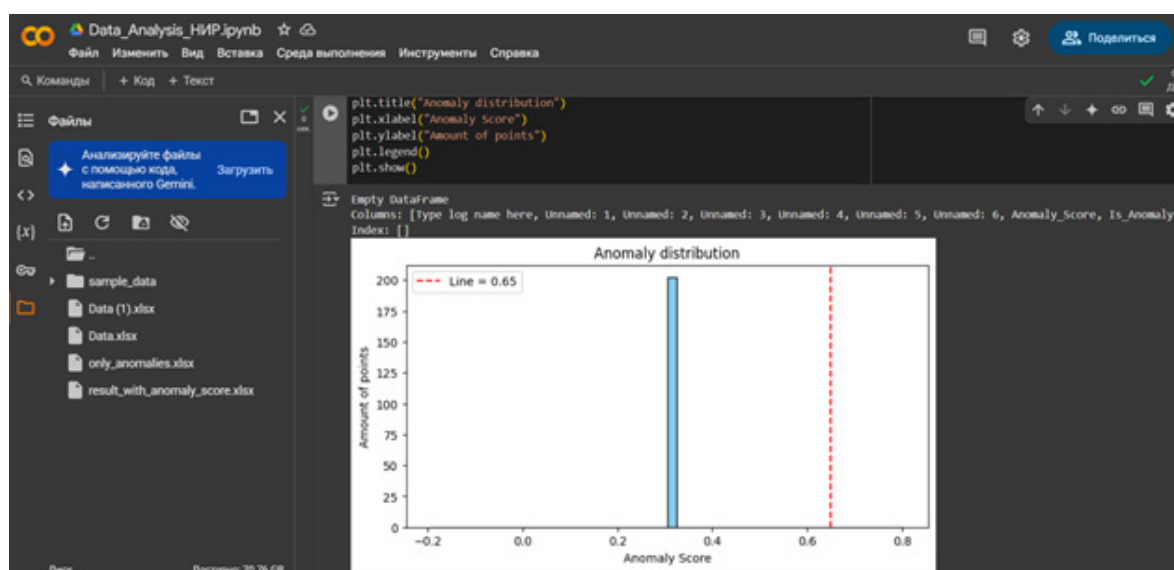


Figure 20 – Matplot of anomaly distribution

However, for safety reasons it would be better to limit access to control of plants from that model. By limiting model from full control of automation systems it's possible to prevent them from damaging the system. The last choice must be made by a professional and qualified person.

This small experiment showed the potential in combining machine learning and artificial intelligence with automation and control of production processes, as well as in other areas. If narrow down the created model to analyzing the data it receives in real time, it is possible to create an effective incident prevention system that will significantly reduce the costs of many companies.

Results and discussion

This section presents and discusses the main results obtained during the research. Figure 7 shows the buildup control system which is based on Programmable Logic Controllers (PLC) by using TCP Modbus protocols. In Figure 8 shown the scheme of process control logic of the technical process. Figure 14 demonstrates the implementation of remote monitoring and control using WebVisu from EcoStruxure Machine Expert. Figure 19 presents a summary table of data collected from sensors, and Figure 20 shows the results of anomaly distribution analysis performed using Python programming language.

The Modbus TCP based architecture demonstrated the ease of integration with industrial sensors and the ability to scale. The key performance indicators of the system include data exchange latency, polling frequency, and channel stability under peak loads. In the future, it is recommended to conduct systematic network measurement and analyze their impact on the performance of control algorithms for technological processes.

The analysis of data from sensors using machine learning methods made it possible to identify patterns and detect deviations from the normal behavior of the equipment. The obtained distribution of anomalies (as shown in Figure 21) indicates the presence of localized spikes on individual sensors and periodic deviations, which makes it possible for surgical intervention. For an objective assessment of the anomaly detector's effectiveness, accuracy and completeness metrics are highly required as well as testing on a dedicated deferred samples.

Conclusion

Modernization of automatic control systems for sulfur production is necessary for efficient production. The state of automation affects the efficiency and functionality of production systems. The issue of improving automated systems is acute. The article contains research on the analysis of an automated sulfur production system using Schneider Electric techniques.

The utilization of mathematical modelling, particularly through the MATLAB software, plays a pivotal role in understanding and optimizing the control system. The study involved examining the mathematical model of the heat exchange process.

During the study, a connection structure was created from several PLCs using Schneider Electric software and the network configuration and system architecture were implemented. During this process, a PID controller was designed, which allowed for better productivity. Automated systems at strategically important facilities and enterprises must comply with all applicable standards and function promptly.

REFERENCES

- 1 Rustamov, M.I., Gaisin, A.S., Mamedov, D.N. A modern reference book on petroleum fuels and their production technologies. Khimik Foundation. Baku, 2005. 640 p.
- 2 Wang, L., Visioli, A. Introduction to the special issue on PID control system design, automatic tuning and applications. International Journal of Control, 97 (1), 1–7 (2024).
- 3 Bavafa-Toosi, Yazdan. Introduction to Linear Control Systems, 2019.

- 4 Borase, R.P., Maghade, D.K., Sondkar, S.Y. et al. A review of PID control, tuning methods and applications. *Int. J. Dynam. Control.*, 9, 818–827 (2021).
- 5 Thuan, D.D., Son, N.H. Stability of stochastic differential-algebraic equations with delay. *International Journal of Systems Science*, 1–6 (2024).
- 6 Yazdan Bavafa-Toosi. Introduction to Linear Control Systems, 2019 4. Christophe P. Basso. Linear Circuit Transfer Functions, 2020.
- 7 EcoStruxure™ Control Expert PLC Simulator Original instructions. Schneider Electric, 2024.
- 8 Danilova, S.S. *Vestnik Nauki*, 8, 78–85 (2022).
- 9 Dyagelev, M.U. The review of aeration systems for biological wastewater treatment. *IOP Conf. Ser.: Earth Environ*, 1–7 (2021).
- 10 Miguel-Escrig, O., Romero-Pérez, J. Event-based discrete PI controllers robustness analysis through sampled describing function technique. *International Journal of Control*, 97 (1), 86–100 (2024).
- 11 Fromm, L. Development of a project for a unit for obtaining commercial sulfur from acid gases using the Claus method with a capacity of 20 thousand tons per year: diploma thesis (2022).
- 12 Rakhmatov, H., Nurboboev, N.T. Improving the efficiency of the Klaus unit for the production of elemental sulfur at the Mubarek gas processing plant. *Sanoatda Raqamli Texnologiyalar*, 2 (01) (2024). URL: <https://ojs.qmii.uz/index.php/sr/article/view/680>.
- 13 Gong, G., Sheng, L., Zhou, D. Robust fault-tolerant stabilisation of uncertain high-order fully actuated systems with actuator faults. *International Journal of Systems Science*, 1–8 (2024).
- 14 Web-source: <https://www.mathworks.com/matlabcentral/answers/279079-how-to-make-bodeplot-of-transfer-function>, 2022.

¹Самигулина З.И.,

PhD, ORCID ID: 0000-0002-5862-6415,

e-mail: z.samigulina@kbtu.kz

^{1*}Шегентай И.Т.,

магистрант, ORCID ID: 0009-0008-8595-7069,

*e-mail: i.shegentay@gmail.com

¹Қазақстан-Британ техникалық университеті, Алматы қ., Қазақстан

КҮКІРТ ӨНДІРУ ТЕХНОЛОГИЯЛЫҚ ПРОЦЕСІН АВТОМАТТАНДЫРУ ЖӘНЕ БАСҚАРУ ЖҮЙЕСІН ДАМУ

Андатпа

Заманауи өнеркәсіптік ұйымдарда күрделі технологиялардың кеңінен қолданылуын ескере отырып, жабдық жағдайын бақылау және диагностикалау маңызды міндеттердің бірі. Бұл зерттеу Schneider Electric M241 және M340 бағдарламаланатын логикалық контроллерлерін (PLC) пайдалану арқылы мұнай-газ өнеркәсібіндегі өндірістік жабдықтарды диагностикалаудың жетілдірілген жүйесін әзірлеуге бағытталған. Зерттеудің алғашқы қадамы – жабдықты пайдалану барысында туындайтын ықтимал ақауларды талдау, сондай-ақ мұнай-газ саласында қолданылатын сигналдарды өңдеу әдістерін зерттеу. Екінші қадам – автоматтандырылған деректерді жинау, параметрлерді бақылау және жабдық жағдайын диагностикалау үшін PLC қолдану. Ұсынылған тәсіл негізгі технологиялық процестерді нақты уақыт режимінде басқаруға, істен шығу ықтималдығын азайтуға және өндірістік жабдықтың сенімділігін арттыруға мүмкіндік береді. Зерттеу барысында әртүрлі деректерді өңдеу стратегияларының өнеркәсіптік жабдықты диагностикалау тиімділігіне ықпалы қарастырылды. PLC негізінде деректерді жинау және талдау әдістері, соның ішінде үздіксіз параметрлер мониторингі, шекті мәндерді бақылау және триггерлік оқиғаларды тіркеу зерттелді. Осы әдістердің негізінде EcoStruxure Machine Expert және Unity Pro бағдарламалық орталарында ақауларды автоматты түрде анықтау және дабыл қалыптастыруды қамтамасыз ететін диагностикалық алгоритмдер әзірленіп, іске асырылды.

Тірек сөздер: Schneider Electric, өнеркәсіптік жабдықтың диагностикасы, контроллерлер, күкірт өндірісі.

¹Самигулина З.И.,

PhD, ORCID ID: 0000-0002-5862-6415,

e-mail: z.samigulina@kbtu.kz

^{1*}Шегентай И.Т.,

магистрант, ORCID ID: 0009-0008-8595-7069,

e-mail: i.shegentay@gmail.com

¹Казахстанско-Британский технический университет, г. Алматы, Казахстан

РАЗРАБОТКА СИСТЕМЫ АВТОМАТИЗАЦИИ И УПРАВЛЕНИЯ ТЕХНОЛОГИЧЕСКИМ ПРОЦЕССОМ ПРОИЗВОДСТВА СЕРЫ

Аннотация

Учитывая сложную технологию, которой оснащены современные промышленные организации, мониторинг и диагностика состояния оборудования являются критически важными задачами. Целью настоящего исследования является разработка усовершенствованной системы диагностики промышленного оборудования в нефтегазовой отрасли с использованием программируемых логических контроллеров (ПЛК) Schneider Electric M241 и M340. Первым шагом в этом процессе является анализ неисправностей, возникающих в процессе эксплуатации оборудования, а также изучение методов обработки сигналов, используемых в нефтегазовой отрасли. Вторым шагом является использование ПЛК для автоматизированного сбора данных, мониторинга параметров и диагностики состояния оборудования. Такой подход позволяет в режиме реального времени контролировать ключевые технологические процессы, снижая вероятность отказов и повышая надежность производственного оборудования. В исследовании изучалось влияние различных стратегий обработки данных на эффективность диагностики промышленного оборудования. Были рассмотрены методы сбора и анализа данных ПЛК, включая непрерывный мониторинг параметров, пороговый контроль и триггерные события. На основе этих методов разработаны и внедрены в среды EcoStruxure Machine Expert и Unity Pro алгоритмы диагностики, обеспечивающие автоматическое обнаружение неисправностей и генерацию аварийных сигналов.

Ключевые слова: Schneider Electric, диагностика промышленного оборудования, контроллеры, производство серы.

Article submission date: 28.04.2025

UDC 004.056: 004.8
IRSTI 27.17.03

<https://doi.org/10.55452/1998-6688-2025-22-4-79-96>

¹Kikbayev N.E.,

Master's student, ORCID ID: 0009-0000-0145-5718,

e-mail: kikbaevnurbek@gmail.com

¹Zhexebay D.M.,

PhD, ORCID ID: 0009-0008-1884-4662,

e-mail: zhexebay92@gmail.com

²Xin Y.,

Professor, ORCID ID: 0000-0001-6169-0795,

e-mail: qyuxiao@purdue.edu

³Tynymbayev S.T.,

Professor, ORCID ID: 0000-0002-7808-5273,

e-mail: s.tynym@gmail.com

³Aitmagambetov A.Z.,

Professor, ORCID ID: 0000-0002-9326-9476,

e-mail: a.aitmagambetov@iitu.edu.kz

¹Abdizhalilova L.B.,

Master's student, ORCID ID: 0009-0000-5965-7195,

abdijalil.lazzat@bk.ru

^{1*}Skabylov A.A.

PhD, ORCID ID: 0000-0002-5196-8252,

*e-mail: Alisher.skabylov@kaznu.edu.kz

¹Al-Farabi Kazakh National University, Almaty, Kazakhstan

²Northwestern Polytechnical University, Xi'an, China

³International Information Technology University, Almaty, Kazakhstan

COMPARATIVE STUDY OF MACHINE LEARNING METHODS FOR DETECTING ANOMALIES IN NETWORK TRAFFIC

Abstract

The demand for intrusion detection systems (IDSs) that can promptly identify both known and new types of attacks is on the rise due to the rapid expansion of cyber threats and the consequent increase in network traffic. The utilization of machine learning techniques to autonomously analyze the behavior of network packets and classify them as normal or malicious is a promising way to address this issue. The objective of this investigation is to assess the suitability of a variety of machine learning algorithms for the resolution of network security issues by employing network data analysis as an illustration. This investigation assesses the efficacy of machine learning models in detecting network intrusions using the UNSW-NB15 dataset. This study's primary objective is to assess the effectiveness of various machine learning models, including Random Forest, K-Nearest Neighbors (KNN), Support Vector Machine (SVM), XGBoost, LightGBM, and Logistic Regression, in network security applications. According to the analysis, all models exhibited high classification accuracy; however, the LightGBM model attained the most remarkable results. This model exhibited the highest values of Accuracy (95.86%), Precision (96.02%), and F1-measure (96.99%), confirming its capacity to effectively manage complex and heterogeneous data. Overall, the study underscores the significance of selecting the most appropriate model based on the security system's objectives and the specifics of the data.

Key words: machine learning, network traffic, LightGBM, cybersecurity, IDS, data analysis.

1 Introduction

The world of today is undergoing a rapid transformation, and the number and complexity of cyber threats are increasing as a result of the increasing digitalization. Organizations and consumers worldwide are increasingly emphasizing information security. The active integration of machine learning and artificial intelligence into network security systems is being driven by the increasing ineffectiveness of traditional protection methods in this environment. Not only do these technologies enhance the precision of malware and network attack detection, but they also adjust to the constantly evolving tactics of attackers. New opportunities for the development of more robust and proactive mechanisms to safeguard data and infrastructure are presented by the introduction of intelligent approaches to information systems protection.

A model for the automatic detection of P-wave seismic signals in the Almaty region was devised by D. Zhexebay and his colleagues using convolutional neural networks (CNNs) [1]. The authors determined that the implementation of deep learning enhances the dependability of earthquake early warning systems and can be effectively employed to reduce the risks associated with natural disasters. K. Moulaei et al. conducted a comparative analysis of various machine learning algorithms to predict lethal outcomes in patients hospitalized with COVID-19 [2]. The authors concluded that machine learning models, particularly RF, can be effectively employed to identify patients at a high risk of lethality and to assist in clinical decision-making.

Advanced machine learning models are becoming more precise and resilient in their ability to map vulnerability to natural disasters, such as flooding. A novel method of Flood Susceptibility Mapping (FSM) was proposed by S. T. Seydi et al with the use of the Cascade Forest Model [3]. Zhao et al employed machine learning algorithms to investigate the capabilities of the Google Earth Engine (GEE) platform in land use and land cover classification (LULC) tasks [4].

Malware detection continues to be a significant obstacle in the realm of contemporary network security, necessitating a comprehensive strategy. The accuracy and comprehensiveness of threat detection can be substantially enhanced through the integration of machine learning techniques with both static and dynamic analysis. M. Ijaz, M. H. Durad, and M. Ismail in their paper, examined the efficacy of machine learning techniques for malware detection through both static and dynamic analysis [5]. The problem of polymorphic malware detection was examined by M. S. Akhtar and T. Feng through the application of various machine learning algorithms [6]. R. Baker del Aguila and his associates investigated the feasibility of employing resource-efficient machine learning models for static malware analysis [7].

Intrusion Detection Systems (IDS) assist in the development of anti-cyberattack defenses by analyzing network traffic and detecting suspicious activity. Some new and complex forms of attacks may not be effectively detected by traditional rule-based methods. Machine learning (ML) techniques are being integrated into intrusion detection systems to enhance their capacity to autonomously identify attacks in order to resolve this issue.

Traditional machine learning (ML), ensemble learning, and deep learning methods were evaluated by C. Zhang et al. [8]. Based on the KDD CUP99 and NSL-KDD datasets, experimental evaluations were conducted using the following algorithms: Decision Tree, naive Bayesian algorithm, SVM, Random Forest, XGBoost, CNN, and RNN. The results indicated that the naive Bayesian algorithm was more effective in detecting new forms of attacks, while the ensemble learning methods demonstrated high accuracy. Furthermore, the efficacy of deep learning methods was determined to be contingent upon their structure, hyperparameters, and training duration. The primary challenges and prospective research directions in network attack detection are also emphasized by the authors.

Z. Ahmad and his colleagues examine the most recent advancements and obstacles associated with network intrusion detection systems (IDS) in their paper [9]. They outline the potential of deep learning (DL) and machine learning (ML) techniques for the effective detection of attacks. The authors underscore the necessity of enhancing the precision of intrusion detection systems,

minimizing false positives, and identifying novel types of attacks. Furthermore, they recommend future research directions and classify intrusion detection systems according to machine learning/deep learning.

In the context of a large network data and the Internet of Things (IoT), M. Asif et al. suggest a MapReduce-based intelligent model (MR-IMID) for security [10]. In order to avert future assaults, this model detects network attacks in real-time and stores the data in a database. The MR-IMID model demonstrated an accuracy of 97.7% during training and 95.7% during validation, surpassing previous methods.

A machine learning (ML)-based intelligent intrusion detection system (IDS) for detecting cyber attacks targeting Internet of Things devices is examined by authors D. Musleh, M. Alotaibi, F. Alhaidari, A. Rahman and R. Mohammad [11]. The investigation assessed numerous machine learning models (including Random Forest, KNN, SVM, and stack models) and various feature extraction algorithms (including DenseNet and VGG-16). The model that demonstrated the highest accuracy (98.3%) when utilizing the IEEE Dataport dataset was a combination of VGG-16 and layering.

The Intrusion Detection System (IDS) model proposed by H. Attou, A. Guezzaz, S. Benkirane, M. Azrour, and Y. Farhaoui [12] is a security-enhancing approach to cloud computing systems that employs Random Forest (RF) and feature engineering. The accuracy of the model was 98.3% on the Bot-IoT dataset and 99.99% on the NSL-KDD dataset. The authors observe that this method outperformed other state-of-the-art methods in terms of precision, accuracy, and completeness.

A machine learning-based intrusion detection system (ML-IDS) model for detecting attacks in IoT networks is presented by Y. Saheed, A. Abiodun, S. Misra, M. Holone, and R. Colomo-Palacios [13]. The UNSW-NB15 dataset was employed in the study, and dimensionality reduction techniques were implemented through PCA and minimum and maximum normalization. The results of the testing of six distinct machine learning algorithms were as follows: PCA-XgBoost achieved an accuracy of 99.99%, PCA-Cat Boost achieved an accuracy of 99.99%, PCA-KNN achieved an accuracy of 99.98%, PCA-SVM achieved an accuracy of 99.98%, PCA-QDA achieved an accuracy of 99.97%, and PCA-NB achieved an accuracy of 97.14%. The authors observe that this method is in competition with existing methods [14].

A. Turukmane and R. Devendiran suggest a sophisticated Intrusion Detection System (IDS) that is automated and based on machine learning as a solution to the detection of network intrusions in their research paper [15]. They employed techniques such as Min-Max normalization, zero-value processing, class inequality removal using ASmoT, and meaningful feature extraction using M-SvD to analyze the CSE-CIC-IDS 2018 and UNSW-NB15 datasets. Furthermore, the Mud Ring Multilayer SVM (M-MultiSVM) was employed to classify the attack classes, and the Opposition-based Northern Goshawk Optimization (ONgO) was employed to select the most appropriate features [16]. Consequently, the proposed system obtained an accuracy of 99.89% (CSE-CIC-IDS 2018) and 97.535% (UNSW-NB15). The K-Nearest Neighbor, Naïve Bayes, Support Vector Machine were tested in detecting 19 different types of attacks based on data from the UNSW-NB15 dataset, and the SVM algorithm showed the best results with an accuracy of 97.78%, exceeding the others. Despite its ability to achieve 98.9% accuracy on the KDD'99 Cup dataset, the SVM algorithm proved less effective than the Random Forest algorithm, which demonstrated the highest result of 99.81% [17]. The study also analyzed the classification of machine learning algorithms used in intrusion detection systems (IDS) in areas such as IoT, Big Data, Fog computing, and 5G networks. However, the following study notes that the advantage of this algorithm is revealed when computational resources are limited [18].

The comparative analysis of state-of-the-art ensemble and classical machine learning algorithms on the real and diverse UNSW-NB15 dataset is the unique feature of this study. The concentration is on the efficacy of the models in the face of cyber threats.

2 Materials and methods

2.1 Description of machine learning

Machine learning (ML) is a subfield of artificial intelligence that enables systems to autonomously learn from data and make predictions or decisions without the need for explicit programming. It is predicated on the development of models that can recognize patterns in historical data and apply those patterns to the analysis of new input data. The machine learning process is comprised of numerous critical stages, each of which is crucial to the development of the effective model depicted in Figure 1.

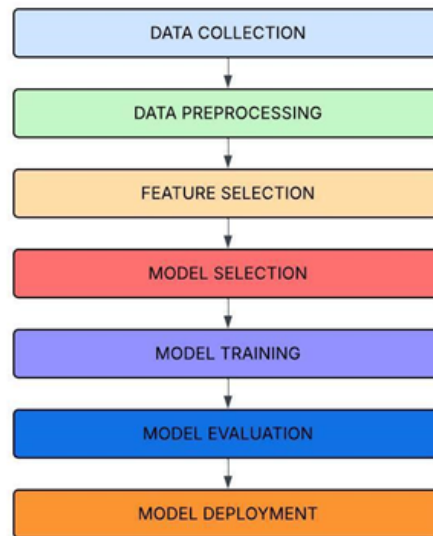


Figure 1 – Block diagram of the machine learning process

The process commences with the acquisition of data from a variety of sources, including sensors, databases, and web services. Subsequently, the data is subjected to preprocessing to eliminate outliers, duplicates, noise, normalize, and code features. Subsequently, feature selection is implemented, which involves the identification of the most critical attributes that influence the outcome in order to enhance the model's precision and simplify its structure. This is followed by the selection of an appropriate learning algorithm based on the data structure and problem type. The subsequent phase involves the model's training, which involves the analysis of the training data to identify the internal dependencies between inputs and outputs. Following the training process, the model is assessed on a test sample using a variety of metrics, including the F1-measure and accuracy, to guarantee its efficacy. The model's final stage is to be deployed in a real environment, where it begins to process real data and generate predictions.

2.2 Data preparation

In 2015, the Australian Centre for Cyber Security developed the UNSW-NB15 dataset, which served as the foundation for the investigation. This dataset was produced by utilizing IXIA PerfectStorm traffic, which offers a more accurate and current perspective on the behavior of both legitimate and malicious network traffic. The dataset comprises a total of 175,341 network sessions, of which 56,000 are normal and 119,341 are attacking. The nine primary categories of attack traffic are as follows: Fuzzers (introduction of incorrect data to disrupt the system), Analysis (analysis and scanning attacks), Backdoors (stealthy remote access), DoS (denial of service attacks), Exploits (vulnerability exploitation), Generic (cryptographic attacks), Reconnaissance (gathering information about the system), Shellcode (injection of malicious code through the shell), and Worms (self-propagating malware).

UNSW-NB15 comprises 49 attributes, each of which delineates a distinct characteristic of a network connection. Basic attributes (e.g. IP addresses, ports, and protocols), traffic content attributes (e.g. HTTP methods and FTP commands), temporal attributes (duration, connection start and end timestamps), flow attributes (amount of data transferred, number of packets), and additional automatically generated attributes (e.g. delay values, packet intervals, and TCP connection attributes) are all possible categories into which these attributes can be roughly categorized. Furthermore, class labels are present, including a binary label for normal traffic (0) and an `attack_cat` label for aggressor traffic (1), which denotes the category of the attack. In order to guarantee the generalizability of the models, the data were partitioned into training and test samples in an 80/20 proportion using `train_test_split` (`random_state=42`).

The `id` and `attack_cat` features were eliminated from the dataset during the preprocessing stage, as they were deemed unsuitable for binary classification tasks. In order to guarantee the accurate performance of feature scale-sensitive models, including Support Vector Machine (SVM) and K-Nearest Neighbors (KNN), categorical variables were transformed using one-hot coding and numeric features were scaled using `StandardScaler`. This data preparation enabled the formation of a balanced and homogeneous sample for the purpose of further training machine learning models.

In this study, the UNSW-NB15 dataset was selected as the primary benchmark for evaluating the performance of machine learning models. To justify this choice, a comparative analysis was conducted with two other widely used datasets in the field of intrusion detection: NSL-KDD and CICIDS2017. The comparison highlights the strengths of UNSW-NB15 in terms of modernity, class balance, and relevance to contemporary cybersecurity challenges.

Table 1 – Comparison of Popular NIDS Datasets

Characteristic	UNSW-NB15	NSL-KDD	CICIDS2017
Release Year	2015	2009 (based on 1999 KDD99)	2017
Number of Features	49 (original) / ~186 (after one-hot)	41	>80
Number of Attack Types	9 categories	4 categories	15+ attack scenarios
Traffic Type	Modern protocols (FTP, SSH, HTTP, DNS)	Outdated protocols (Telnet, ICMP)	Realistic mixed traffic
Class Balance	Moderate (~60% attack traffic)	Highly imbalanced	Highly imbalanced
Threat Relevance	Modern threats	Outdated threats	Modern threats
Dataset Size	~2 million packets	~150,000 records	>3 million packets
Realism	High (generated via IXIA tools)	Low (synthetic benchmark)	Very high (real-world captures)

As shown in the comparison in Table 1, NSL-KDD is an improved version of the outdated KDD99 dataset. However, it does not reflect modern network behavior or attack patterns. CICIDS2017 offers high realism and a diverse set of attack scenarios but suffers from extreme class imbalance and excessive volume, which complicates training and evaluation processes.

In contrast, UNSW-NB15 provides a balanced trade-off between size, realism, and diversity of attack types. It includes up-to-date network traffic and protocols, a reasonable number of well-defined classes, and moderate class balance, making it ideal for building and benchmarking intrusion detection models.

2.3 Model training and evaluation

The system initially receives input network traffic, which is data regarding network connections, including both legitimate and malicious sessions. This data is gathered using specialized tools, such as IXIA PerfectStorm, to provide a realistic and current representation of the network behavior in Figure 2.

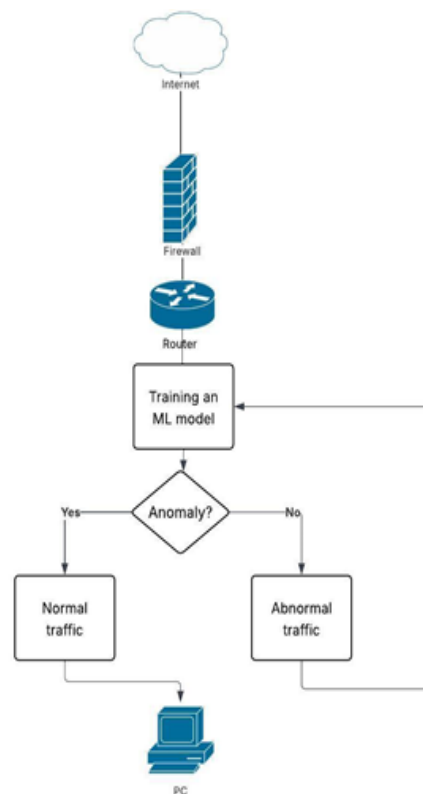


Figure 2 – Process of anomaly detection in network traffic using machine learning

The traffic is subsequently subjected to a data preprocessing stage. The set is purged of irrelevant features, including connection identifiers and assault categories, as they do not offer any valuable information for binary classification tasks. Numerical features are scaled using standard normalization, while categorical variables are transformed using one-hot encoding. This process guarantees the accurate operation of machine learning algorithms, particularly those that are vulnerable to the scope of the data, such as the support vector method or K-nearest neighbors.

Machine learning models are trained using the data that has been preprocessed. The study employs six distinct algorithms: Random Forest, K-Nearest Neighbors, Support Vector Machine, XGBoost, LightGBM, and Logistic Regression. From interpretability and fundamental accuracy to high performance and robustness to overtraining, each has its own parameters and advantages. Training is conducted on a labeled sample that is divided into normal and assault traffic.

The model that has been trained subsequently acquires new data and performs classification. The decision phase is responsible for determining whether a specific network connection is normal or anomalous. The connection is classified as anomalous if the features correspond to malevolent behavior patterns. Therefore, the system is capable of efficiently filtering network traffic, thereby detecting potential hazards in near real-time.

Six machine learning algorithms were selected for comparison:

Random Forest

The Random Forest Classifier is implemented by employing a fixed random number generator `sid` (`random_state=42`) and 100 decision trees (`n_estimators=100`). This method of ensemble construction achieves high accuracy and robustness against overfitting by constructing multiple trees and averaging their predictions.

`RandomForestClassifier(n_estimators=100, random_state=42)`

K-Nearest Neighbors (KNN)

The parameter `n_neighbors=5` is employed in conjunction with the K nearest neighbors method. It categorizes objects based on the most labels among the adjacent neighbors in the feature space. It is imperative to scale the data prior to training, as the distances between nodes have a direct impact on the outcome.

`KNeighborsClassifier(n_neighbors=5)`

Support Vector Machine (SVM)

The probability parameter = True and the linear kernel (`kernel='linear'`) were implemented in the support vector algorithm to facilitate the construction of the ROC curve. SVM is capable of effectively handling binary classification tasks in the presence of a substantial number of features.

`SVC(kernel='linear', probability=True, random_state=42)`

XGBoost

In order to construct a robust gradient-enhancing model, the XGBoost algorithm is employed with the following parameters: `n_estimators=200`, `learning_rate=0.1`, and `max_depth=6`. The loss during the learning phase is estimated using the `eval_metric='logloss'` parameter. This algorithm exhibits a high degree of accuracy and a strong capacity for generalization.

`XGBClassifier(n_estimators=200, learning_rate=0.1, max_depth=6, eval_metric='logloss', random_state=42)`

LightGBM

LightGBM, like XGBoost, is a gradient-based booleaning algorithm; however, it operates more efficiently on data of substantial magnitude. The parameters `n_estimators=200`, `learning_rate=0.1`, and `max_depth=10` are employed to achieve a balance between performance and accuracy.

`LGBMClassifier(n_estimators=200, learning_rate=0.1, max_depth=10, random_state=42)`

Logistic regression

In order to ensure convergence with a substantial number of features, the logistic regression model was implemented with an increased number of iterations (`max_iter=1000`). Logistic regression is a fundamental model that is frequently employed as an initial benchmark and is easily interpreted.

`LogisticRegression(max_iter=1000, random_state=42)`

Each model was trained on a standardized training sample and subsequently tested on the deferred portion. The evaluation was conducted using the following metrics:

- ♦ Accuracy – The total proportion of correct predictions.
- ♦ Precision – The accuracy of attack classification.
- ♦ Recall (completeness) - the model's ability to find all attacks.
- ♦ F1-score is the harmonic mean between precision and recall.
- ♦ ROC AUC is the area under the ROC curve showing the quality of the binary classification at different thresholds.
- ♦ Confusion Matrix – for the visual evaluation of classification errors.

A comparative ROC curve was also generated, which enabled a visual comparison of the models' quality in terms of AUC by calculating TPR (True Positive Rate) and FPR (False Positive Rate) values for each model and interpolating them over the total interval.

3 Results and discussion

3.1 Results of the ML algorithms

Based on the results of the experiments, the LightGBM model demonstrated the most effective performance, as illustrated in Table 2. The model's exceptional capacity to differentiate between assaults and normal traffic was demonstrated by its high classification accuracy of 95.86% and an area under the ROC curve (AUC-ROC) value of 94.6%. Furthermore, the model demonstrated a precision of 96.02%, recall of 97.98%, and F1-count of 96.99%, indicating a high equilibrium between the accuracy and completeness of predictions.

The effectiveness of LightGBM in the network attack detection assignment is confirmed by these results, which are based on the features presented in Figure 3.

Table 2 – Results of the ML algorithms

Model	Accuracy	Precision	Recall	F1-score	AUC-ROC
Random Forest	95,83	95,96	98,0	96,97	94,6
KNN	93,77	94,45	96,53	95,48	92,2
SVM	93,27	91,12	99,86	95,29	89,5
XGBoost	95,74	95,91	97,92	96,9	94,5
LightGBM	95,86	96,02	97,98	96,99	94,6
Logistic Regression	93,42	92,09	98,83	95,34	90,3

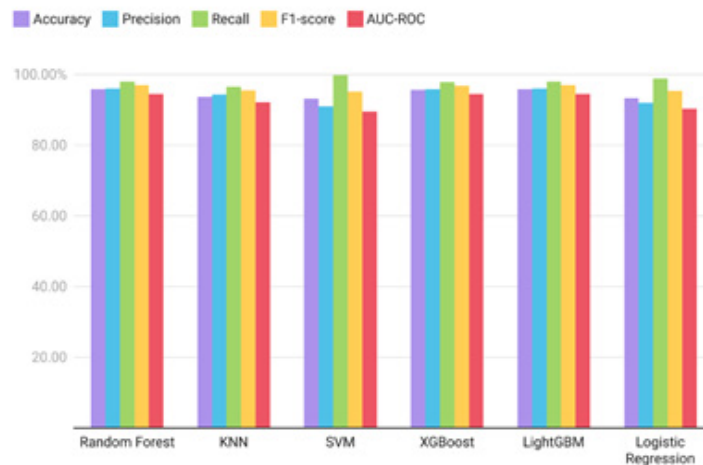


Figure 3 – Results of the ML algorithms

LightGBM showed the highest AUC-ROC value (94,6) in Figure 4. In general, LightGBM was the most effective model. The study's findings indicated that machine learning models are highly effective in the detection of attacks. The LightGBM model, in particular, demonstrated exceptional performance on all critical metrics, owing to its ensemble nature and capacity to effectively adjust to the complexity of the data. This demonstrates that it will be highly beneficial in real-world applications, particularly when analyzing large volumes of network traffic.

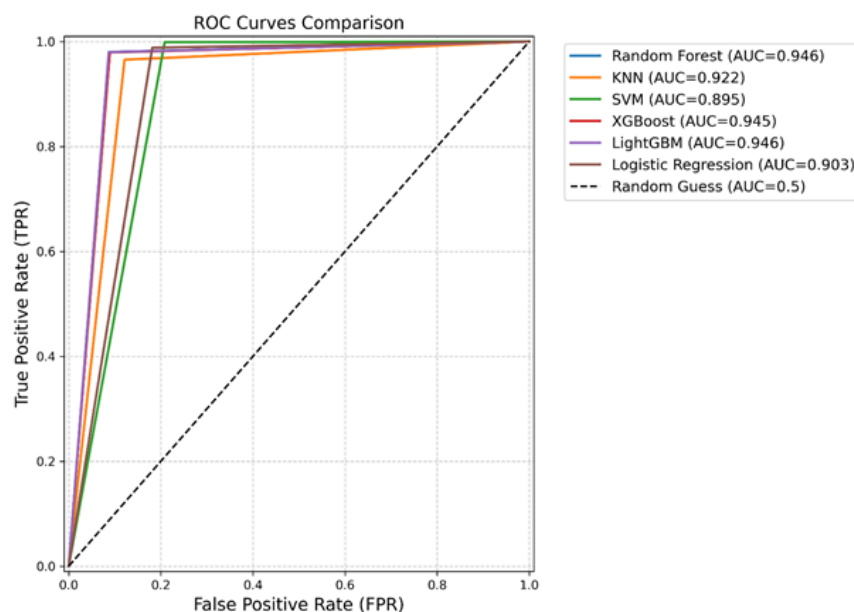


Figure 4 – Comparison of ROC curves for the six models

Similar high performance was also obtained by the Random Forest and XGBoost models. This is a result of their structural similarity (ensembles of decision trees) and their capacity to iteratively reduce errors. The SVM model was effective in maximizing assault coverage (Recall); however, it was inferior in terms of overall performance (AUC-ROC, Precision). Although the logistic regression and KNN algorithms are straightforward and basic models, they demonstrated relatively limited performance in identifying complex types of attacks, as illustrated in Figure 5.

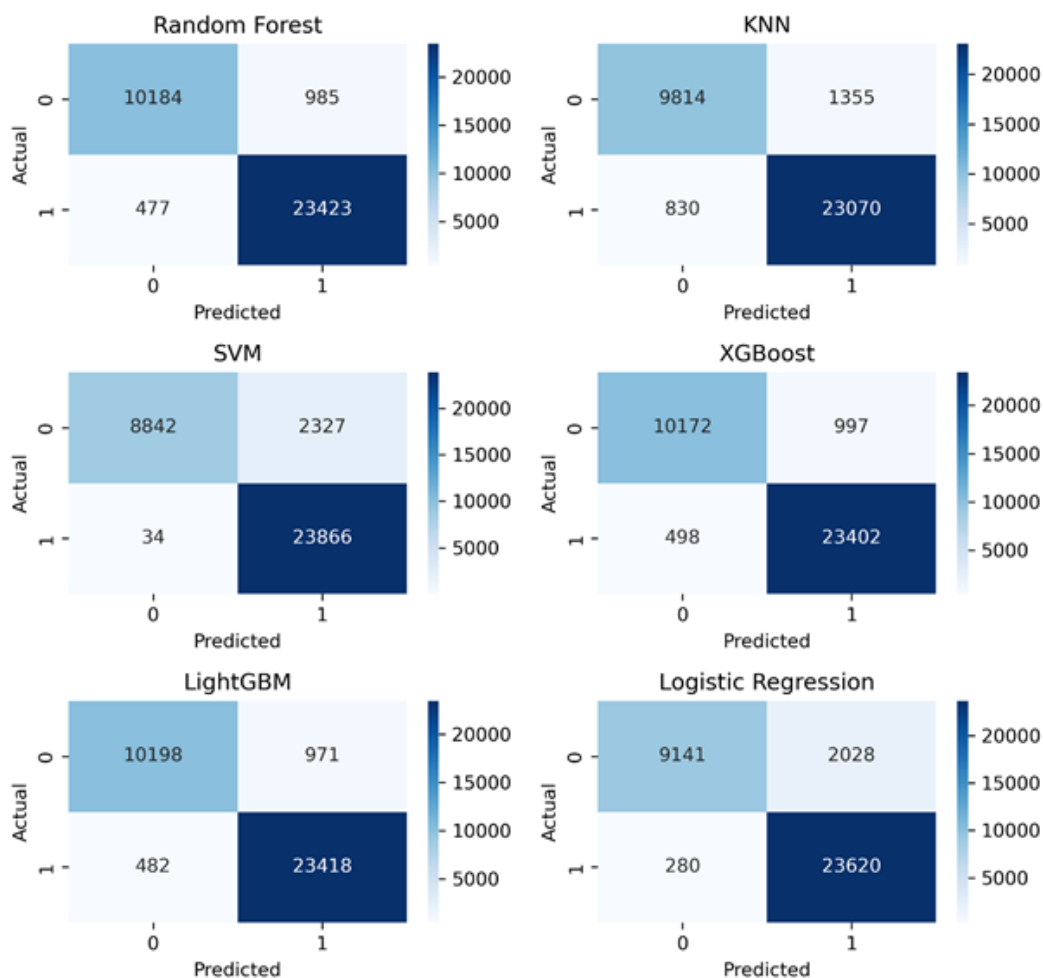


Figure 5 – Confusion matrix analysis for the six models

It is important to acknowledge that the UNSW-NB15 dataset utilized in the study encompasses specific categories of network attacks; however, it may be deficient in certain aspects of real-time complexity. Additional testing is necessary to conduct a comprehensive assessment of the performance of these models in real-world scenarios. Furthermore, it is advised that future research make use of alternative datasets (e.g., CICIDS, NSL-KDD) and assess the efficacy of the models' application to real-time systems.

Therefore, the findings of this investigation demonstrate that reinforcement models, including LightGBM, are ideal for practical application in the context of automated attack detection, thereby facilitating the development of effective network security decisions.

An analysis of classification errors by attack category (attack_cat) was conducted to conduct a more thorough assessment of the quality of the models. False positive (FP) and false negative (FN) classifications were given particular attention, as they are essential components of intrusion detection systems.

Table 3 provides a summary of the analysis' findings, including the quantity of FP and FN errors for each category of attacks.

Table 3 – Classification errors by attack category

Attack category	False Positives	False Negatives
Fuzzers	0	416
Analysis	0	40
DoS	0	2
Exploits	0	18
Generic	0	1
Reconnaissance	0	3
Shellcode	0	2
Normal	971	0

The analysis indicates that the Exploits category has the maximum number of FNs, which may be attributed to the diverse attack behavior patterns in this group outlined in Figure 6. This suggests that it is necessary to enhance the sensitivity of models to this category.

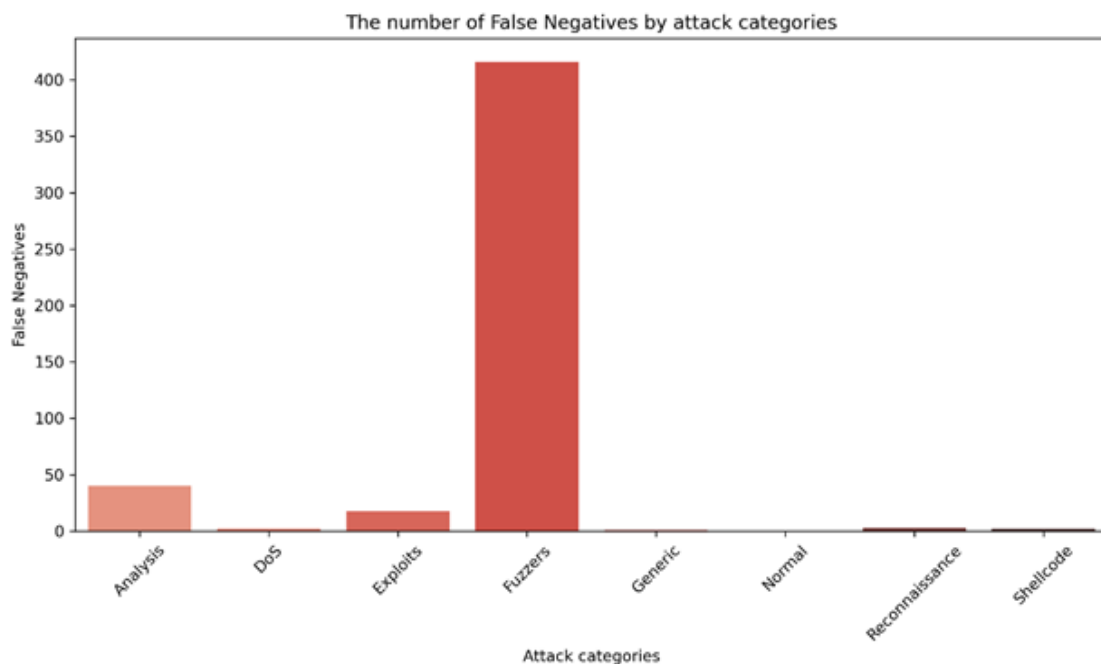


Figure 6 – Graph of the number of False Negatives by attack category

3.2 Evaluation of efficiency by resource consumption

The results of the measurements of peak memory usage, prediction time, and training time are summarized in Table 4.

For the practical application of the models, it is crucial to evaluate not only accuracy but also memory and time efficiency, as illustrated in Figure 7.

Table 4 – Comparison of resource intensity of models

Model	Training time (s)	Prediction time (s)	Memory (MB)
Random Forest	19.296	0.466	114.5
KNN	0.091	19.070	213.9
SVM	15833.081	60.857	416.5
XGBoost	2.820	0.030	1.5
LightGBM	1.906	0.071	4.1
Logistic Regression	8.979	0.007	212.3

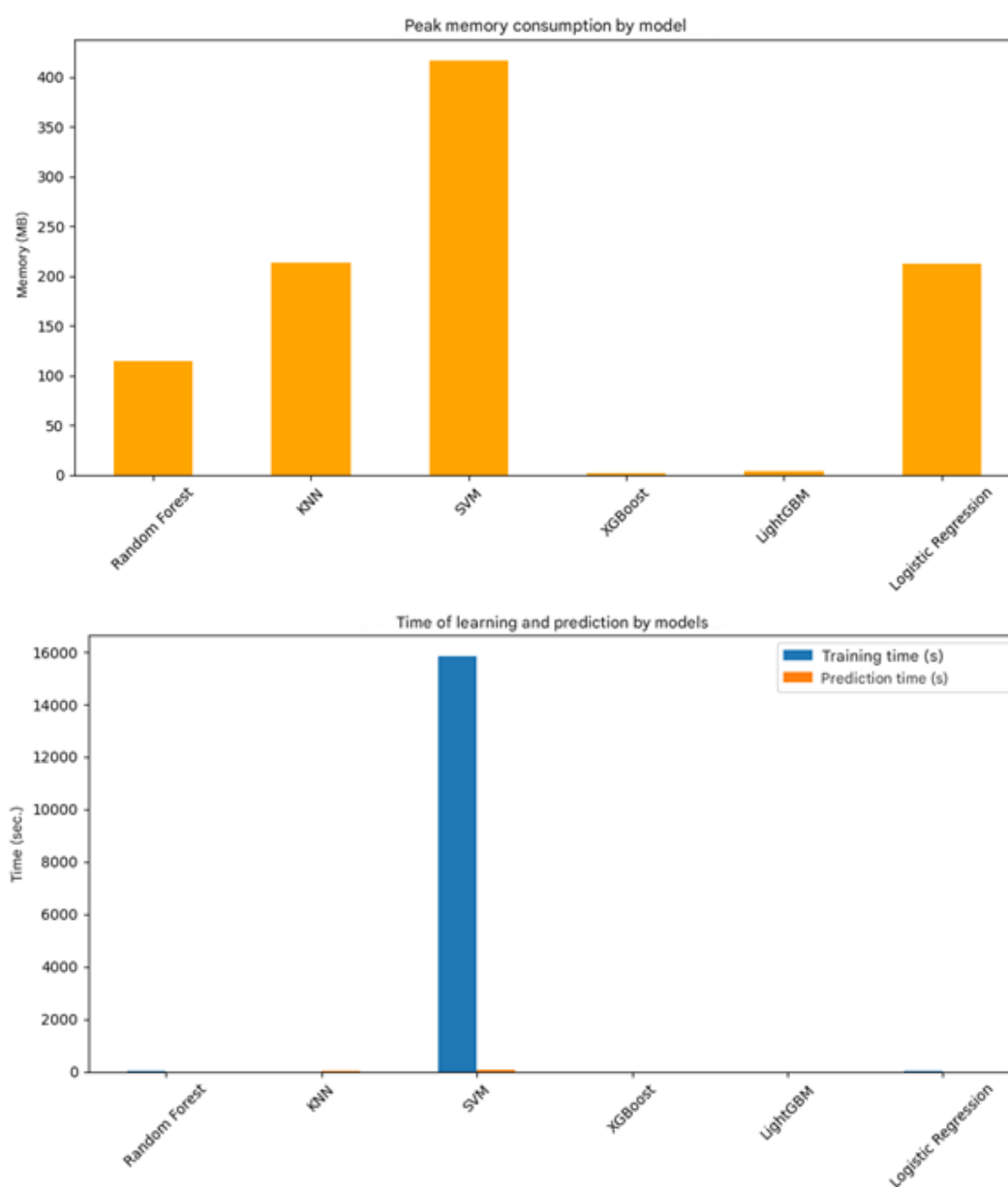


Figure 7 – Comparative plot of learning and memory time

The results indicate that LightGBM exhibits the highest accuracy/performance ratio, with minimal training time and relatively low resource consumption.

3.3 Configuring hyperparameters

Heuristics and expert settings were implemented to enhance the models' precision through manual hyperparameter selection. The parameters utilized for each model and their fitting procedure are summarized in Table 5.

Table 5 – Model hyperparameter settings

Model	Selection method	Parameters	Values
Random Forest	Manual selection	n_estimators	100
KNN	Manual selection	n_neighbors	5
SVM	Manual selection	kernel, probability	linear, True
XGBoost	Manual selection	n_estimators, max_depth, learning_rate	200, 6, 0.1
LightGBM	Manual selection	n_estimators, max_depth, learning_rate	200, 10, 0.1
Logistic Regression	Manual selection	max_iter	1000

To ensure optimal model performance, hyperparameter tuning was performed using Optuna, a Bayesian optimization framework. The study optimized n_estimators, learning_rate, and max_depth for LightGBM, as well as n_estimators, max_depth, and min_samples_split for Random Forest over 30 optimization trials each, using 3-fold cross-validation and ROC-AUC as the objective function in Table 6.

Table 6 – Results of LightGBM and Random Forest

Model	Accuracy	Precision	Recall	F1-score	AUC-ROC	Best Parameters
LightGBM	95,97	96,09	98,06	97,07	99,40	n_estimators=279, learning_rate=0.0910, max_depth=14
Random Forest	95,92	95,81	98,32	97,05	99,40	n_estimators=297, max_depth=26, min_samples_split=5

Due to their superior baseline performance, hyperparameter optimization was initially applied to LightGBM and Random Forest classifiers. Other models were evaluated with standard or manually selected parameters, as their optimization yielded marginal or less impactful improvements.

Both LightGBM and Random Forest, after Bayesian optimization via Optuna, achieved almost identical performance with an AUC of 99,4. LightGBM demonstrated slightly higher precision (96,09) and a better F1-score (97,07), while Random Forest had slightly better recall (98,32). These results confirm the high effectiveness of ensemble models in detecting network anomalies in the UNSW-NB15 dataset.

3.4 Importance of attributes

In order to enhance the interpretability of the models, feature importance diagrams were generated for the LightGBM and Random Forest models contained in Table 7.

Table 7 – Top 10 traits by importance (LightGBM)

№	sign	Importance value
1	smean	575
2	sbytes	527
3	ct_srv_src	432
4	ct_srv_dst	266
5	ct_dst_src_ltm	246
6	sload	229
7	dbytes	201
8	synack	198
9	stcpb	195
10	dmean	187

The 10 most critical features, as indicated by the LightGBM version, are illustrated in the Figure 8.

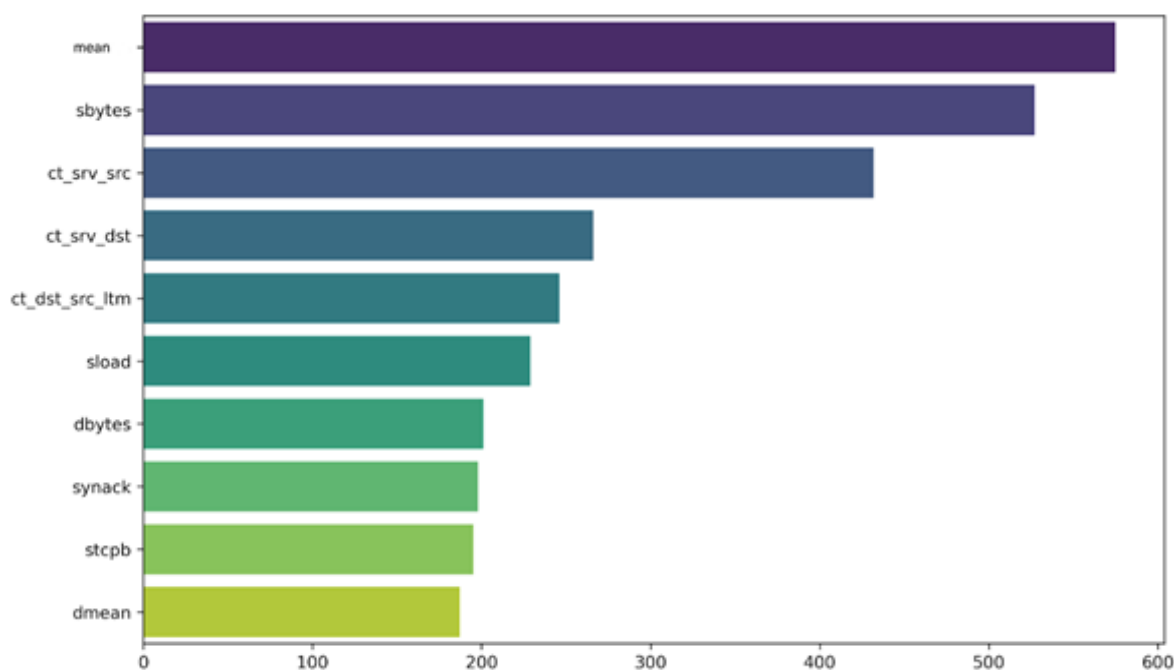


Figure 8 – Top 10 attributes by importance of LightGBM

Critical attributes include:

- ♦ sbytes (bytes sent)
- ♦ dbytes (bytes received)
- ♦ ct_state_ttl (connection state and TTL)
- ♦ ct_dst_sport_ltm (port destination frequency)
- ♦ sttl (packet lifetime)

These features have the greatest impact on the final solution of the model, which confirms their importance in the task of detecting anomalous behavior in the network.

3.5 Error Analysis

A refined evaluation of classification results underscores the significance of analyzing false positives (FP) and false negatives (FN) in intrusion detection. For the LightGBM model, 950 FPs and 463 FNs were observed, while the Random Forest model yielded 1028 FPs and 403 FNs. Although overall error rates were moderate, false negatives are particularly critical as they represent undetected attacks.

Key causes of false negatives include overlapping feature distributions between benign and malicious traffic, underrepresentation of specific attack types, and high intra-class variability. Conversely, false positives may arise from benign behavior mimicking attacks under certain conditions, increased noise due to high-dimensional features, and classifier uncertainty near decision boundaries.

To mitigate these errors, the following strategies are recommended:

- ♦ Employ class balancing techniques (e.g., SMOTE, class weight adjustment) to reduce FNs.
- ♦ Apply feature selection or dimensionality reduction to minimize FPs.
- ♦ Optimize classification thresholds using validation metrics.
- ♦ Evaluate model performance per attack type to identify and address specific weaknesses.

Several related works have evaluated machine learning models on the UNSW-NB15 dataset, yet their reported performance metrics vary depending on the selected models, features, and experimental setups.

For example, the study by Yakub Kayode Saheed et al. applied PCA and XGBoost to the UNSW-NB15 dataset and achieved an accuracy of 99.99%. However, the study did not report detailed F1-scores or confusion matrices, making it difficult to assess classification robustness across all classes [19].

Another study by Turukmane and Devendiran used a multi-class SVM ensemble on the CSE-CIC-IDS2018 dataset, reporting an F1-score of 99.89%. While the performance is high, the dataset used was significantly larger and more diverse, which limits the comparability to UNSW-NB15 results [20].

In contrast, our optimized LightGBM and Random Forest models, evaluated strictly on the UNSW-NB15 dataset, achieved F1-scores of 97.07% and 97.05%, respectively, along with ROC AUC scores of 0.994, which are highly competitive. Additionally, our models demonstrated strong generalization without relying on excessive feature engineering or complex ensembles.

This comparison emphasizes the effectiveness and simplicity of the proposed approach, showcasing that with proper hyperparameter optimization, standard tree-based models can achieve state-of-the-art results on modern intrusion detection datasets.

Conclusion

The performance of a variety of machine learning algorithms, such as Random Forest, K-Nearest Neighbors (KNN), Support Vector Machine (SVM), XGBoost, LightGBM, and Logistic Regression, was compared in this study. All of the models that were evaluated exhibited exceptional performance in the classification of network traffic; however, certain algorithms were particularly noteworthy.

After applying Bayesian hyperparameter optimization using Optuna, both LightGBM and Random Forest models achieved significant performance improvements across all key evaluation metrics. The LightGBM model demonstrated the best overall performance, with an accuracy of 95.97%, precision of 96.09%, recall of 98.06%, and an F1-score of 97.07%, alongside a near-perfect ROC AUC score of 0.994. Random Forest model achieved nearly equivalent results, with an accuracy of 95.92%, precision of 95.81%, recall of 98.32%, F1-score of 97.05%, and the same ROC AUC score of 0.994. The high adaptability of this model to complex heterogeneous data structures, as well as its overall efficacy, are demonstrated by these figures. Gradient boosting, a high learning rate, the efficient handling of large quantities of data, and the ability to minimize errors at each iteration step are the primary benefits of LightGBM. In addition, the model is particularly well-suited for real-

time cyber threat detection systems due to its ability to perform well on categorical features and its reduced resource consumption.

Random Forest and XGBoost also demonstrated satisfactory performance, with an F1-estimation that exceeded 96.9% and an AUC-ROC value of approximately 94.5%, in addition to LightGBM. These findings substantiate the reliability of ensemble methods and their exceptional prediction accuracy in network security tasks.

The SVM model's recall was 99.86%, which was an intriguing characteristic that suggested its capacity to accurately identify all affirmative cases. Additionally, it demonstrated an exceptionally high level of sensitivity to attacks. Nevertheless, its AUC-ROC and overall accuracy were inferior to those of the ensemble models. The Logistic Regression and KNN models, despite their simplicity, also obtained satisfactory results; however, they demonstrated less robustness when confronted with more intricate types of attacks.

The research demonstrated that LightGBM is the most balanced model in terms of speed, responsiveness, and accuracy. These attributes render it an ideal choice for practical implementation in network threat detection systems. The significance of meticulously selecting a machine learning model in real-world applications is underscored by the paper, which considers the complexity of data, the availability of computational resources, and the accuracy requirements.

REFERENCES

- 1 Zhexebay, D., Skabylov, A., Ibraimov, M., Khokhlov, S., Agishev, A., Kudaibergenova, G., Orazakova, A., & Agishev, A. Deep Learning for Early Earthquake Detection: Application of Convolutional Neural Networks for P-Wave Detection. *Applied Sciences*, 15(7), 3864 (2025). <https://doi.org/10.3390/app15073864>.
- 2 Moulaei, K., Shanbehzadeh, M., Mohammadi-Taghiabad, Z., Mousavi, S. F., & Jafari, S. Comparing machine learning algorithms for predicting COVID-19 mortality. *BMC Medical Informatics and Decision Making*, 22(1), 2 (2022). <https://doi.org/10.1186/s12911-021-01742-0>.
- 3 Seydi, S. T., Kanani-Sadat, Y., Hasanlou, M., Sahraei, R., Chanussot, J., & Amani, M. Comparison of Machine Learning Algorithms for Flood Susceptibility Mapping. *Remote Sensing*, 15(1), 192 (2023). <https://doi.org/10.3390/rs15010192>.
- 4 Zhao, Z., Islam, F., Waseem, L. A., Tariq, A., Nawaz, M., Islam, I. U., Bibi, T., Rehman, N. U., Ahmad, W., Aslam, R. W., Raza, D., & Hatamleh, W. A. Comparison of three machine learning algorithms using Google Earth Engine for Land Use Land Cover classification. *Rangeland Ecology & Management*, 92, 129–137 (2024). <https://doi.org/10.1016/j.rama.2023.10.007>.
- 5 Ijaz, M., Durad, M. H., & Ismail, M. Static and dynamic malware analysis using machine learning. In 2019 16th International Bhurban Conference on Applied Sciences and Technology (IBCAST) (pp. 687–691). IEEE (2019). <https://doi.org/10.1109/IBCAST.2019.8667136>.
- 6 Akhtar, M. S., & Feng, T. Malware Analysis and Detection Using Machine Learning Algorithms. *Symmetry*, 14(11), 2304 (2022). <https://doi.org/10.3390/sym14112304>.
- 7 Baker del Aguila, R., Contreras Pérez, C. D., Silva-Trujillo, A. G., Cuevas-Tello, J. C., & Nunez-Varela, J. Static Malware Analysis Using Low-Parameter Machine Learning Models. *Computers*, 13(3), 59 (2024). <https://doi.org/10.3390/computers13030059>.
- 8 Zhang, C., Jia, D., Wang, L., Wang, W., Liu, F., & Yang, A. Comparative research on network intrusion detection methods based on machine learning. *Computers & Security*, 121, 102861 (2022). <https://doi.org/10.1016/j.cose.2022.102861>.
- 9 Ahmad, Z., Khan, A. S., Shiang, C. W., Abdullah, J., & Ahmad, F. Network intrusion detection system: A systematic study of machine learning and deep learning approaches. *Transactions on Emerging Telecommunications Technologies*, 32(1), e4150 (2021). <https://doi.org/10.1002/ett.4150>.
- 10 Asif, M., Abbas, S., Khan, M. A., Fatima, A., Khan, M. A., & Lee, S.-W. MapReduce based intelligent model for intrusion detection using machine learning technique. *Journal of King Saud University – Computer and Information Sciences*, 34(10, Part B), 9723–9731 (2022). <https://doi.org/10.1016/j.jksuci.2021.12.008>.
- 11 Musleh, D., Alotaibi, M., Alhaidari, F., Rahman, A., & Mohammad, R.M. Intrusion detection system using feature extraction with machine learning algorithms in IoT. *Journal of Sensor and Actuator Networks*, 12(2), 29 (2023). <https://doi.org/10.3390/jsan12020029>.

12 Attou, H., Guezzaz, A., Benkirane, S., Azrour, M., & Farhaoui, Y. Cloud-based intrusion detection approach using machine learning techniques. *Big Data Mining and Analytics*, 6(3), 311–320 (2023). <https://doi.org/10.26599/BDMA.2022.9020038>.

13 Saheed, Y.K., Abiodun, A.I., Misra, S., Holone, M.K., & Colomo-Palacios, R. A machine learning-based intrusion detection for detecting Internet of Things network attacks. *Alexandria Engineering Journal*, 61(12), 9395–9409 (2022). <https://doi.org/10.1016/j.aej.2022.02.063>.

14 Moustafa, N., & Slay, J. UNSW-NB15: A comprehensive data set for network intrusion detection systems (UNSW-NB15 network data set). 2015 Military Communications and Information Systems Conference (MilCIS), 1–6 (2015). <https://doi.org/10.1109/MilCIS.2015.7348942>.

15 Turukmane, A.V., & Devendiran, R. M-MultiSVM: An efficient feature selection assisted network intrusion detection system using machine learning. *Computers & Security*, 137, 103587 (2024). <https://doi.org/10.1016/j.cose.2023.103587>.

16 Ajagbe, S.A., & Alabi, O.O. Comparative Study of Machine Learning Models Using UNSW Datasets. In *The 45th Annual Conference of the South African Institute of Computer Scientists and Information Technologists* (p. 69). URL: https://saicsit2024.mandela.ac.za/saicsit2024/media/Store/documents/SAICSIT_PGS_v1-3.pdf#page=72.

17 Hussain, A., Khatoon, A., Aslam, A., & Khosa, M. A comparative performance analysis of machine learning models for intrusion detection classification. *Journal of Cybersecurity*, 6,1 (2024). <https://www.proquest.com/openview/08a4c605b57abbca71467cce40765b15/1?pq-origsite=gscholar&cbl=4585457>.

18 Mishra, N., & Mishra, S. A Review of Machine Learning-based Intrusion Detection System. *EAI Endorsed Transactions on Internet of Things*, 10 (2024). <https://doi.org/10.4108/eetiot.5332>.

19 Tahri, R., Benslimane, Y., Rifi, M., & Maqnaoui, M. Intrusion detection system using machine learning algorithms. *ITM Web of Conferences*, 46, 02003 (2022). <https://doi.org/10.1051/itmconf/20224602003>.

20 Saranya, T., Sridevi, S., Deisy, C., Chung, T. D., & Ahamed Khan, M.K. Performance analysis of machine learning algorithms in intrusion detection system: A review. *Procedia Computer Science*, 171, 1251–1260 (2020). <https://doi.org/10.1016/j.procs.2020.04.133>.

¹Кикбаев Н.Е.,

магистрант, ORCID ID: 0009-0000-0145-5718,
e-mail: kikbaevnurbek@gmail.com

¹Жексебай Д.М.,

PhD, ORCID ID: 0009-0008-1884-4662,
e-mail: zhexebay92@gmail.com

²Синь Ю.,

профессор, ORCID ID: 0000-0001-6169-0795,
e-mail: qyuxiao@purdue.edu

³Тынымбаев С.Т.,

профессор, ORCID ID: 0000-0002-9326-9476,
e-mail: s.tynym@gmail.com

³Айтмагамбетов А.З.,

профессор, ORCID ID: 0000-0002-9326-9476,
e-mail: a.aitmagambetov@iitu.edu.kz

¹Абдижалилова Л.Б.,

магистрант, ORCID ID: 0009-0000-5965-7195,
e-mail: abdijalil.lazzat@bk.ru

^{1*}Скабылов А.А.,

PhD, ORCID ID: 0000-0002-5196-8252,
*e-mail: Alisher.skabylov@kaznu.edu.kz

¹Әл-Фараби атындағы Қазақ ұлттық университеті, Алматы қ., Қазақстан

²Солтүстік-Батыс политехникалық университеті, Сиан қ., Қытай

³Халықаралық ақпараттық технологиялар университеті, Алматы қ., Қазақстан

ЖЕЛІЛІК ТРАФИКТЕГІ АНОМАЛИЯЛАРДЫ АНЫҚТАУ ҮШІН МАШИНАЛЫҚ ОҚЫТУ ӘДІСТЕРІН САЛЫСТЫРМАЛЫ ЗЕРТТЕУ

Аңдатпа

Киберқауіптердің қарқынды өсуіне және соның салдарынан желілік трафиктің артуына байланысты белгілі және жаңа шабуыл түрлерін жылдам анықтай алатын шабуылдарды анықтау жүйелеріне (IDS) сұраныс артып келеді. Желілік пакеттердің әрекетін автономды түрде талдау және оларды қалыпты немесе зиянды деп жіктеу үшін машиналық оқыту әдістерін пайдалану – бұл мәселені шешудің перспективалы тәсілі. Бұл зерттеудің мақсаты – иллюстрация ретінде желі деректерін талдауды қолдана отырып, желілік қауіпсіздік мәселелерін шешуде әртүрлі машиналық оқыту алгоритмдерінің тиімділігін бағалау. Зерттеу барысында UNSW-NB15 деректер жинағы негізінде желіге заңсыз кіруді анықтаудағы машиналық оқыту үлгілерінің өнімділігі тексерілді. Негізгі назар Random Forest, K-Nearest Neighbors (KNN), Support Vector Machine (SVM), XGBoost, LightGBM және Logistic Regression модельдеріне аударылды. Талдау нәтижесінде барлық модельдер жоғары классификация дәлдігін көрсетті; алайда LightGBM моделі ең үздік нәтижелерге қол жеткізді. Атап айтқанда, ол дәлдік бойынша – 95,86%, нақтылық бойынша – 96,02% және F1-өлшем бойынша – 96,99% нәтижелерін көрсетті. Бұл оның күрделі әрі біртекті емес деректерді тиімді басқару мүмкіндігін растады. Жалпы алғанда, зерттеу желілік қауіпсіздік жүйелерінде қолданылатын үлгілерді таңдаудың маңыздылығын айқындады. Нәтижелер IDS жүйелерін жобалауда нақты мақсаттар мен деректердің ерекшеліктеріне сәйкес ең қолайлы машиналық оқыту үлгісін таңдаудың тиімділігін дәлелдейді.

Тірек сөздер: машиналық оқыту, желілік трафик, LightGBM, киберқауіпсіздік, IDS, деректерді талдау.

¹Кикбаев Н.Е.,

магистрант, ORCID ID: 0009-0000-0145-5718,

e-mail: kikbaevnurbek@gmail.com

¹Жексебай Д.М.,

PhD, ORCID ID: 0009-0008-1884-4662,

e-mail: zhexebay92@gmail.com

²Синь Ю.,

профессор, ORCID ID: 0000-0001-6169-0795,

e-mail: qyuxiao@purdue.edu

³Тынымбаев С.Т.,

профессор, ORCID ID: 0000-0002-9326-9476,

e-mail: s.tynym@gmail.com

³Айтмагамбетов А.З.,

профессор, ORCID ID: 0000-0002-9326-9476,

e-mail: a.aitmagambetov@iitu.edu.kz

¹Абдижалилова Л.Б.,

магистрант, ORCID ID: 0009-0000-5965-7195,

e-mail: abdijalil.lazzat@bk.ru

^{1*}Сқабылов А.А.

PhD, ORCID ID: 0000-0002-5196-8252,

e-mail: Alisher.skabylov@kaznu.edu.kz

¹Казахский национальный университет им. аль-Фараби, г. Алматы, Казахстан

²Северо-Западный политехнический университет, г. Сиань, Китай

³Международный университет информационных технологий, г. Алматы, Казахстан

СРАВНИТЕЛЬНОЕ ИССЛЕДОВАНИЕ МЕТОДОВ МАШИННОГО ОБУЧЕНИЯ ДЛЯ ВЫЯВЛЕНИЯ АНОМАЛИЙ В СЕТЕВОМ ТРАФИКЕ

Аннотация

Спрос на системы обнаружения вторжений (IDS), которые могут оперативно определять как известные, так и новые типы атак, растет из-за быстрого расширения киберугроз и последующего увеличения сетевого

трафика. Использование методов машинного обучения для автономного анализа поведения сетевых пакетов и классификации их как нормальных или вредоносных является многообещающим способом решения этой проблемы. Целью данного исследования является оценка пригодности различных алгоритмов машинного обучения для решения проблем сетевой безопасности путем использования анализа сетевых данных в качестве иллюстрации. В данном исследовании оценивается эффективность моделей машинного обучения при обнаружении сетевых вторжений с использованием набора данных UNSW-NB15. Основная цель этого исследования – оценить эффективность различных моделей машинного обучения, включая случайный лес, метод К-ближайших соседей (KNN), опорную векторную машину (SVM), XGBoost, LightGBM и логистическую регрессию, в приложениях сетевой безопасности. Согласно анализу, все модели продемонстрировали высокую точность классификации; однако модель LightGBM достигла самых значительных результатов. Эта модель продемонстрировала самые высокие значения точности (95,86%), точности (96,02%) и F1-меры (96,99%), что подтверждает ее способность эффективно управлять сложными и неоднородными данными. В целом исследование подчеркивает важность выбора наиболее подходящей модели на основе целей системы безопасности и специфики данных.

Ключевые слова: машинное обучение, сетевой трафик, LightGBM, кибербезопасность, IDS, анализ данных.

Article submission date: 29.06.2025

UDC 004.89
IRSTI 28.23.37

<https://doi.org/10.55452/1998-6688-2025-22-4-97-106>

¹**Abdildaeva A.A.,**

PhD, acting Professor, ORCID: 0000-0002-6381-9350,

e-mail: assel.abdildaeva@kaznu.edu.kz

^{1*}**Nurtugan G.B.,**

PhD student, ORCID: 0000-0001-8824-9502,

*e-mail: g.nurtugun@gmail.com

²**Wojtkiewicz K.**

PhD, assistant Professor, ORCID: 0000-0002-7851-4330,

e-mail: krystian.wojtkiewicz@pwr.edu.pl

¹Al-Farabi Kazakh National University Almaty, Kazakhstan

²Wroclaw University of Science and Technology (WUST) Wroclaw, Poland

DEVELOPMENT AND OPTIMIZATION OF NEURAL NETWORK MODELS WITH ATTENTION MECHANISMS FOR INTRADAY PRICE FORECASTING FOR EUR/USD

Abstract

The study examines the problem of intraday forecasting of the EUR/USD currency pair using various neural network architectures, in particular models integrating attention mechanisms. Three neural network architectures were studied: the basic LSTM model, the LSTM model with the Bahdanau attention mechanism, and the Transformer model with the self-attention mechanism. The experiment was conducted on historical minute data for the period from January 2020 to December 2022. The results showed that attentional models are significantly superior to the basic LSTM architecture. The best results were obtained by the Transformer model (MSE=0.185, MAE=0.297, RMSE=0.431, MAPE=7.3%). A detailed analysis confirmed the stability and accuracy of the Transformer model. The identified advantages of attention models justify their prospects for use in algorithmic trading and require further research to optimize and adapt to real trading conditions. In particular, further research may be aimed at integrating attention models with trading strategies and risk management systems, as well as studying their behavior in the face of sudden changes in market volatility. In addition, it is proposed to explore the possibilities of combining attention architectures with other forecasting methods to increase the overall stability and reliability of forecasts in practical trading.

Keywords: intraday forecasting, Forex, exchange rates, EUR/USD, neural networks, attention mechanism, Transformer, LSTM, self-attention.

Introduction

Accurate forecasting of prices for currency pairs, especially on the scale of intraday trading, is a critically important task for participants in the foreign exchange market. The Forex market, where the EUR/USD currency pair is traded, is characterized by high volatility and significant noise levels, especially over short time intervals such as minutes. This significantly complicates the process of forecasting short-term price movements, setting traders and analysts the task of finding effective tools to capture complex and rapidly changing patterns [1].

Traditional forecasting methods, including statistical approaches (for example, ARIMA or exponential smoothing models), often prove to be insufficiently effective when working with high-frequency Forex data. Even classical neural network models such as LSTM and convolutional networks (CNN), despite their power and ability to detect nonlinear dependencies, are sometimes

unable to quickly and accurately adapt to sudden market changes. This is due to their limited ability to focus on the most significant elements of time sequences.

In this regard, this study aims to improve the quality of intraday price forecasting of the EUR/USD currency pair through the use of modern neural network architectures that integrate attention mechanisms. Attention mechanisms that have proven effective in areas such as natural language processing and sequence analysis have the potential to highlight the most significant and relevant segments of historical data. In the context of intraday forecasting, this can lead to a significant increase in prediction accuracy by selectively focusing the model on important time points.

The novelty of the research lies in the fact that the mechanisms of attention, including transformer-type architectures and hybrid models combining CNN and LSTM with attention, are still insufficiently studied in the field of algorithmic trading on the Forex market, especially on small time scales. Most of the research in the field of deep learning applied to finance has traditionally focused on daily and weekly intervals and is mainly focused on the stock market. Intraday Forex data and approaches using attention architectures have been studied much less. Thus, this study is aimed at filling the existing gap, assessing how effectively attention mechanisms can improve the quality of short-term forecasts of the minute interval in the foreign exchange market.

It is assumed that the results of this study will provide new knowledge and clearly demonstrate whether modern neural network architectures with attention mechanisms can significantly exceed the accuracy of traditional approaches based on intraday currency data. This will allow us to better understand the possibilities of applying attention models in the practice of algorithmic trading and lay the foundation for further developments in this area.

Traditional Neural Network Models for Intraday Forex

Long-term short-term memory (LSTM) and similar architectures have become the most popular models for predicting financial time series. LSTM captures time dependencies well and by 2020 has demonstrated superiority over classical methods in the task of short-term price forecasting. Thus, Yildirim et al., LSTM was used to predict the direction of movement of the EUR/USD pair using a hybrid architecture: one network was trained on macroeconomic indicators, the other on technical indicators [2]. Combining the results of two LSTMs through a special rule made it possible to improve the accuracy of the direction classification in comparison with a single LSTM for all signs. The authors also introduced a neutral «no change» class for small fluctuations and showed an increase in the final profit metric (profit_accuracy) to ~73.6%. This approach reduced the number of erroneous transactions, filtering out insignificant price fluctuations. In another study, a two-input LSTM model was proposed, where fundamental and technical indicators are presented separately; this approach also showed an increase in forecast accuracy compared to models using only one type of feature [3].

Other works have investigated various modifications of recurrent networks. In particular, a comparison of LSTM and its bidirectional version of BiLSTM demonstrated some advantage of Bidirectional-LSTM in the accuracy of forecasting exchange rates [4]. Hossain and colleagues proposed a hybrid GRU-LSTM for predicting the prices of four major pairs (EUR/USD, GBP/USD, etc.) at intraday intervals of 10 and 30 minutes [5]. Their two-layer network contains the GRU layer first, and then the LSTM; this combination surpassed the accuracy of the individual LSTM and GRU models, especially in the 10-minute interval. According to MSE, RMSE and MAE, the hybrid showed the smallest error (for 10-minute forecasts) among all comparative models, and in terms of the coefficient of determination R^2 it also turned out to be the best, which indicates its more stable results. These results confirmed the effectiveness of combining different types of RNNs to capture complex patterns of intraday fluctuations. Additionally, it was shown that the complication of the LSTM architecture itself in the form of its multilayer version also increases accuracy: the two-layer model (Stacked LSTM) for the AUD/USD pair surpassed the single-layer LSTM and a number of alternative approaches in terms of error indicators [6].

In addition to recurrent networks, convolutional neural networks (CNNs) were also used to extract local patterns from price ranges. However, the direct application of CNNs or deep multi-layer networks to a financial range can lead to instability due to sensitivity to noise [7]. Therefore, CNN is

often combined with RNN: for example, in one approach for Forex, it was proposed to first compress the input data of 1-D CNN, and then process the sequence with LSTM (CNN-LSTM). Such hybrids are designed to combine CNN's ability to recognize short patterns with LSTM's ability to work with long dependencies. For example, Markova et al. (2023) developed a CNN-LSTM auto-encoder model for predicting 5-minute EUR/USD prices, which demonstrated very high forecast accuracy ($RMSE \approx 0.0036$ on test data) [8]. In another study, the CNN+LSTM combination reduced the standard error of the forecast by about 9% compared to LSTM alone, while increasing the stability and learning rate of the model (~40% gain in training time) [9]. In addition, approaches are being explored taking into account the macroeconomic context: for example, Pornwattanavichai et al. (2022) proposed the BERTFOREX cascade model, where fundamental indicators (such as inflation, GDP, indices, etc.) are first processed using BERT and an autoencoder and then combined with technical indicators; such a hybrid achieved about 79.4% accuracy in forecasting course directions [10].

It is important to note that the success of the model based on historical data does not always guarantee a practical profit when trading costs are taken into account. A study by Ito et al. analyzed the prediction of intraday returns of currencies using LSTM based on events in the bid stack [11]. The model was good at predicting the direction of change in the mid-exchange rate (mid-quote) in minute increments, ahead of a number of other algorithms. However, when taking into account the spread and commissions, such a predictive model ceased to bring economic benefits, in fact confirming the hypothesis of market efficiency. This indicates a limitation of many purely predictive models: without taking into account market frictions, their results may be overestimated. Nevertheless, the very fact that the neural network is able to detect signals in the order flow is valuable – it indicates the presence of predictable patterns, even if they are not directly monetized in the presence of transaction costs.

It should also be mentioned that some comparative studies question the universality of LSTM. For example, Zafeiriou and Kalles tested several architectures on a short Forex trend and found that a special multi-layer perceptron network simulating a technical analyst worked no worse than LSTM with much less computing time [12]. This highlights that the success of models depends on fine-tuning and structure: increasing complexity does not always automatically lead to a better result without taking into account the specifics of the data.

Attention-oriented and hybrid approaches

Current trends in forecasting financial series are related to the introduction of attention mechanisms, which allow the model to focus on the most informative parts of the sequence. In the context of currency series, attention is able to identify important time steps (for example, reactions to news or certain market phases) and thereby improve the quality of the forecast compared to conventional LSTM/GRU, which «remember» everything. From 2020–2021, works began to appear adapting the successful Transformer architecture from the field of natural language processing to financial forecasting tasks. Transformer is completely based on the self-attention mechanism for modeling dependencies in a sequence and eliminates RNN memory length limitations.

One of the first examples of Transformer's use in intraday Forex is the study by Grądzki and Wójcik [13]. Their goal was to test whether the Transformer architecture could outperform modern deep models based on high-frequency currency data. The authors note that intraday trading in the Forex market using deep learning has not been studied enough, so they first identified the best benchmark (it turned out to be ResNet-LSTM), and then compared Transformer with it. The results showed that Transformer has a high predictive power based on Forex minute data and slightly surpasses the carefully tuned ResNet-LSTM in terms of forecast quality. This is an extremely important conclusion: even taking into account that the improvement was not radical, the fact that the attention model is superior confirms the prospects of this approach. Interestingly, Other authors also note the potential of Transformer models for high-frequency trading: for example, a decoder-free model trained on the EUR/USD and GBP/USD minute series reached a loss function (cross-entropy) value below 0.2, which indicates the possibility of building profitable strategies based on attention architectures [14]. In accordance with these observations, Fischer et al. (2024) note that Transformer is very suitable

for currency forecasting, however, in order to be confidently superior to LSTM, it requires the use of multiple input features (i.e., taking into account several simultaneously predicted time series and factors) [15]. Interestingly, in the work of Grądzki and Wójcik, the transformer performed better on larger timeframes (up to 12 hours), whereas there is too much noise in the very small data. In general, this study demonstrates the viability of Transformer for intraday currency forecasting and provides a basis for further improvements (for example, cost accounting in real trading – the authors also conducted a backtest of strategies and took into account commissions when evaluating the model).

Another notable trend is hybrid architectures with attention, combining different types of networks. For example, in the related field of cryptocurrencies, Peng et al. proposed the ACLMC model, which combines a convolutional network and an LSTM with an attention mechanism to predict trends of several cryptocurrencies at once at a high frequency [16]. The key idea was to use multi-frequency data and stabilize the target variable using a triple trend classification. As a result, the built-in attention helped the model extract useful information at different time scales and significantly reduced the number of false trades compared to traditional methods. This hybrid model has improved a number of financial quality metrics and surpassed several basic algorithms without attention. In addition, the use of improved attention mechanisms directly in recurrent layers is justified: for example, the integration of a special GRU-enhanced attention mechanism into the LSTM structure has improved the accuracy and stability of predictions in a volatile stock market [17]. Similar results show that the combination of CNN/LSTM with attention is able to simultaneously take into account short-term patterns and long-term dependencies, which is especially important for noisy intraday series.

Complex Transformer architectures are also emerging-derivatives adapted to time series. Zhao and colleagues conducted a detailed comparison of the classic Transformer with its Informer and Temporal Fusion Transformer (TFT) modifications on several currency pairs (for example, NZD/USD, NZD/CNY, etc.) [18]. The results showed that the most advanced TFT model gave the highest prediction accuracy (coefficient R^2 to 0.94) with minimal RMSE/MAE errors. Transformer was inferior in accuracy without improvements, but the alternative Informer model trained and converged much faster, albeit with a slight decrease in accuracy. Thus, there is a trade-off between the quality of the forecast and the computational efficiency of the attention models. In addition, Zhao et al. We found that the inclusion of additional market indicators, such as the VIX volatility index, improves the accuracy of TFT forecasts. This underscores the importance of a multimodal approach: attention mechanisms can easily be expanded to multiple data sources (news, volatility indicators, etc.), which has already been demonstrated in the latest models combining price series with text news signals and opens up opportunities for further improving the quality of intraday forecasts. More complex hybrid schemes based on Transformer have recently been proposed: for example, a dual attention mechanism, where one subnet processes price ranges with improved self-attention (with masking and multi-head), and a parallel subnet analyzes related factors through ConvLSTM+BiGRU with self-attention. Combining such networks has significantly improved forecasting accuracy and surpassed five modern models on a variety of financial datasets [19].

The use of attention layers on top of classic LSTMs is also being explored in the stock market. Thus, a modification of Att-LSTM – LSTM with an optimized attention mechanism was proposed for predicting stock prices, and an improvement in accuracy compared to conventional LSTM was shown [20]. Such results indirectly confirm the value of the attention mechanism for currency ranges as well: focusing on the most significant changes should enhance the model's ability to predict sharp intraday movements in the exchange rate.

Discussion and identified gaps

An analysis of modern literature shows that neural network approaches dominate in forecasting exchange rates, especially on high-frequency data. Classical RNN architectures (LSTM/GRU and their combinations) have been successfully used for short-term and intraday forecasts, surpassing traditional models in accuracy. At the same time, they are not without drawbacks: the limited ability to take into account long-term dependencies and important market features (shocks, news) stimulated

the introduction of attention mechanisms and completely new architectures (Transformers). Modern attention models have already shown competitiveness and sometimes superiority over the best recurrent networks, especially in the tasks of classifying the direction of price movement and detecting signals in noisy data. However, the literature on attention methods specifically for intraday Forex is still scarce. Only a few papers directly address this area, whereas most studies either focus on daily intervals or study related markets (stocks, cryptocurrencies). In particular, Grądzki and Wójcik emphasize that deep learning based on intraday Forex data remains an insufficiently researched area. Our review confirms this assessment: although some successes (for example, the use of Transformer) have already been achieved, the mechanisms of attention in the context of intraday currency forecasting are still not fully understood. It is not fully clear how best to integrate attention to account for cross-market effects, news flows, and how these models perform in real trading conditions with slippage and fees.

Thus, there is an obvious gap in research related to the in-depth use of attention mechanisms for high-frequency currency ranges. Filling this gap may include the development of specialized attention architectures or hybrids that take into account the specifics of Forex (for example, daily cycles of trading sessions, spikes in activity when macro statistics are released), and comparing them with already proven LSTM/GRU models. The conducted studies substantiate the prospects of such a path: they demonstrate the advantages of existing approaches and at the same time point out the limitations, overcoming which (through the mechanisms of attention) will be the task of further scientific work. This logically leads to the hypothesis of the current study, aimed at using attention mechanisms to improve the accuracy of the intraday forecast - not approving it in advance, but following from the analyzed achievements and gaps in the literature.

Materials and methods

In this study, historical intraday data of the EUR/USD currency pair with a minute frequency for the period from January 2020 to December 2022 were used to evaluate the effectiveness of neural network models. The data has been preprocessed using the following steps:

- ♦ Data resampling up to a minute interval (to detect gaps);
- ♦ Deleting weekend data (Saturdays and Sundays that appeared during resampling);
- ♦ Interpolation of missing values using the linear method and subsequent filling using the forward fill method.

The data was divided into three samples: training (up to and including December 2021), validation (from January to September 2022) and test (from October to December 2022). The signs for the models were the opening, maximum, minimum and closing prices. The closing price was chosen as the predicted indicator.

Model Architecture

The following neural network architectures were considered in the study:

- ♦ Basic LSTM: The architecture of the model consisted of two recurrent layers of the LSTM type (Long Short-Term Memory) with dimensions of 64 and 32 neurons, respectively. A 0.2 dropout was applied between the layers to prevent overfitting. The network output used a fully connected layer with a single neuron responsible for predicting the closing price.

- ♦ LSTM with the attention mechanism: This model used a single LSTM layer with a dimension of 64 neurons and a dropout level of 0.2. Next, the Bahdanau attention mechanism was used, which allowed the model to dynamically distribute the attention weights for each time step. This mechanism was implemented by two fully connected layers (32 neurons with the tanh activation function and 1 neuron for calculating scoring). The final forecast was formed by summing the LSTM outputs weighted by attention weights.

- ♦ Transformer (Encoder-only): This architecture was based entirely on the self-attention mechanism. It used a multi-headed attention with 4 heads and a representation dimension (d_{model}) of 64. LayerNormalization and positional embeddings were also included in the model. After the

self-attention layer, the Feed-Forward Neural Network block was used (128 neurons with ReLU activation function and additional dropout 0.2 regularization). The final forecast was carried out using a global average pooling layer (GlobalAveragePooling1D) and a fully connected layer with one neuron.

All models were trained using the standard error loss function (MSE) and the Adam optimizer. To prevent overfitting, early stopping (EarlyStopping) and maintaining the best model weights (ModelCheckpoint) were used. The quality of the models was assessed using the MSE, MAE, RMSE and MAPE metrics in a test sample.

Thus, the proposed methodology made it possible to conduct a systematic comparison of neural network architectures and identify the contribution of the attention mechanism to improving the accuracy of intraday price forecasting of the EUR/USD currency pair.

Results and discussion

Table 1 – Model metrics based on test data

Model	MSE	MAE	RMSE	MAPE
Basic LSTM	1.773636	1.158659	1.331779	32.238757
LSTM+Attention	0.616269	0.745160	0.785028	22.013667
Transformer	0.185482	0.297064	0.430676	7.299603

The results of the experiment show pronounced differences in the effectiveness of the considered neural network architectures. The best performance was demonstrated by the Transformer model, which has a self-attention mechanism, which significantly surpassed the other studied models in all evaluation metrics (Table 1). The mean square error (MSE) of the Transformer model was 0.185, which is significantly less than the LSTM+Attention model (MSE=0.616) and the basic LSTM (MSE=1.774). A similar trend is observed for other key metrics: Transformer showed minimal average absolute errors (MAE=0.297) and standard deviation (RMSE=0.431), as well as the smallest average absolute percentage error (MAPE=7.3%).

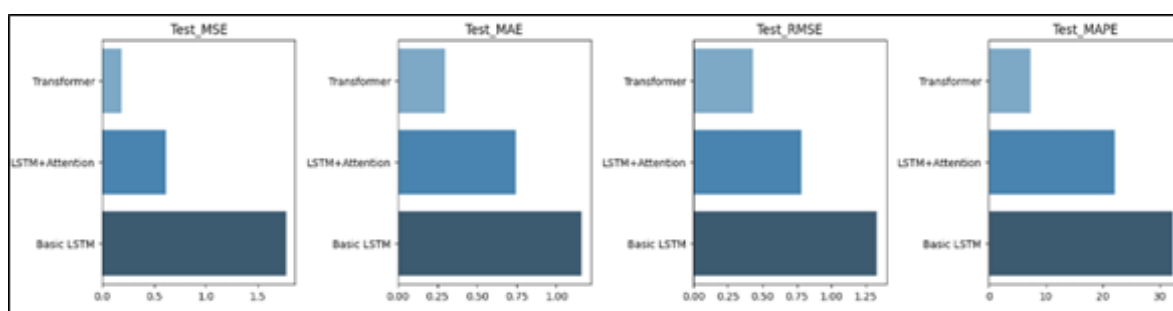


Figure 1 – Comparison of model indicators based on test data

The attentional LSTM model also significantly improved the prediction quality compared to the basic LSTM, demonstrating a significant reduction in errors (MSE=0.616, MAE=0.745, RMSE=0.785, and MAPE=22.01%). The basic LSTM model, on the contrary, performed worse (Figure 1), which indicates its insufficient ability to take into account complex and rapidly changing market conditions.

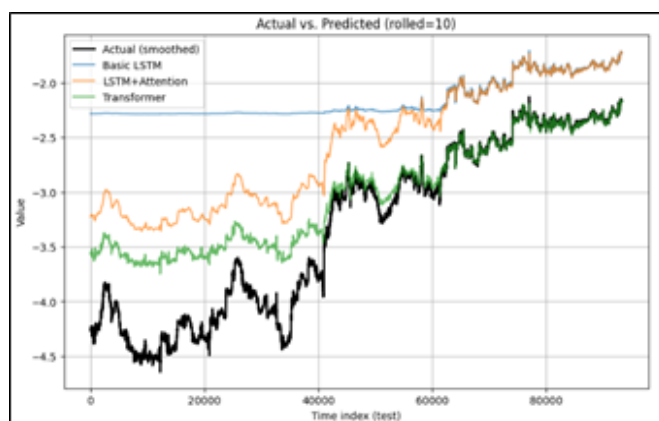


Figure 2 – Model forecasts

An additional analysis of the visual graphs (Figure 2) of the real and predicted values confirmed the numerical results. The predictions of the Transformer model are as close as possible to the real values, while the remnants of its predictions are minimal and do not show a pronounced systematic error (Figure 3). The LSTM model with the attention mechanism also shows stable behavior, but its deviations from real values are noticeably greater compared to Transformer. At the same time, the basic LSTM model systematically overestimates forecasts, as can be seen from both the remainder graph and the error distribution.

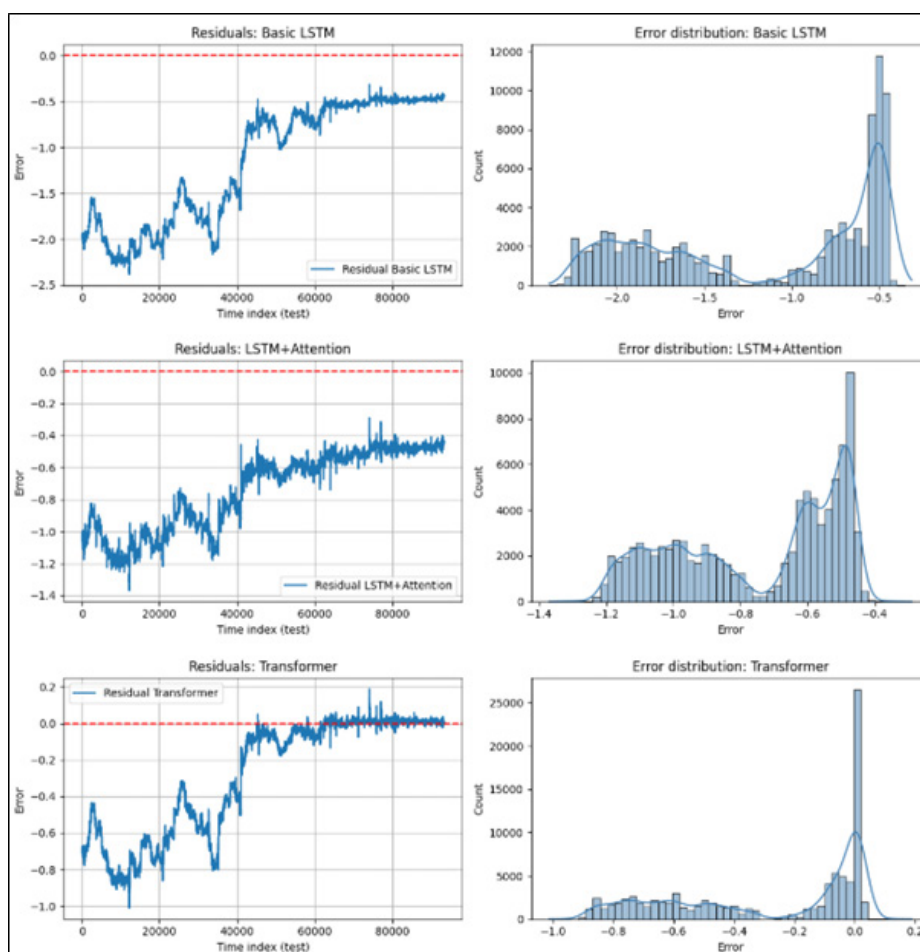


Figure 3 – Error distribution and residuals

The results obtained suggest that the attention mechanisms, in particular, the Transformer architecture, have a high potential for use in the tasks of intraday forecasting of exchange rates. The strong superiority of the Transformer model over the basic LSTM underscores the importance of dynamically managing attention to various segments of historical data and confirms the relevance of adaptive neural network architectures.

It is worth noting that even the LSTM model, complemented by a simple Bahdanau attention mechanism, is significantly superior to the classical recurrent architecture without attention, which confirms the hypothesis of the study on the expediency of using attention mechanisms in forecasting high-frequency financial series.

Nevertheless, the success of the Transformer model is due not only to the attention mechanism itself, but also to the complexity of the architecture, which requires significant computing resources and careful tuning of hyperparameters. This opens up opportunities for further research aimed at optimizing and adapting attention models for practical use in trading. It is also important to take into account that the results may vary depending on currency pairs and time intervals, which requires additional testing and analysis of the resilience of models to changing conditions.

Conclusion

As a result of the research, it was proved that the use of neural network models with an attention mechanism is a promising direction for the tasks of intraday forecasting of exchange rates. The Transformer model, which implements the self-attention mechanism, demonstrated the best results for all the studied metrics based on the data of the EUR/USD currency pair.

The results obtained confirm the hypothesis that attention mechanisms can significantly improve the accuracy of short-term forecasting based on high-frequency data, opening up prospects for their wide application in algorithmic trading. At the same time, it should be borne in mind that the success of such models depends on their computational efficiency and hyperparameter settings. Future research may be aimed at expanding the number of currency pairs under consideration, exploring various attention architectures and developing methods for their effective configuration.

REFERENCES

- 1 Dakalbab, F., Kumar, A., Abu Talib, M., and Nasir, Q. Advancing Forex prediction through multimodal text-driven model and attention mechanisms. *Intelligent Systems with Applications*, 23, 200518 (2025). <https://doi.org/10.1016/j.iswa.2025.200518>
- 2 Yildirim, D.C., Toroslul, I.H., and Fiore, U. Forecasting directional movement of Forex data using LSTM with technical and macroeconomic indicators. *Financial Innovation*, 7 (1), 1–21 (2021). <https://doi.org/10.1186/s40854-020-00220-2>
- 3 Saghafi, A., Bagherian, M., and Shokoohi, F. Forecasting Forex EUR/USD closing prices using a dual-input deep learning model with technical and fundamental indicators. *Mathematics*, 13 (9), 1472 (2025).
- 4 García, F., Martínez-César, S., Pérez-Rosés, H., et al. Foreign exchange forecasting models: LSTM and BiLSTM comparison. *Engineering Proceedings*, 68 (1), 19 (2024).
- 5 Islam, M.S., and Hossain, E. Foreign exchange currency rate prediction using a GRU–LSTM hybrid network. *Soft Computing Letters*, 3, 100009 (2021).
- 6 Ayitey Junior, M., Appiahene, P., and Appiah, O. Forex market forecasting with two-layer stacked long short-term memory neural network (LSTM) and correlation analysis. *Journal of Electrical Systems and Information Technology*, 9 (1), 14 (2022).
- 7 Bormpotsis, C., Sedky, M., and Patel, A. Predicting Forex currency fluctuations using a novel bio-inspired modular neural network. *Big Data and Cognitive Computing*, 7 (3), 152 (2023).
- 8 Markova, M. Forex time series forecasting using hybrid convolutional neural network/long short-term memory network model. In: *International Conference on New Trends in the Applications of Differential Equations in Sciences*. Cham: Springer International Publishing, 2022, pp. 295–305.

- 9 Shi, H., Li, Y., Wang, Z., et al. A CNN–LSTM based deep learning model with high accuracy and robustness for carbon price forecasting: A case of Shenzhen's carbon market in China. *Journal of Environmental Management*, 352, 120131 (2024).
- 10 Pornwattanavichai, A., Maneeroj, S., and Boonsiri, S. BERTFOREX: Cascading model for Forex market forecasting using fundamental and technical indicator data based on BERT. *IEEE Access*, 10, 23425–23437 (2022).
- 11 Ito, K., Iima, H., and Kitamura, Y. LSTM forecasting foreign exchange rates using limit order book. *Finance Research Letters*, 47, 102517 (2022).
- 12 Zafeiriou, T., and Kalles, D. Comparative analysis of neural network architectures for short-term FOREX forecasting. *Intelligent Decision Technologies*, 18 (3), 2177–2190 (2024).
- 13 Gradzki, P., and Wojcik, P. Is attention all you need for intraday Forex trading? *Expert Systems*, 41 (2), e13317 (2024).
- 14 Kantoutsis, K.T., Mavrogianni, A.N., and Theodorakatos, N.P. Transformers in high-frequency trading. *Journal of Physics: Conference Series*, 2701 (1), 012134 (2024).
- 15 Fischer, T., Sterling, M., and Lessmann, S. FX-spot predictions with state-of-the-art transformer and time embeddings. *Expert Systems with Applications*, 249, 123538 (2024).
- 16 Peng, P., Liu, X., Zhang, Y., et al. Attention-based CNN–LSTM for high-frequency multiple cryptocurrency trend prediction. *Expert Systems with Applications*, 237, 121520 (2024).
- 17 Zhang, J., and Yan, L. GRU-enhanced attention mechanism for LSTM in hybrid CNN–LSTM models for stock prediction. *Journal of Global Trends in Social Science*, 2 (3), 10–17 (2025). <https://doi.org/10.70731/rzvs8j53>
- 18 Zhao, L., and Yan, W.Q. Prediction of currency exchange rate based on transformers. *Journal of Risk and Financial Management*, 17 (8), 332 (2024).
- 19 Hadizadeh, A., Tarokh, M.J., and Ghazani, M.M. A novel transformer-based dual attention architecture for the prediction of financial time series. *Journal of King Saud University – Computer and Information Sciences*, 37, 72 (2025). <https://doi.org/10.1007/s44443-025-00045-y>
- 20 Liu, Q., Hu, Y., and Liu, H. Enhanced stock price prediction with optimized ensemble modeling using multi-source heterogeneous data: Integrating LSTM attention mechanism and multidimensional gray model. *Journal of Industrial Information Integration*, 42, 100711 (2024).

¹Абдилдаева А.А.,

PhD, профессор м.а., ORCID: 0000-0002-6381-9350,

e-mail: assel.abdildaeva@kaznu.edu.kz

^{1*}Нуртуган Г.Б.,

докторант, ORCID: 0000-0001-8824-9502,

*e-mail: g.nurtugun@gmail.com

²Войткевич К.,

PhD, доцент, ORCID: 0000-0002-7851-4330,

e-mail: krystian.wojtkiewicz@pwr.edu.pl

¹Әл-Фараби атындағы Қазақ ұлттық университеті, Алматы қ., Қазақстан

²Вроцлав ғылым және технологиялар университеті (WUST), Вроцлав қ., Польша

EUR/USD ЖҰБЫ БОЙЫНША ТӘУЛІКТІК БАҒА БОЛЖАУҒА АРНАЛҒАН НАЗАР МЕХАНИЗМДЕРІ БАР НЕЙРОНДЫҚ ЖЕЛІ МОДЕЛЬДЕРІН ДАМУ ЖӘНЕ ОҢТАЙЛАНДЫРУ

Аңдатпа

Зерттеу әртүрлі нейрондық желілік архитектураларды, атап айтқанда назар аудару механизмдерін біріктіретін модельдерді (attention) пайдалана отырып, EUR/USD валюта жұбының бағамын бір күндік болжау мәселесін қарастырады. Үш нейрондық желілік архитектура зерттелді: LSTM базалық моделі, бағдану зейін механизмі бар LSTM моделі және өзін-өзі зейіндеу механизмі бар трансформатор моделі (self-attention). Эксперимент 2020 жылдың қаңтарынан 2022 жылдың желтоқсанына дейінгі кезеңдегі

тарихи минуттық деректерде жүргізілді. Нәтижелер назар аудару механизмі бар модельдер LSTM негізгі архитектурасынан айтарлықтай жоғары екенін көрсетті. Ең жақсы нәтижелер Transformer моделімен алынды ($MSE=0.185$, $MAE=0.297$, $RMSE=0.431$, $MAPE=7.3\%$). Егжей-тегжейлі талдау transformer моделінің тұрақтылығы мен дәлдігін растады. Attention модельдерінің анықталған артықшылықтары олардың алгоритмдік саудада қолдану перспективасын негіздейді және нақты сауда жағдайларын оңтайландыру және бейімдеу үшін қосымша зерттеулерді қажет етеді. Атап айтқанда, қосымша зерттеулер attention модельдерін сауда стратегиялары мен тәуекелдерді басқару жүйелерімен біріктіруге, сондай-ақ нарықтық құбылмалылықтың күрт өзгеруі жағдайында олардың мінез-құлқын зерттеуге бағытталуы мүмкін. Сонымен қатар, практикалық саудадағы болжамдардың жалпы тұрақтылығы мен сенімділігін арттыру үшін attention архитектураларын басқа болжау әдістерімен біріктіру мүмкіндіктерін зерттеу ұсынылады.

Тірек сөздер: тәуліктік болжау, Forex, айырбас бағамдары, EUR/USD, нейрондық желілер, назар механизмі, Transformer, LSTM, өзіндік назар (self-attention).

¹**Абдилдаева А.А.,**

PhD, и.о. профессора, ORCID: 0000-0002-6381-9350,

e-mail: assel.abdildaeva@kaznu.edu.kz

^{1*}**Нуртуган Г.Б.,**

докторант, ORCID: 0000-0001-8824-9502,

*e-mail: g.nurtugun@gmail.com

²**Войткевич К.,**

PhD, доцент, ORCID: 0000-0002-7851-4330,

e-mail: krystian.wojtkiewicz@pwr.edu.pl

¹Казахский национальный университет им. аль-Фараби, г. Алматы, Казахстан

²Вроцлавский университет науки и технологий (WUST), г. Вроцлав, Польша

РАЗРАБОТКА И ОПТИМИЗАЦИЯ НЕЙРОННЫХ СЕТЕВЫХ МОДЕЛЕЙ С МЕХАНИЗМАМИ ВНИМАНИЯ ДЛЯ ВНУТРИДНЕВНОГО ПРОГНОЗИРОВАНИЯ ЦЕНЫ ПАРЫ EUR/USD

Аннотация

В исследовании рассматривается проблема внутридневного прогнозирования курса валютной пары EUR/USD с использованием различных нейросетевых архитектур, в частности моделей, интегрирующих механизмы внимания (attention). Были исследованы три нейросетевые архитектуры: базовая модель LSTM, модель LSTM с механизмом внимания Bahdanau и модель Transformer с механизмом самовнимания (self-attention). Эксперимент проводился на исторических минутных данных за период с января 2020 по декабрь 2022 гг. Результаты показали, что модели с механизмом внимания значительно превосходят базовую архитектуру LSTM. Наилучшие результаты были получены моделью Transformer ($MSE=0.185$, $MAE=0.297$, $RMSE=0.431$, $MAPE=7.3\%$). Подробный анализ подтвердил стабильность и точность модели Transformer. Выявленные преимущества attention-моделей обосновывают их перспективность для применения в алгоритмической торговле и требуют дальнейших исследований для оптимизации и адаптации к реальным торговым условиям. В частности, дальнейшие исследования могут быть направлены на интеграцию attention-моделей с торговыми стратегиями и системами управления рисками, а также изучение их поведения в условиях резких изменений рыночной волатильности. Кроме того, предлагается исследовать возможности комбинирования attention-архитектур с другими методами прогнозирования для повышения общей устойчивости и надежности прогнозов в практическом трейдинге.

Ключевые слова: внутридневное прогнозирование, Forex, валютные курсы, EUR/USD, нейронные сети, механизм внимания, Transformer, LSTM, самовнимание.

Article submission date: 30.06.2025

УДК 004.056
МРНТИ 81.93.29

<https://doi.org/10.55452/1998-6688-2025-22-4-107-118>

¹*Самуйлова А.,
магистрант, ORCID ID: 0009-0003-5806-8542,
*e-mail: anastassiyasamuilova@gmail.com

¹Казахский национальный университет им. аль-Фараби, г. Алматы, Казахстан

СРЕДСТВА ВЫЯВЛЕНИЯ УЯЗВИМОСТЕЙ ИНФОРМАЦИОННОЙ БЕЗОПАСНОСТИ НА ОСНОВАНИИ ДАННЫХ ТЕМАТИЧЕСКИХ ИНТЕРНЕТ-РЕСУРСОВ

Аннотация

По мере усложнения киберугроз традиционные методы обнаружения уязвимостей теряют эффективность. Цель данной работы – разработка и апробация подхода к выявлению уязвимостей на основе анализа данных с тематических интернет-ресурсов: форумов, блогов и социальных сетей. Эти источники содержат большое количество неструктурированной информации, что требует применения методов интеллектуального анализа данных. В работе используется интеграция современных технологий: предобученной языковой модели SecBERT (Security Bidirectional Encoder Representations from Transformers), предназначенной для задач в области кибербезопасности, и адаптивной нейро-нечеткой системы вывода DENFIS (Dynamic Evolving Neural-Fuzzy Inference System). Предлагаемая система позволяет фильтровать нерелевантные сообщения, выделять индикаторы компрометации и потенциальные угрозы. Применение нечеткой логики дает возможность эффективно обрабатывать неопределенную и неполную информацию. Проведенные эксперименты подтвердили высокую точность классификации и устойчивость нечеткой кластеризации ($FPC = 0,93$; $PE = 0,28$; $XB = 0,042$). Система продемонстрировала способность к своевременному выявлению признаков киберугроз и обладает потенциалом масштабирования для задач мониторинга и предсказания атак. Результаты свидетельствуют о его потенциале в повышении скорости реагирования на киберугрозы и усилении защиты информационных систем.

Ключевые слова: уязвимости информационной безопасности, тематические интернет-ресурсы, трансформеры, SecBERT, нейро-нечеткая логика.

Введение

В современных условиях стремительного развития цифровых технологий и роста объемов информации увеличивается вероятность кибератак и выявления новых уязвимостей в информационных системах. Традиционные методы обнаружения угроз, такие как сигнатурный анализ, поведенческий анализ, анализ аномалий и статистические подходы, обладают существенными ограничениями, поскольку они ориентированы на известные шаблоны атак или аномалии, выявленные в прошлом [1]. Однако учитывая сложность и динамичность современных угроз, требуется более гибкий и адаптивный подход, способный эффективно анализировать разнородные источники информации, включая специализированные интернет-ресурсы, хакерские форумы, базы данных уязвимостей и новостные агрегаторы [2].

На сегодняшний день ключевыми источниками данных о киберугрозах остаются такие базы, как NVD или CVE, а также отчеты государственных и частных организаций, где сведения структурируются экспертами и исследовательскими центрами [3]. Однако эти источники имеют очевидный недостаток: процесс верификации и внесения данных может занимать значительное время, в результате чего появляется временной лаг между выявлением угрозы и ее официальной регистрацией [4]. В противоположность этому специализированные форумы и анонимные площадки предоставляют информацию о новых уязвимостях и потенциальных

векторах атак практически в режиме реального времени [5]. Их высокая популярность среди специалистов в области информационной безопасности делает такие ресурсы ценным источником для мониторинга и анализа киберугроз, позволяя фиксировать и прогнозировать тенденции еще до их официального признания [6].

Актуальность данного исследования обусловлена необходимостью создания автоматизированного подхода, который позволит эффективно извлекать полезную информацию о киберугрозах из открытых источников [5]. Традиционные методы обнаружения уязвимостей, основанные на статическом анализе данных, не способны оперативно реагировать на появление новых типов атак, в то время как интеграция технологий трансформеров и нейро-нечеткой логики может существенно повысить качество прогнозирования и выявления потенциальных угроз [2].

Объектом исследования выступают информационные системы, функционирующие в условиях активного использования интернет-ресурсов. Предметом – методы автоматического выявления уязвимостей с помощью интеллектуального анализа данных. Основная цель работы – разработать комплексный метод, который позволит повысить точность и оперативность обнаружения киберугроз и потенциальных рисков.

Для реализации данной цели необходимо решить следующие задачи:

- ♦ проанализировать существующие методы выявления уязвимостей и выявить их ограничения в условиях постоянно меняющейся киберугрозы;
- ♦ разработать методику, сочетающую трансформеры для извлечения семантических признаков и нейро-нечеткую логику для работы с неопределенностью в данных;
- ♦ провести эксперименты на реальных данных, собранных с форумов и баз уязвимостей;
- ♦ провести сравнительный анализ предложенного метода с традиционными подходами, такими как сигнатурный анализ, статистический анализ и методы машинного обучения.

Методологическая основа исследования включает в себя методы глубокого обучения, применяемые для классификации и категоризации сообщений об уязвимостях, нейро-нечеткую логику, обеспечивающую обработку неопределенных данных и вероятностную оценку угроз, а также статистический анализ, позволяющий выявлять закономерности в распространении уязвимостей и их потенциальное влияние на системы [5].

Современные исследования в области анализа киберугроз демонстрируют эффективность методов машинного обучения и глубокого анализа текста в выявлении потенциальных угроз [2]. В работах [1] и [5] показано, что применение нейросетевых моделей для анализа сообщений на хакерских форумах позволяет выявлять новые уязвимости и векторы атак задолго до их эксплуатации злоумышленниками. Одним из ключевых направлений исследований является анализ индикаторов компрометации (IoC), который используется в платформах AlienVault OTX, VirusTotal и MISP для сопоставления выявленных угроз с реальными атаками [7].

Особенно перспективным направлением считается использование трансформеров, таких как BERT, и его специализированных версий – SecBERT, DarkBERT. Эти модели показали высокую эффективность в обработке естественного языка в контексте задач кибербезопасности [8]. Исследования [9] также подтверждают, что механизмы внимания (attention mechanism) значительно повышают точность извлечения ключевых элементов из текстов, что критически важно при анализе форумов с неструктурированной информацией. В статье [10] представлена модель SecureBERT 2.0, обученная на обширном и тщательно подобранном корпусе данных, включающем более 13,6 млрд текстовых токенов и 53,3 млн токенов кода. Корпус включает публичные отчеты о безопасности, уведомления об уязвимостях, открытые статьи по кибербезопасности, технические книги и рецензируемые научные работы, что позволяет модели получать богатое контекстное представление о кибербезопасности в различных поддоменах, включая сетевую архитектуру, управление доступом, механизмы аутентификации и криптографические протоколы. Модель достигает F1-score 0.945 с высоким recall (0.965, доля

корректно выявленных сущностей) и точностью (0.927, доля правильно классифицированных положительных объектов), что позволяет эффективно выявлять киберугрозы и ключевые сущности при минимуме ложноположительных срабатываний. В статье [11] представлена модель DarkBERT, обученная на текстах Даркнета из датасетов DUTA-10K и CoDA, включающих страницы с onion-доменов, HTML-теги title и body, а также различные активности пользователей Даркнета. Модель использует архитектуру RoBERTa и обучалась как на сырых, так и на предобработанных данных. На датасете DUTA F1-score составляет 80.01 на сырых данных и 79.98 на предобработанных, с precision и recall около 80, а на CoDA достигает 94.25 на сырых и 94.42 на предобработанных данных.

Однако несмотря на успехи глубокого обучения, оно остается недостаточным для полноценной работы с динамически изменяющимися угрозами. В этой связи в ряде исследований предлагается интеграция с другими методами. Так, в обзоре [12] авторы рассматривают применение больших языковых моделей (LLM) в кибербезопасности и отмечают важность решения вопросов этики, объяснимости моделей и их интеграции с другими методами, а для повышения эффективности предлагают сотрудничество с экспертами по когнитивным наукам и психологии. Помимо этого, для работы с неопределенностью и динамичностью угроз в ряде исследований предлагается использование нейро-нечеткой логики [13]. Подходы, основанные на адаптивных нейро-нечетких системах вывода (ANFIS, DENFIS), позволяют повысить точность прогнозирования угроз без необходимости полной переобучаемости модели. Это особенно актуально в условиях постоянной смены тактик атак и появления новых типов уязвимостей [14].

Материалы и методы

В данном исследовании применяется интегрированный подход, основанный на сборе, обработке и анализе данных с тематических интернет-ресурсов с использованием методов веб-скрапинга, предобученных моделей и нейро-нечеткой логики, согласно рисунку 1.



Рисунок 1 – Архитектура разрабатываемой модели

На первом этапе осуществляется парсинг данных с хакерских форумов и других специализированных интернет-ресурсов. Используется Python-библиотека BeautifulSoup, Requests, Praw для извлечения постов из Reddit и Sqlite3 для динамического сбора данных с сайтов, требующих взаимодействия пользователя. Собранные данные сохраняются в SQL-базе данных для последующей обработки. Структура базы данных показана на рисунке 2.

На втором этапе осуществляется предобработка текстовой информации, включая очистку данных, удаление дубликатов и фильтрацию нерелевантных сообщений. Применяется лемма-

тизация и удаление стоп-слов с использованием библиотек NLTK и spaCy. Данные подвергаются первичной фильтрации с использованием ключевых слов, связанных с информационной безопасностью, таких как exploit, vulnerability, malware.



Рисунок 2 – Структура базы данных

Для классификации сообщений и определения релевантности используется предобученная модель трансформеров SecBERT. SecBERT – это языковая модель на основе архитектуры BERT, показанной на рисунке 3, адаптированная для обработки текстов в сфере кибербезопасности. Он использует механизм самовнимания, который позволяет учитывать контекстные зависимости между словами и выявлять значимые элементы в тексте. Модель принимает на вход токенизированный текст и пропускает его через слои трансформеров, после чего формируются эмбединги слов. Для представления входных данных используется векторное представление токенов $X = [x_1, x_2, \dots, x_i]$, где x_i – это эмбединг i -го токена. Эти эмбединги передаются через слои самовнимания, определяемые следующим уравнением:

$$A = \text{softmax} \left(\frac{QK^T}{\sqrt{d_k}} \right) * V \quad (1)$$

где Q , K , V – это матрицы запросов, ключей и значений, а d_k – размерность ключей. Итоговый скрытый слой H формируется после нескольких таких операций, проходя через механизмы нормализации и активации. Для классификации угроз выходное представление H передается в полносвязный слой с функцией активации softmax, который вычисляет вероятность принадлежности сообщения к категории угроз.

Модель была обучена на специализированном корпусе, включающем такие источники, как APTnotes, Stucco-Data, CASIE, а также данные из задачи SemEval-2018 Task 8 (SecureNLP). Для повышения точности работы с текстами данной тематики в модели используется специализированный словарь (secvocab), сформированный на основе терминологии, характерной для кибербезопасности. Помимо стандартной версии SecBERT, существуют также ее модификации, такие как SecRoBERTa, созданные на основе архитектуры RoBERTa и обладающие схожими возможностями. Для решения задачи классификации в данной работе модель была дообучена на размеченном корпусе, содержащем 41 тысячу текстов, относящихся и не относящихся к области информационной безопасности, и продемонстрировала точность 98–99% на данном корпусе.

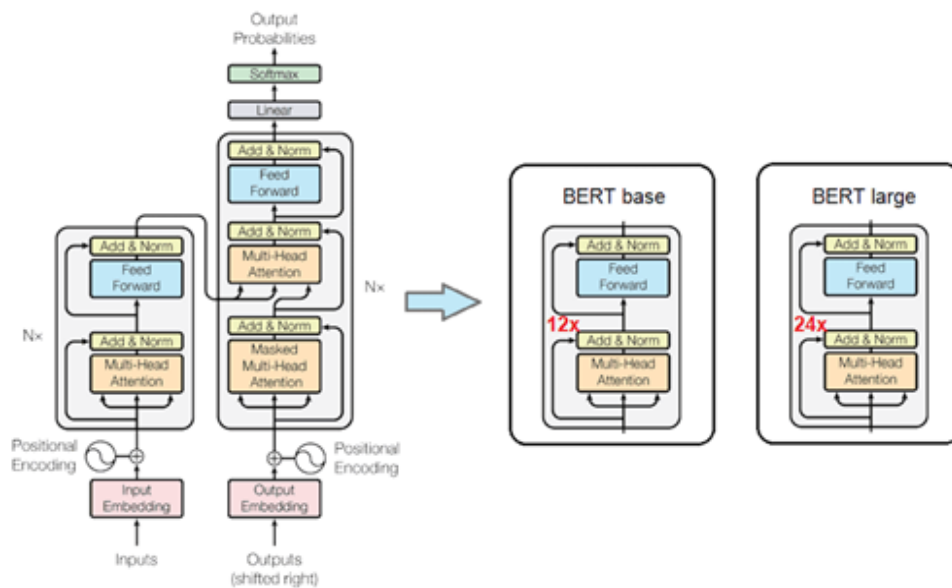


Рисунок 3 – Архитектура BERT модели

Отфильтрованные данные сохраняются в отдельную базу данных как возможные прецеденты и затем передаются в систему нейро-нечеткой логики для дополнительного анализа. Нечеткая логика представляет собой метод обработки информации, позволяющий моделировать неопределенность и приближенные значения, в отличие от классической логики, где переменные принимают строго определенные значения [15]. В данной работе применяется адаптивная нейро-нечеткая система вывода DENFIS, которая сочетает элементы нечеткой логики и нейросетевых алгоритмов, динамически адаптируясь к изменениям входных данных. Архитектура данной системы представлена на рисунке 4 и отражает процесс формирования правил на основе входных данных в реальном времени с использованием механизма эволюционного обучения [15]. Примеры правил показаны в таблице 1.

Таблица 1 – Правила нечеткой логики

1	Если количество комментариев на форуме высокое и уровень рейтинга авторов высокий, то вероятность возникновения угрозы информационной безопасности высокая.
2	Если количество комментариев на форуме высокое и уровень рейтинга авторов средний, то вероятность возникновения угрозы информационной безопасности средняя.
3	Если количество комментариев на форуме высокое и уровень рейтинга авторов низкий, то вероятность возникновения угрозы информационной безопасности низкая.
4	Если количество комментариев на форуме среднее и уровень рейтинга авторов высокий, то вероятность возникновения угрозы информационной безопасности высокая.
5	Если количество комментариев на форуме среднее и уровень рейтинга авторов средний, то вероятность возникновения угрозы информационной безопасности средняя.
6	Если количество комментариев на форуме среднее и уровень рейтинга авторов низкий, то вероятность возникновения угрозы информационной безопасности низкая.
7	Если количество комментариев на форуме низкое и уровень рейтинга авторов высокий, то вероятность возникновения угрозы информационной безопасности средняя.
8	Если количество комментариев на форуме низкое и уровень рейтинга авторов средний, то вероятность возникновения угрозы информационной безопасности низкая.
9	Если количество комментариев на форуме низкое и уровень рейтинга авторов низкий, то вероятность возникновения угрозы информационной безопасности низкая.

Принцип работы DENFIS основан на том, что входные данные оцениваются через нечеткие функции принадлежности (например, гауссовых или треугольных), которые определяют степень соответствия каждого параметра правилам, аналогичным приведенным в таблице 1. Каждое значение, например рейтинг автора или количество комментариев, проверяется относительно этих функций, и вычисляется степень принадлежности к каждому возможному условию правила. Затем минимальное расстояние между входными данными и центрами кластеров используется для активации наиболее релевантных правил и формирования итоговой оценки уровня угрозы. Входные параметры включают уровень достоверности сообщения, рейтинг автора и уровень обсуждаемости темы, то есть количество комментариев и активность пользователей. Выходной параметр представляет собой оценку уровня угрозы, включающую низкий, средний и высокий уровни.

Система адаптивна: при поступлении новых данных DENFIS корректирует свои правила функции принадлежности и обновляет центры кластеров, автоматически уточняя условия существующих правил и создавая новые комбинации характеристик постов.

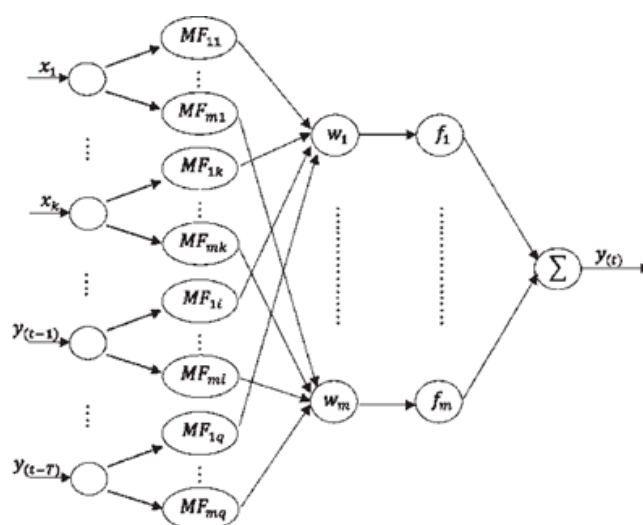


Рисунок 4 – Архитектура DENFIS

Полученные результаты визуализируются и представляются в виде отчетов, содержащих информацию о выявленных угрозах, их классификации и вероятностной оценке значимости. Данные могут быть использованы для автоматизированного мониторинга угроз в реальном времени и принятия решений по реагированию на потенциальные кибератаки. Таким образом, предложенная методика сочетает в себе веб-скрапинг, обработку данных, анализ с использованием предобученной модели и нейро-нечеткую логику, что позволяет эффективно выявлять новые угрозы и обеспечивать их классификацию с учетом вероятностной оценки риска.

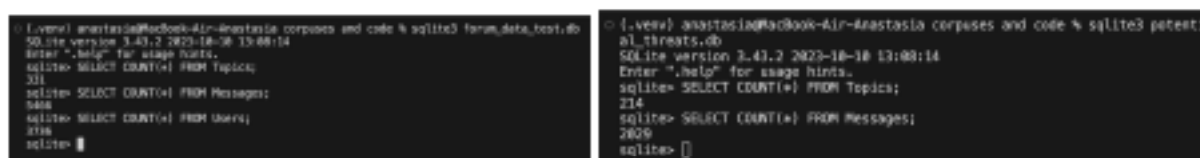
Результаты и обсуждение

Гипотеза данного исследования заключалась в том, что тематические интернет-ресурсы могут служить достоверным источником данных для автоматизированного выявления уязвимостей информационной безопасности при условии применения современных методов анализа неструктурированной информации. В ходе эксперимента был реализован интегрированный подход, сочетающий предобученную трансформерную модель SecBERT и нейро-нечеткую систему DENFIS, что позволило эффективно классифицировать посты и формировать правила на основе кластеризации. Для оценки точности кластеризации использовались метрики FPC, XB и PE: FPC (Fuzzy Partition Coefficient) отражает степень четкости принадлежности объектов к кластерам, XB (Xie-Beni index) оценивает компактность и разделимость кластеров, а PE (Partition Entropy) измеряет неопределенность распределения объектов по кластерам.

Собранный датасет включал более 300 тем, согласно рисунку 5a, и свыше 5000 комментариев с шести доступных форумов, включая Reddit, HackerNews, Antichat и других специализированных ресурсов, посвященных вопросам информационной безопасности, и использовался как тестовый корпус для проверки работы модели. Однако процесс веб-скрапинга сопровождался рядом ограничений, так как некоторые форумы были заблокированы государственными учреждениями, а другие требовали авторизацию для просмотра содержимого, что ограничило доступный объем данных.

Предварительно обученная модель SecBERT успешно выделила 214 релевантных тем, согласно рисунку 5b, соответствующих признакам потенциальных угроз. Пример таких релевантных постов представлен на рисунке 6. Далее с применением алгоритма fuzzy c-means были выделены три устойчивых кластера и рассчитаны их центры для дальнейшей классификации собранных данных.

В ходе работы также был рассмотрен альтернативный метод классификации данных на основе онтологий, как показано в исследовании [16]. Однако данный подход обладает рядом существенных ограничений. Онтологическая модель плохо справляется с новыми угрозами, так как требует заранее заданной структуры знаний. Точность классификации напрямую зависит от полноты онтологии и ее соответствия современным угрозам. Онтологии требуют постоянного пополнения и ручного контроля, что делает их менее гибкими в условиях динамически меняющейся киберугрозы. Кроме того, онтологии строятся на фиксированных взаимосвязях между понятиями, что может привести к ошибкам классификации при наличии неоднозначных или неизвестных терминов. Также процесс разработки и поддержки онтологии требует значительных затрат ресурсов, в отличие от машинного обучения, где модели могут автоматически подстраиваться под новые угрозы.



a) До фильтрации

b) После фильтрации

Рисунок 5 – Результаты фильтрации при помощи SecBERT

На последнем этапе данные передаются в DENFIS для определения уровня угрозы обсуждений. Сначала происходит подключение к базе данных, содержащей информацию о потенциальных инцидентах информационной безопасности. С помощью SQL-запроса из таблиц с постами и пользователями извлекаются необходимые данные: идентификатор и название темы, рейтинг автора, количество сообщений, время создания поста и имя автора. После извлечения информация преобразуется в удобный для анализа формат – двумерный массив, в котором в качестве признаков используются только числовые значения: рейтинг автора и количество сообщений. Это позволяет эффективно применить алгоритм кластеризации для формирования исходных правил для дальнейшей работы нейро-нечеткой модели.

Алгоритм fuzzy c-means вычисляет центры кластеров и матрицу принадлежности объектов к кластерам. Полученные результаты выводятся в консоль и визуализируются на графике, как показано на рисунке 7: объекты, принадлежащие к каждому кластеру, отображаются разными цветами, а центры кластеров обозначаются крестиками.

В результате работы алгоритма получилось три центра кластеров: (3.98824988e+05, 1.00000024e+00), (1.71489384e+03, 6.88635187e+01) и (1.58592931e+04, 1.75610911e+00). Каждый из этих центров отражает характерные группы объектов, отличающиеся, в частности, по масштабу первого признака и умеренно варьирующиеся по второму.

```

476[1]A new vulnerability on IPv6 parsing in linux|5|351
8|3521|2024-11-12T23:54:51Z
478[1]New Vulnerability Found in Apps Using Wi-Fi|17|352
4|3530|2013-10-29T12:07:24Z
480[1]Intel Downfall: New Vulnerability Affecting AVX2/A
VX-512|1|3532|3533|2023-08-08T17:06:24Z
481[1]New WiFi Vulnerability: The SSID Confusion Attack|
8|3534|3537|2024-05-15T04:50:01Z
482[1]Schneier: Details on a New FGP Vulnerability|2|353
8|3540|2018-05-17T15:51:17Z
483[1]The Accidental Discovery of a New Vulnerability in
Google's OAuth Implementation [video] (2023)|3|3541|3544
|2024-03-12T15:43:22Z
491[1]Ed Felten & Team Disclose 4 New CSRF Vulnerabiliti
es, Can Transfer Funds From ING|1|47|47|2008-09-20T21:10
:14Z
492[1]SaltStack reveals new critical vulnerabilities, pa
tch now|1|3558|3558|2020-11-03T19:39:57Z
493[1]Disclosure of three 0-day iOS vulnerabilities|456|
3559|3612|2021-09-24T00:20:01Z
494[1]Vulnerability in the Mac Zoom client allows malici
ous websites to enable camera|456|3613|3720|2019-07-08T2
2:17:47Z
495[1]Exploiting vulnerabilities in Cellebrite UFED and
Physical Analyzer|332|3721|3766|2021-04-21T16:20:45Z
496[1]Miracles: An Apple M1 covert channel vulnerability
|277|3767|3813|2021-05-26T03:02:00Z

```

Рисунок 6 – Пример релевантных постов

Суммарно все три кластера, выделенные в ходе анализа, демонстрируют устойчивость и обоснованность: высокое значение FPC (0.93), низкий PE (0.28) и низкий XB (0.042) показывают, что выделенные группы действительно отражают различные типы обсуждений, подтвердив высокую четкость и достоверность кластеризации. Первый кластер при этом характеризуется очень крупным значением первого признака и небольшим вторым, второй кластер включает объекты со значительно меньшим первым признаком и более высокой величиной второго, а третий занимает промежуточное положение по первому признаку с незначительно возросшим вторым.

На основании данной кластеризации были сформированы изначальные правила для последующего применения алгоритма DENFIS, при котором каждому кластеру сопоставляется соответствующая метка, а сами метки затем учитываются в процессе нечеткого неструктурированного моделирования.

После этапа кластеризации данные проходят дополнительную обработку, цель которой – вычисление совокупного балла поста, отражающего его значимость и потенциальный уровень угрозы. Для этого из каждого поста извлекаются два ключевых показателя: рейтинг автора и число сообщений в теме. Чтобы сделать их сопоставимыми, применяется нормализация. Количество сообщений нормализуется линейно, а рейтинг – с помощью логарифмической трансформации, которая позволяет нивелировать влияние экстремально высоких значений и повысить чувствительность к колебаниям в нижнем диапазоне.

После нормализации оба показателя объединяются с равными весами, то есть суммируются и делятся пополам, а затем полученное значение масштабируется в диапазоне от 0 до 100. Этот итоговый балл представляет собой комбинированную меру важности, которая учитывает как активность (количество сообщений), так и авторитетность (рейтинг автора). Таким образом, высокий балл свидетельствует о том, что пост получил высокий рейтинг и/или содержит большое количество сообщений, что может указывать на повышенный уровень обсуждения и потенциальную угрозу.

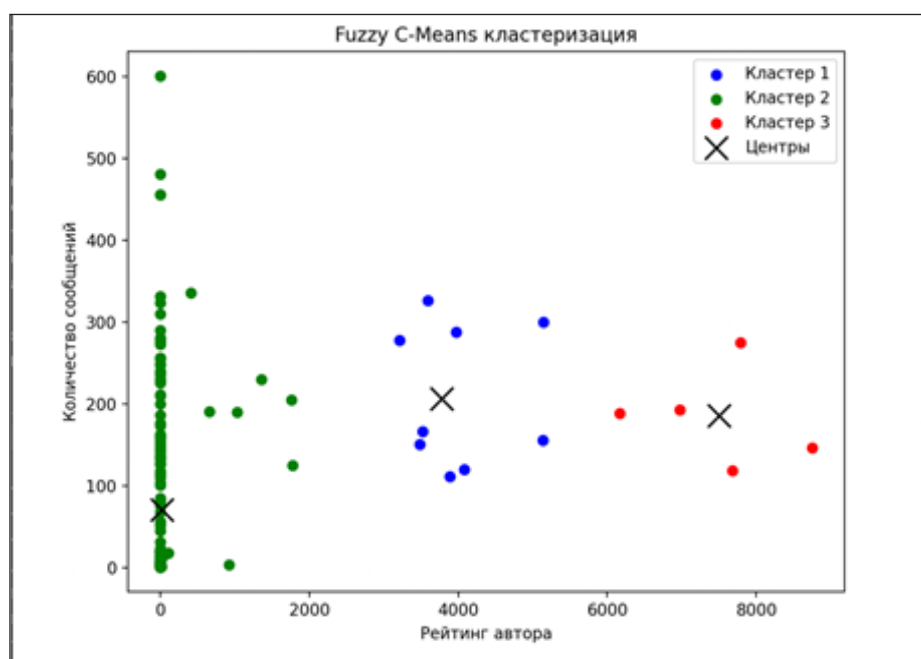


Рисунок 7 – Результат кластеризации

Далее в зависимости от итогового балла посту присваивается метка: если значение балла превышает высокий порог (например, 66), то пост получает метку high; если балл находится в среднем диапазоне (от 33 до 66) – метку medium; а если ниже – метку low. Эта классификация позволяет разделить посты по уровням угрозы, что является важным этапом для последующего анализа. Результаты этой классификации представлены на рисунке 8, где показано распределение количества постов по соответствующим меткам.

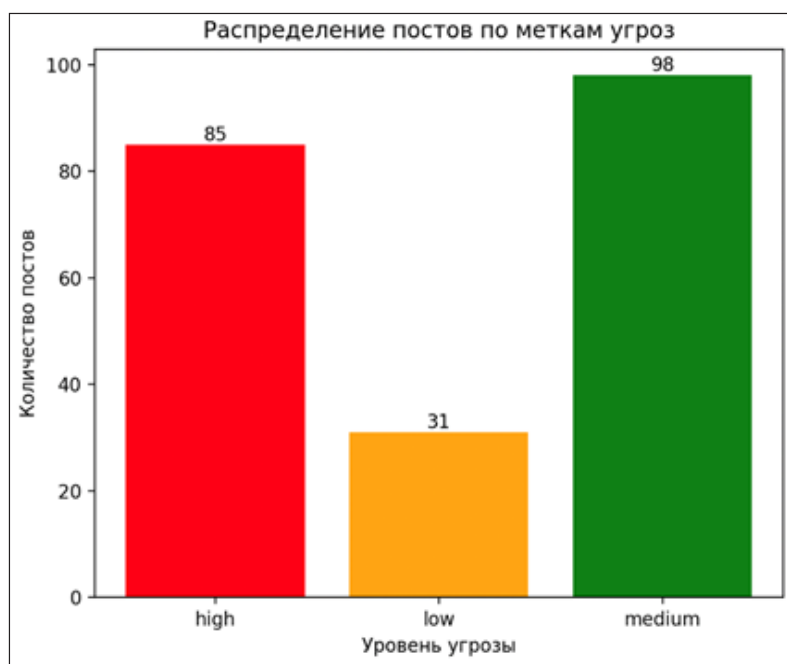


Рисунок 8 – Результат нейро-нечеткого алгоритма

Полученные таким образом результаты подтверждают, что предварительное групповое разбиение данных на три нечетких кластера дает достаточно устойчивые «прототипы», которые впоследствии могут быть эффективно использованы для генерации правил в DENFIS.

Заключение

В ходе проведенного исследования удалось успешно достичь поставленной цели – разработать интегрированный подход к обнаружению киберугроз на основе анализа тематических интернет-ресурсов с использованием методов кластеризации и нейро-нечеткой логики. Созданная система позволяет объединять все ключевые этапы – от извлечения и предварительной обработки данных до кластеризации с помощью алгоритма fuzzy c-means и формирования начальных правил для последующей работы нейро-нечеткой модели DENFIS. Полученные результаты демонстрируют высокую эффективность модели, что подтверждается следующими краткими выводами:

Извлечение и предварительная обработка данных позволили сформировать набор признаков (рейтинг автора и количество сообщений), обеспечивающих корректное представление исходной информации.

Применение алгоритма нечеткой кластеризации привело к выделению устойчивых прототипов, характеризующих высокими значениями коэффициента нечеткого разбиения ($FPC \approx 0.93$), что указывает на достаточно четкое распределение объектов.

Дополнительные метрики качества кластеризации – Partition Entropy (0.28) и индекс Xie-Beni (0.042) – подтвердили компактность и разнесенность полученных кластеров, что свидетельствует о высокой определенности группировки данных.

На основании полученных кластерных центров были сформированы первоначальные правила для нейро-нечеткой модели DENFIS, позволяющие присваивать постам метки high, medium или low в зависимости от совокупного балла, рассчитанного на основе нормализованных значений рейтинга и количества сообщений.

Итоговая интеграция предварительной кластеризации с динамической адаптацией правил в системе DENFIS обеспечивает оперативное и точное определение уровня угрозы обсуждений, что является критически важным для своевременного реагирования на потенциальные кибератаки.

В дальнейшем планируется расширение функциональности предлагаемой системы за счет интеграции дополнительных источников данных, таких как специализированные социальные сети, новые тематические интернет-ресурсы и форумы. Также будет вестись работа по улучшению этапа предварительной обработки, чтобы повысить качество извлечения информации из неструктурированных текстов. Особое внимание планируется уделить разработке адаптивных механизмов обновления правил DENFIS с применением онлайн-обучения, что позволит системе быстрее реагировать на появление новых угроз. Также рассматривается возможность внедрения современных трансформерных моделей для более глубокой семантической обработки текстов, что повысит точность классификации постов и обнаружения угроз. Эти направления будущей работы направлены на дальнейшее повышение оперативности, точности и надежности системы мониторинга, что имеет решающее значение для обеспечения информационной безопасности в условиях динамично меняющейся цифровой среды.

ЛИТЕРАТУРА

- 1 Sommer, R., and Paxson, V. Outside the Closed World: On Using Machine Learning for Network Intrusion Detection. Proceedings of the IEEE Symposium on Security and Privacy, 305–316 (2010). <https://doi.org/10.1109/SP.2010.25>.
- 2 Bilge, L., and Dumitras, T. Before We Knew It: An Empirical Study of Zero-Day Attacks in the Real World. Proceedings of the ACM Conference on Computer and Communications Security (CCS), 83–94 (2012). <https://doi.org/10.1145/2382196.2382284>.

- 3 National Institute of Standards and Technology (NIST). National Vulnerability Database (NVD) (2010). <https://nvd.nist.gov/>.
- 4 Zhao, X., Wang, X., and Li, X. Vulnerability Disclosure and Information Verification Delays in Cybersecurity. *Journal of Cybersecurity Research*, 3 (2), 45–60 (2015).
- 5 Wang, W., and Lu, Y. Mining Cyber Threat Intelligence from the Dark Web. *IEEE Transactions on Information Forensics and Security*, 13 (2), 275–286 (2018). <https://doi.org/10.1109/TIFS.2017.2761918>.
- 6 O'Connor, N., and Torabi, A. Cyber Threat Intelligence: An Introduction. *IEEE Security & Privacy*, 13 (3), 19–27 (2015).
- 7 Cadar, C., Dunbar, D., and Engler, D. KLEE: Unassisted and Automatic Generation of High-Coverage Tests for Complex Systems Programs. *Proceedings of the USENIX Symposium on Operating Systems Design and Implementation (OSDI)*, 209–224 (2008).
- 8 Yadav, S., Sharma, A., and Gupta, M. Indicators of Compromise Analysis Using Threat Intelligence Platforms. *Journal of Cybersecurity*, 5 (4), 210–223 (2019). <https://doi.org/10.5555/1855741.1855756>.
- 9 Devlin, J., Chang, M.W., Lee, K., and Toutanova, K. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. *Proceedings of NAACL HLT*, 4171–4186 (2019). <https://doi.org/10.18653/v1/N19-1423>.
- 10 Aghaei, E., Jain, S., Arun, P., and Sambamoorthy, A. SecureBERT 2.0: Advanced Language Model for Cybersecurity Intelligence. Cisco AI, San Jose, CA, USA (2025). {eaghaei, sjain2, parun, asambamo}@cisco.com.
- 11 Jin, Y., Jang, E., Cui, J., Chung, J.-W., Lee, Y., and Shin, S. DarkBERT: A Language Model for the Dark Side of the Internet. KAIST, Daejeon, South Korea; S2W Inc., Seongnam, South Korea (2023). {ijjinjin, claude}@kaist.ac.kr; {genesith, geeoon19, jwchung, lee}@s2w.inc. <https://aclanthology.org/2023.acl-long.415.pdf>.
- 12 Güven, M. A Comprehensive Review of Large Language Models in Cyber Security. *International Journal of Computational and Experimental Science and Engineering*, 10 (3), 507–516 (2024). <https://doi.org/10.22399/ijcesen.469>.
- 13 Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A.N., Kaiser, L., and Polosukhin, I. Attention Is All You Need. *Advances in Neural Information Processing Systems (NeurIPS)*, 5998–6008 (2017). <https://doi.org/10.48550/arXiv.1706.03762>.
- 14 Jang, J.S.R. ANFIS: Adaptive-Network-Based Fuzzy Inference System. *IEEE Transactions on Systems, Man, and Cybernetics*, 23 (3), 665–685 (1993). <https://doi.org/10.1109/21.256541>.
- 15 Kasabov, N., and Song, Q. DENFIS: Dynamic Evolving Neural-Fuzzy Inference System and Its Application for Time-Series Prediction. *IEEE Transactions on Fuzzy Systems*, 10 (2), 144–154 (2002). <https://doi.org/10.1109/91.995117>.
- 16 Poletaev, V.S. Informatsionno-analiticheskaya sistema prognozirovaniya ugroz i uyazvimostey informatsionnoy bezopasnosti na osnove analiza dannykh tematiceskikh internet-resursov [Information-Analytical System for Predicting Cybersecurity Threats and Vulnerabilities Based on Analysis of Thematic Internet Resources] (Ulyanovsk: UISU, 2024), 172 p. (in Russian)

¹*Самуйлова А.,

магистрант, ORCID ID: 0009-0003-5806-8542,

*e-mail: anastassiyasamuilova@gmail.com

¹Өл-Фараби атындағы Қазақ ұлттық университеті, Алматы қ., Қазақстан

ИНТЕРНЕТ РЕСУРСТАРЫ ДЕРЕКТЕРІ НЕГІЗІНДЕ АҚПАРАТТЫҚ ҚАУІПСІЗДІКТІҢ ОСАЛ ТҮСТАРЫН АНЫҚТАУ ҚҰРАЛДАРЫ

Аңдатпа

Киберқауіптердің күрделенуіне байланысты дәстүрлі осалдықтарды анықтау әдістері өз тиімділігін жоғалта бастады. Бұл зерттеудің мақсаты – форумдар, блогтар және әлеуметтік желілер сияқты тақырыптық интернет-ресурстар деректерін талдау негізінде осалдықтарды анықтауға арналған тәсілді әзірлеу және сынақтан өткізу. Аталған дереккөздер құрылымдалмаған ақпараттың үлкен көлемін қамтиды, сондықтан

зияткерлік деректерді талдау әдістерін қолдану қажет. Жұмыста заманауи технологиялар біріктірілген: киберқауіпсіздікке бейімделген алдын ала оқытылған тілдік модель – Security Bidirectional Encoder Representations from Transformers (SecBERT) және динамикалық дамитын нейро-анық емес шығару жүйесі – Dynamic Evolving Neural-Fuzzy Inference System (DENFIS). Ұсынылған жүйе маңызсыз хабарламаларды сүзгіден өткізуге, қауіп туралы индикаторларды және ықтимал осалдықтарды анықтауға мүмкіндік береді. Анық емес логиканы қолдану белгісіз және толық емес ақпаратты тиімді өңдеуге жағдай жасайды. Жүргізілген эксперименттер жоғары дәлдікпен жіктеу және тұрақты шынжырлы кластерлеу нәтижелерін көрсетті (FPC = 0.93; PE = 0.28; XB = 0.042). Жүйе киберқауіптердің белгілерін уақтылы анықтай алатынын көрсетті және мониторинг пен шабуылдарды болжау міндеттері үшін ауқымдауға мүмкіндік бар екенін дәлелдеді. Зерттеу нәтижелері бұл тәсілдің киберқауіптерді уақтылы анықтауға және ақпараттық жүйелердің қауіпсіздігін арттыруға ықпал ете алатынын көрсетеді.

Тірек сөздер: ақпараттық қауіпсіздіктің осал тұстары, тақырыптық интернет-ресурстар, трансформерлер, SecBERT, нейро-анық емес логика.

¹*Samuilova A.,

Master's student, ORCID ID: 0009-0003-5806-8542,

*e-mail: anastassiasamuilova@gmail.com

¹Al-Farabi Kazakh National University, Almaty, Kazakhstan

TOOLS FOR IDENTIFYING INFORMATION SECURITY VULNERABILITIES BASED ON DATA FROM INTERNET RESOURCES

Abstract

As cyber threats become more complex, traditional vulnerability detection methods lose their effectiveness. The purpose of this work is to develop and test an approach to identifying vulnerabilities based on the analysis of data from thematic Internet resources: forums, blogs and social networks. These sources contain a large amount of unstructured information, which requires the use of data mining methods. The work uses the integration of modern technologies: the pre-trained SecBERT language model (Security Bidirectional Encoder Representations from Transformers), designed for cybersecurity tasks, and the adaptive neuro-fuzzy inference system DENFIS (Dynamic Evolving Neural-Fuzzy Inference System). The proposed system allows you to filter irrelevant messages, highlight indicators of compromise and potential threats. The use of fuzzy logic makes it possible to efficiently process vague and incomplete information. Experiments confirmed high classification accuracy and stable fuzzy clustering performance (FPC = 0.93; PE = 0.28; XB = 0.042). The system demonstrated the ability to promptly detect signs of cyber threats and has scalability potential for monitoring and attack prediction tasks. The results indicate its potential in increasing the speed of response to cyber threats and strengthening the protection of information systems.

Keywords: information security vulnerabilities, thematic Internet resources, transformers, SecBERT, neuro-fuzzy logic.

Дата поступления статьи в редакцию: 19.05.2025

ӘОЖ 004.93
FTAXP 28.23.15

<https://doi.org/10.55452/1998-6688-2025-22-4-119-130>

¹**Есмұхамедов Н.С.,**

докторант, ORCID: 0009-0006-0652-3082,
e-mail: yesmukhamedov.yeskendyr@gmail.com

^{1*}**Сапақова С.З.,**

ф.-м.ғ.к., қауымдастырылған профессор, ORCID ID: 0000-0001-6541-6806,
*e-mail: sapakovasz@gmail.com

²**Қожамкулова Ж.Ж.,**

PhD, қауымдастырылған профессор, ORCID ID: 0000-0002-9780-9767,
e-mail: zh.kozhamkulova@aes.kz

³**Даниярова Д.Р.,**

т.ғ.к., қауымдастырылған профессор, ORCID: 0000-0002-6833-6462,
e-mail: duriya.daniyarova@mail.ru

¹**Арманкызы Р.,**

магистр, лектор, ORCID ID: 0009-0009-2236-559X,
e-mail: armankyzyrenata@gmail.com

¹Халықаралық ақпараттық технологиялар университеті, Алматы қ., Қазақстан

²Г. Дукеев атындағы КЕАҚ АЭЖБУ, Алматы қ., Қазақстан

³Халықаралық білім беру корпорациясы, Алматы қ., Қазақстан

ДИАБЕТТІК РЕТИНОПАТИЯНЫ АНЫҚТАУДА КӨЗ ТҮБІ СУРЕТТЕРІН АЛДЫН АЛА ӨНДЕУ ЖӘНЕ ТАЛДАУ ӘДІСТЕРІ

Аңдатпа

Бұл жұмыс диабеттік ретинопатияны анықтауға арналған көз түбі суреттерін алдын ала өңдеу және талдау әдістерін зерттеуге арналған. Диабеттік ретинопатия (ДР) – қант диабетімен ауыратын науқастарда жиі кездесетін көз ауруы, және оның ерте диагностикасы көру қабілетінің жоғалуын болдырмауға мүмкіндік береді. Зерттеу барысында көз түбі суреттерін өңдеу мен талдау үшін заманауи әдістер, соның ішінде Deep Learning негізінде құрылған EfficientNetB0 архитектурасы қолданылды. Алдын ала өңдеу үшін сурет аугментациясы (айналдыру, масштабтау, кесу, контрастты арттыру) және нормализация әдістері енгізілді. EfficientNetB0 архитектурасын пайдалану барысында екі тәсіл сыналды: базалық қабаттарды қатырып оқыту және жоғарғы қабаттарды ашу арқылы қосымша бейімдеу (fine-tuning). Нәтижелер метрикалар бойынша бағаланды. Бірінші әдісте тест жиыны үшін дәлдік (precision) 65%, екінші әдісте 75% болды. Бірінші әдіс бойынша валидациялық жиынның дәлдігі 63%, ал екінші әдіс бойынша 71% жетті. Қайталауды анықтау (recall) метрикасы бірінші әдісте тест жиыны үшін 60%, екінші әдісте 74% көрсетті. Жалпы, fine-tuning әдісінің көрсеткіштері жоғары болды. Бұл әдістерді пайдалану диабеттік ретинопатияны тиімді диагностикалау үшін суреттерді өңдеу мен классификациялаудың сапасын арттыруға мүмкіндік береді. Зерттеудің жаңашылдығы – тиімділігі жоғары EfficientNetB0 архитектурасын қолдану және бейімдеудің әртүрлі тәсілдерін талдау. Алынған нәтижелер ДР диагностикасында автоматтандырылған жүйелердің сапасын жақсартуға және модельдің энергия тиімділігін арттыруға мүмкіндік береді. Ұсынылған әдістер көз ауруларын ерте кезеңде анықтауда жоғары потенциалға ие.

Тірек сөздер: диабеттік ретинопатия, машиналық оқыту, деректерді талдау, лазерлік коагуляция, фундус-суреттер.

Кіріспе

Диабеттік ретинопатия (ДР) қант диабетімен ауыратын науқастарда жиі кездесетін көз ауруы, ол көру қабілетінің жоғалуының негізгі себептерінің бірі. ДР ерте кезеңдерде

анықталса, оның дамуын тоқтатуға немесе баяулатуға болады, ал дәл диагностикалау көруді сақтап қалуға мүмкіндік береді. Бұл әдебиеттік шолудың мақсаты – диабеттік ретинопатияны диагностикалауда қолданылатын қазіргі заманауи әдістер мен технологияларды талдау. Әсіресе, терең оқыту және бейімделген нейрондық желілердің қолданылуы қарастырылады. Мысалы, авторлар [1–3] мақалада диабеттік ретинопатияны диагностикалауға бағытталған терең оқыту үлгілерін әзірлеу үшін кескінді өңдеу әдістерін пайдалану талқыланады. Көз түбінің фотосуреттерін талдайтын және дәл диагноз қою үшін қажетті негізгі үлгілерді танитын нейрондық желілерге назар аударылады. Арнайы өңдеу әдістері көрсетілмегенімен, кескінді жақсарту, қан тамырлары мен зақымдануларды бөлектеу үшін сегменттеу және құрылым мен түс сияқты ерекшеліктерді алу күтілуде. Зерттеу толық модельді оқытуға мүмкіндік беретін «Бразилиялық көп мақсатты офтальмология деректер жинағына» негізделген. Мақалада сонымен қатар диабеттік ретинопатия мен онымен байланысты ауруларды диагностикалауға арналған модельдерді оқытудың үш тәсілі баяндалады, диагностикалық дәлдікті арттыруда кескінді өңдеу әдістерінің маңыздылығы көрсетіледі. Келесі [2] жұмыста диабеттік ретинопатия үшін лазерлік коагуляцияны емдеуді жақсарту үшін көз түбінің кескіндерін талдауда терең нейрондық желілерді пайдалануды зерттейді. Зерттеу екі нейрондық желіні пайдаланады: біріншісі – көз түбінің кескіндеріндегі анатомиялық құрылымдарды анықтау үшін, екіншісі – лазерлік емдеуді қажет ететін ісіну аймақтарын анықтау үшін. Екі желінің шығыстарын біріктіру арқылы әдіс оңтайлы емдеу аймақтарын дәл анықтауға, медициналық араласулардың дәлдігін арттыруға мүмкіндік береді. Қысқаша айтқанда, бұл зерттеуде қолданылатын әдістер анатомиялық құрылымдар мен ісіну аймақтарын шығаруға баса назар аудара отырып, көз түбінің кескіндерін талдау үшін терең нейрондық желілерді қолдануды қамтиды. Бұл нәтижелерді біріктіру диабеттік ретинопатия үшін лазерлік коагуляцияны емдеудің дәлдігін арттыруға бағытталған. Авторлар [3–7] диабеттік ретинопатияның (DR) автоматтандырылған диагностикасы үшін деректерді жинау және өңдеудің жүйелі тәсілі сипатталған, ол бірнеше негізгі қадамдарды қамтиды. Кескінді алу клиникалық маңыздылықпен кескін сапасын қамтамасыз ете отырып, офтальмологиялық тәжірибелермен бірлесіп, офтальмоскопия көмегімен жүзеге асырылады. Содан кейін кескіндер әзірленген алгоритм арқылы орталықтандырылған өңдеуге арналған арнайы серверге жүктеледі, мұнда графикалық интерфейс медицина мамандары үшін жүктеп алуды жеңілдетеді. Жүктелген кескіндер оптикалық диск және зақымданулар сияқты DR негізгі ерекшеліктерін анықтауға бағытталған морфологиялық алгоритмдерді пайдалана отырып, бірқатар өңдеу қадамдарынан өтеді. Мамандар бұрын қолмен тексеруді қажет ететін процесті автоматтандыру арқылы диагностикалық дәлдігі мен тиімділігін арттыратын терең оқыту әдістері мен нейрондық желілерді пайдалана отырып, автоматтандырылған талдауға басты назар аударылады. Бүкіл процесс диабеттік ретинопатияны ерте анықтауды жақсартуға бағытталған, ол қант диабеті және соған байланысты көру қабілетінің жоғалуын азайтуға көмектеседі, осылайша емдеудің тиімді нәтижелерін береді.

Келесі жұмыста [4] терең оқыту және кескінді алдын ала өңдеу әдістерін қолдана отырып, диабеттік ретинопатияны (DR) болжау бойынша бірнеше негізгі қорытындылар берілген. Біріншіден, ол көрудің нашарлауын болдырмау үшін DR ерте анықтау маңыздылығын көрсетеді, бұл емнің нәтижелерін және пациенттердің өмір сүру сапасын айтарлықтай жақсартып алады. Екіншіден, зерттеу дәл талдау үшін шуды азайту және айқындықты арттыру арқылы кескін сапасын жақсартатын Sobel, Winer, Gauss және жергілікті емес орташа сүзгілер сияқты кескінді алдын ала өңдеудің әртүрлі әдістерінің тиімділігіне назар аударады [5]. Терең конволюциялық нейрондық желі (CNN) кескінді өңдеуден кейін DR болжау үшін қолданылады, аурумен байланысты күрделі үлгілерді ашады. Алдын ала өңдеу және CNN болжау әдістерінің үйлесімі кескін сапасын және болжау дәлдігін арттыра отырып [6–7], DR талдауының перспективалы тәсілін көрсетеді. Зерттеу нәтижелері денсаулық сақтау провайдерлеріне көмектесетін және емдеу нәтижелерін жақсартатын және аурумен байланысты ауыртпалықты азайтатын DR үшін

компьютерлік диагностикалық жүйелердің әлеуетін көрсетеді. Қорытындылай келе, мақалада кескінді алдын ала өңдеу әдістері мен терең оқыту үлгілерінің интеграциясы диабеттік ретинопатияны болжау мен диагностикасының дәлдігін айтарлықтай жақсартып, медициналық шешімдерді жақсартуға жол ашатынын айтады. Авторлар [4–10] алдын ала өңдеу әдістері және CNN моделі сияқты әдістемелерге назар аудара отырып, диабеттік ретинопатияны болжауды зерттеу үшін қандай ретинальды кескін жиынтығы пайдаланылғаны көрсетілмеген. Онда CNN моделі бақыланатын оқыту үшін қажетті таңбаланған деректер жиынында оқытылатыны айтылады, бірақ бұл деректер жиынының нақты атауы немесе көзі көрсетілмейді. Алдын ала өңдеу әдістерінің маңыздылығына баса назар аударылғанымен, пайдаланылған деректер жиынының сипаттамалары немесе шығу тегі ескерілмейді, бұл оны идентификациялауды шектейді. Қорытындылай келе, қарастырылған жұмыстарда қолданылатын әдістерге назар аудара отырып, белгілі бір кескіндер жиынтығы туралы ақпарат берілмейді, бірақ мұндай зерттеулердегі жалпы деректер жиыны туралы ақпарат алу үшін EyePACS немесе Kaggle диабеттік ретинопатияны анықтау сияқты жалпыға қолжетімді жинақтарды іздеуге болады, олар жиі кездеседі және ДР зерттеулерінде қолданылады.

Авторлардың [6–7] мақаласында диабеттік көз ауруларын (DED) жіктеу үшін бірнеше жалпыға қолжетімді деректер жинақтары пайдаланылады. Зерттеуде Messidor, Messidor-2, DRISHTI-GS және GitHub жинақтарының суреттері пайдаланылады, оларда ретинопатия дәрежесі және макулярлы ісіну қаупі бар офтальмолог таңбаланған кескіндер бар. Әрбір сурет қан кетулер мен микроаневризмалар сияқты белгілі бір зақымдардың болуына негізделі таңбалады. Keras ImageDataGenerator көмегімен геометриялық түрлендірулерді (айдау, айналдыру және қию) қоса алғанда, жаттығу деректерінің өлшемін үлкейту үшін деректерді кеңейту әдістері қолданылады. Қалыпты үлгілер арасында ЖСҚ жеңіл жағдайларын анықтауға ерекше назар аударылады, кескіндердегі нәзік айырмашылықтарды ажырату үшін үлгілерді үйрету үшін пайдаланылатын деректер жиынының маңыздылығын көрсетеді. Сонымен, зерттеу диабеттік көз ауруларын анықтауға бағытталған терең үлгілердің өнімділігін жақсарту үшін көптеген деректер жинақтары мен алдын ала өңдеу әдістерін пайдалануды көрсетеді. Авторлардың [6–7] мақаласы әртүрлі әдістерді қолдана отырып, диабеттік көз ауруларын (ДР) анықтауды жақсартудың жүйелі тәсілін сипаттайды [8–11]. Негізгі әдістерге көз түбінің сапасын жақсартуға арналған кескінді өңдеудің дәстүрлі әдістері, қызығушылық аймақтарын оқшаулау үшін сегменттеу, аудару, айналдыру және Keras ImageDataGenerator арқылы кескінді қию арқылы деректерді ұлғайту жатады. Зерттеу сонымен қатар DED классификациясы үшін ResNet50, VGG-16 және Xception сияқты алдын ала дайындалған конволюционды нейрондық желілерді (CNN) пайдаланады [12–15], сонымен қатар салыстыру үшін өзінің CNN архитектурасын үйретеді. Гиперпараметрлерді баптау олардың үлгі өнімділігіне әсерін бағалайды және эксперимент нәтижелері дәлдік, ерекшелік және сезімталдық үшін талданады. Қорытындылай келе, дәстүрлі өңдеу әдістері мен жаңа CNN архитектурасының үйлесімі ЖСҚ жіктеу тапсырмаларында ең жақсы нәтижелерді көрсетеді, бұл осы ауруларды ерте анықтауды жақсартуға көмектеседі.

Материалдар мен әдістер

Диабеттік ретинопатияны (DR) жіктеу тапсырмасында болжаулардың жоғары дәлдігі мен негізділігін қамтамасыз ететін бірқатар әдістер мен алгоритмдер, нейрондық желі қолданылды. Диабеттік ретинопатия – көздің түбін бейнелеу арқылы диагноз қойылған тордың прогрессивті ауруы [16]. Қиындық аурудың әртүрлі кезеңдері арасындағы көрнекі ұқсастықта жатыр, ол күрделі кескінді талдау мен деректерді өңдеу әдістерін қажет етеді. Кескіндерді жіктеу үшін заманауи машиналық оқыту әдістері қолданылды [17]. Модельді жалпылауды жақсартатын, олардың көлемін және өзгермелілігін арттыру үшін деректерді кеңейту (айналдыру, масштабтау, жылжу, жарықтықты өзгерту) қолданылды. Оқыту бес сыныптың ықтималдығын болжау үшін

шығыс softmax қабаты бар `init_model` арқылы инициализацияланған терең конволюциялық нейрондық желіні қолдану арқылы жүзеге асырылды. Қатені азайту үшін кросс-энтропия жоғалту функциясы (`CrossEntropyLoss`) пайдаланылды [18].

Оңтайландыру оқу үдерісін тұрақтандыратын қадамдық оқыту жылдамдығын жоспарлаушымен (`StepLR`) толықтырылған Adam алгоритмімен орындалды. Өнімділікті бағалау үшін класстық теңгерімсіздікті есепке алу үшін квадраттық салмақтары бар Коэннің Каппа метрикасы пайдаланылды. Дәлдік, еске түсіру және F1 ұпайлары да талданған. Ерте тоқтату шамадан тыс орнатудың алдын алды, ал ең жақсы салмақтарды сақтау нәтижелердің қайталануын қамтамасыз етті [19].

Бекітілген қатпарлар саны бар кросс-валификация үлгінің беріктігін тексерді. Бұл әдістерді қолдану тұрақты өнімділікті қамтамасыз етті және одан әрі оңтайландыру бағыттарын анықтады, соның ішінде сыныптық теңгерімсіздікпен күресу [20].

1. Кросс-валидация

Кросс-валидация күрделілік пен деректер шектеулеріне қарамастан үлгі өнімділігін сенімді бағалауды қамтамасыз етті. ДР классификациясы контекстінде:

- ♦ Кросс-валидация аурудың неғұрлым ауыр кезеңдері (мысалы, 3 және 4
- ♦ сыныптар) сирек кездесетін теңгерімсіз класс таралуының әсерін азайтты [15].
- ♦ Модельдің әрбір к-қатардағы орташа өнімділігі оның жаңа деректерге жалпылау мүмкіндігі туралы объективті түсінік береді:

$$M = \frac{1}{k} \sum_{i=1}^k M_i, \quad (1)$$

мұндағы M_i – i -ші қабаттағы метрика. Бұл әсіресе маңызды, өйткені жеңіл кезеңді ауыр кезең ретінде қате жіктеу артық емдеуге әкелуі мүмкін, ал ауыр кезеңді өткізіп жіберу кеш араласуға әкелуі мүмкін.

2 Конволюция. Модельдегі конволюциялық қабаттар көз түбі кескіндерін талдауда маңызды рөл атқарды және жергілікті ерекшеліктерін анық көрсетеді, мысалы: диабеттік ретинопатияға тән ісіну, қан кету және микротамырлық ауытқулар конвульсиялар арқылы автоматты түрде шығарылды [17]:

$$(x, y) = \sum_{i=1}^m \sum_{j=1}^n K(i, j) \cdot I(x + i, y + j), \quad (2)$$

мұнда, $I(x+i, y+j)$ – бастапқы кескіннің пиксельдік мәні (кіріс бейненің қарқындылығы); $K(i, j)$ – тұқыртпа ядросы немесе фильтрдің салмақ коэффициенттері; m, n – сүзгінің (ядроның) биіктігі мен ені; (x, y) – шығыс кескіннің (feature map) координаталары. Бұл операция нәтижесінде бастапқы кескіннің ерекшеліктері (мысалы: жиектер, текстуралар) бөлініп алынады.

3. Softmax функциясы. DR сатысын 5 сыныптың біріне жіктеу үшін модельдің шығыс мәндерін сыныптардың әрқайсысына жататын ықтималдықтарға түрлендіретін softmax функциясы қолданылды:

$$P(y = k | x) = \frac{\exp(z_k)}{\sum_{j=1}^C \exp(z_j)}, \quad (3)$$

мұнда, $P(y=k | x)$ – берілген кіріс x үшін оның k -сыныпқа жату ықтималдығы; z_k – модельдің соңғы қабатындағы k -сыныпқа сәйкес келетін логит (шығыс активацияға дейінгі мән); C – сыныптар саны (мысалы, диабеттік ретинопатияны бес сатыға жіктеу кезінде $C=5$); $\exp(\cdot)$ – экспоненциалдық функция, ол барлық логиттерді оң мәндерге түрлендіреді.

Бұл модельдік болжамдарды «ретинопатияның жоқтығынан» (0-сынып) «ауыр кезеңге» (4-сынып) дейінгі кезеңдердің ықтималдығы ретінде түсіндіруге мүмкіндік берді. Softmax

модельге белгісіздікті азайтып, ең жоғары ықтималдығы бар сыныпқа назар аударуға көмектеседі [18–19].

4. Айқас энтропия жоғалту функциясы

DR жіктеу мәселесі теңдестірілмеген сыныптарды қамтитындықтан, кросс-энтропия жоғалту функциясы шынайы класс белгілері мен болжамды ықтималдықтар арасындағы сәйкессіздікті азайтуға көмектесті:

$$L = -\frac{1}{N} \sum_{i=1}^N \sum_{k=1}^C y_{i,k} \log(\hat{y}_{i,k}), \quad (4)$$

мұндағы $y_{i,k}$, егер объект k класына жататын болса, 1 мәнін, ал басқаша жағдайда 0 мәнін қабылдайды. Бұл модельде сирек, бірақ маңызды сыныптарды дәл анықтауға мүмкіндік береді (мысалы, 4-сынып, DR ауыр сатысы).

5. Adam Optimizer

Модельді үйрету үшін Adam оптимизаторы пайдаланылды [21]. DR жіктеу мәселесінде ол жылдам және тұрақты конвергенцияға ықпал еткен градиенттердің адаптивті бақылауын қамтамасыз етті:

$$\theta_{t+1} = \theta_t - \eta \frac{\hat{m}_t}{\sqrt{\hat{v}_t + \epsilon}}. \quad (5)$$

Бұл деректердің біркелкі еместігіне байланысты градиенттер сирек болуы мүмкін күрделі көз түбінің кескіндерін талдау кезінде өте маңызды (мысалы, микроаневризмалар сияқты кішігірім зақымданулар сирек кездеседі).

6. Коэннің каппасы

Диабеттік ретинопатия диагностикалық тапсырмасында модель болжамдары мен негізгі шындық белгілері арасындағы келісімді бағалау үшін квадраттық салмақтары бар Коэннің Каппа метрикасы пайдаланылды [22]:

$$\kappa = \frac{p_o - p_e}{1 - p_e}. \quad (6)$$

Квадрат салмақтар:

$$w_{ij} = 1 - \left(\frac{i - j}{C - 1} \right)^2, \quad (7)$$

іргелес сыныптар арасындағы қателер (мысалы, 2 және 3) қашықтағы сыныптар арасындағы қателерден (мысалы, 0 және 4) аз жазаланғанын қамтамасыз етті. Бұл клиникалық мағынаға ие, өйткені жеңіл кезеңдер арасындағы ауысудағы қателер ауыр жағдайларды жіберіп алудан гөрі маңызды емес.

7. Ерте тоқтау

Егер валидация жинағындағы Коэннің Каппа метрикасы k дәуірде жақсармаса, шамадан тыс фитингтің алдын алу үшін жаттығуды тоқтату үшін ерте тоқтату қолданылды: егер $M_t - M_{t-k} < \epsilon$, онда жаттығу тоқтайды.

Бұл деректердегі шуылға шамадан тыс бейімделуді болдырмай, оңтайлы модельді оқытуды қамтамасыз етті.

8. Деректерді кеңейту

Деректерді кеңейту мыналарды қамтиды:

- ◆ Айналдыру (0): Көз түбінің бұрыштарының өзгеруін ескере отырып, кескіндер кездейсоқ бұрышқа айналдырылды.

♦ Масштабтау: Көру өрісіндегі айырмашылықтарды имитациялау үшін кескіндердің өлшемін өзгерту.

♦ Жарықтықтың өзгеруі: түсіру кезінде жарықтандырудың өзгермелілігін ескереді.

Бұл әдістер модельге нақты жағдайға бейімделуге және диабеттік ретинопатияның негізгі ерекшеліктерін жақсырақ көрсетуге мүмкіндік берді. Сипатталған әдістерді қолдану деректердегі қиындықтарды ескере отырып, диабеттік ретинопатияны диагностикалау мәселесін тиімді шешуге мүмкіндік берді, мысалы, класстық теңгерімсіздік, шағын жергілікті белгілер және аурудың кезеңдері арасындағы ұқсастықтар. Таңдалған тәсілдердің математикалық негіздемесі олардың дұрыстығын және клиникалық маңызды нәтижелерді беру мүмкіндігін растайды.

Деректер жинағы және олардың сипаттамасы. Осы зерттеудің бөлігі ретінде оқу және сынақ жиынтықтарына бөлінген көз түбінің кескіндерін қамтитын деректер жиынтығы пайдаланылды. Жаттығу жиынында 35 126 сурет, ал сынақ жиынында 3 662 сурет болды. Әрбір суретте диабеттік ретинопатияның кезеңдерін анықтайтын бес санаттың біріне сәйкес келетін диагностикалық белгі болды: 0 – ауру жоқ, 1 – ерте кезең, 2 – орташа кезең, 3 – ауыр кезең, 4 – пролиферативті кезең (1-сурет).

Оқыту жиынындағы сабақтардың бөлінуі:

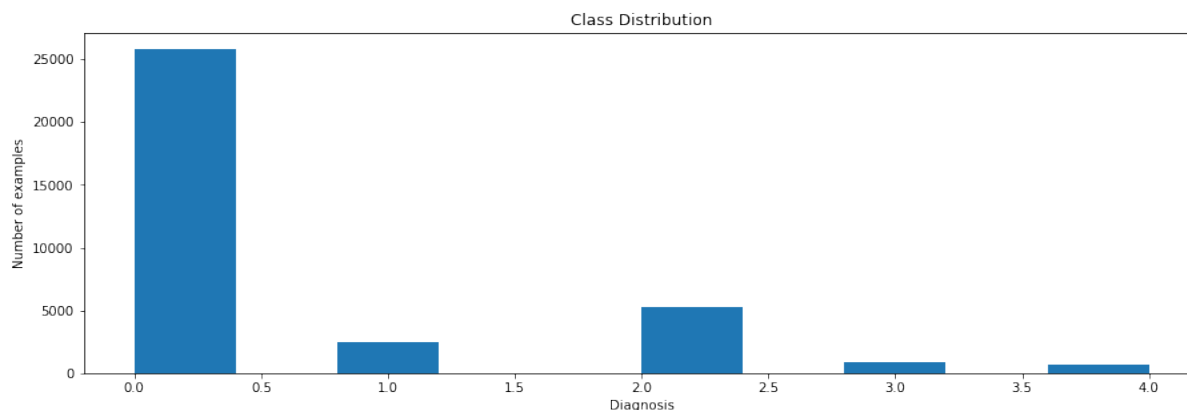
- ♦ 0-сынып (ретинопатия жоқ): 25 810 сурет (73,5%)
- ♦ 1-сынып (бастапқы кезең): 2443 сурет (7,0%)
- ♦ 2-сынып (орташа кезең): 5292 сурет (15,1%)
- ♦ 3-сынып (ауыр кезең): 873 сурет (2,5%)
- ♦ 4-сынып (пролиферативті кезең): 708 сурет (2,0%)

Сынақ топтамасындағы деректердің бөлінуі:

- ♦ 0-сынып (ретинопатия жоқ): 1805 сурет (49,3%)
- ♦ 1-сынып (бастапқы кезең): 370 сурет (10,1%)
- ♦ 2-сынып (орташа кезең): 999 сурет (27,3%)
- ♦ 3-сынып (ауыр кезең): 193 сурет (5,3%)
- ♦ 4-сынып (пролиферативті кезең): 295 сурет (8,0%)



Сурет 1 – Көз түбі кескіндерінің сыныптар бойынша жіктелуі



Сурет 2 – Сыныптардағы кескіндердің үлестірімінің гистограммасы

Бірдей көз түбінің 3-сурет екі нұсқасы көрсетілген: бастапқы сурет (сол жағында) және алдын ала өңделген сурет (оң жағында). Бастапқы сурет деректерді өңделмеген күйінде көрсетеді, ал өңделген сурет бірқатар сапа жақсарту кезеңдерінен өткен, бұл деректерді дәл интерпретациялау үшін маңызды.



Сурет 3 – Қалыпты және ДР кездесетін кескіндер

Алдын ала өңдеу бірнеше негізгі кезеңдерден тұруы мүмкін:

Айқындық деңгейін күшейту. Көз түбінің маңызды құрылымдарын, мысалы, қан тамырларын немесе тор қабығының зақымдануларын анық көрсету үшін суреттің контрастын арттыру. Бұл әсіресе диабеттік ретинопатия сияқты ауруларды диагностикалауда маңызды.

Жарықтың әсерін тұрақтандыру. Жарықтандырудың теңсіздігін азайту, бұл суреттердің анықтығын қамтамасыз етеді.

Артефактілер мен шуылдарды тегістеу. Суретте пайда болатын қажетсіз шуылдар мен артефактілерді жою, бұл суреттердің сапасын арттырып, кейінгі талдауға дайын етеді.

Түстік балансын өзгерту. Суреттегі маңызды құрылымдарды айқындау және оны түсінікті көрсету үшін түстік балансты реттеу.

Бұл қадамдар көз түбінің ауруларын, мысалы, диабеттік ретинопатия мен глаукоманы диагностикалау үшін маңызды компьютерлік көру тапсырмаларында қолданылатын суреттерді дайындау үшін қажет. Суреттердің дұрыс өңделуі автоматтандырылған диагностикалық жүйелердің дәлдігін және сенімділігін айтарлықтай арттырады, бұл айқын әрі ақпаратты деректер алуға мүмкіндік береді.

Деректерді өңдеу:

1. Класстың теңгерімсіздігі. Жаттығулар кешенінде жалпы деректердің 73,5%-ын құрайтын 0-сынып (ретинопатия жоқ) басым, ал ауыр кезеңдері (3 және 4-сыныптар) айтарлықтай аз, сәйкесінше 2,5% және 2,0%-ды көрсетеді 2-сурет. Сыныптың теңгерімсіздігін есепке алу үшін салмақты жоғалту функциялары және деректерді көбейту әдістері қолданылды.

2. Деректерді көбейту. Жаттығу деректерінің өзгермелілігін арттыру үшін кескінді айналдыру, масштабтау, жылжу және жарықтықты өзгерту сияқты әдістер қолданылды. Бұл әдістер модельді жалпылауды жақсартты және көз түбі кескіндерінің өзгергіштігін ескерді.

3. Сынақ үлгісінің мақсатын айқындау. Оқу процесінде пайдаланылмаған деректер бойынша модельдің өнімділігін объективті бағалау үшін тест жинағы бөлек бөлінді.

Пайдаланылған деректер жинағы ретинопатиясыз жағдайлар басым және ауыр кезеңдері сирек кездесетін нақты клиникалық тәжірибеге жақын жағдайларда ұсынылған модельдің өнімділігін зерттеуге мүмкіндік берді. Бұл ерекшеліктер класстық теңгерімсіздікпен тиімді күресетін және сирек, бірақ клиникалық маңызды жағдайлардың жіктелуінің дәлдігін жақсартатын әдістердің қажеттілігін көрсетті.

Нәтижелер мен талқылау

Зерттеуде бейнелерді классификациялау үшін EfficientNetB0 нейрондық желісі қолданылды. Бұл әдіс алдын ала оқытылған модельдерге негізделген Transfer Learning тәсілін пайдалануды қамтиды. Оқыту процесінде екі кезең орындалды: базалық қабаттарды қатыру және жоғарғы қабаттарды ашып, қосымша бейімдеу. EfficientNetB0 моделінің гиперпараметрлері алдын ала тәжірибе арқылы таңдалды. Оқыту жылдамдығы, батч мөлшері, оптимизатор түрі және регуляризация параметрлері бірнеше эксперименттік конфигурацияда бағаланды. Торлы іздеу (grid search) және кездейсоқ іздеу (random search) әдістері арқылы ең тиімді комбинация таңдалды. Кросс-валидация нәтижелері модельдің тұрақтылығы мен қайта өндірілуін қамтамасыз етті, бұл зерттеу нәтижелерін басқа деректер жиынтықтарында бағалауға мүмкіндік береді. Модельдің генерализациясын жақсарту үшін аугментация әдістері енгізілді (айналдыру, масштабтау, қиғаштау және т.б.), ал оқыту тиімділігін арттыру үшін Adam оптимизаторымен ReduceLROnPlateau және EarlyStopping сияқты коллбэктер қолданылды. Соңғы қабатта softmax активациясы арқылы бейнелерді мультикласстық классификациялау жүзеге асырылды.

Нәтижелер Precision, Recall, F1-Score және Accuracy метрикаларымен

бағаланды. Бұл тәсіл деректерді дәл және тиімді классификациялауға мүмкіндік берді. Эксперименттер барысында модельдің жұмысын бағалау үшін үш класстан тұратын деректер жиынтығы қолданылды: Train, Test, және Val. Негізгі метрикалар – Precision, Recall және F1-Score – әр класс үшін де, сондай-ақ жалпы түрде (қорытынды метрикалар) есептелді. Нәтижелер 4-суретте көрсетілген.

Results by Class:				
Class	Precision	Recall	F1-Score	Support
Test	0.65	0.60	0.62	615
Train	0.85	0.90	0.87	5004
Val	0.68	0.65	0.66	619
Summary Metrics:				
	Precision	Recall	F1-Score	Support
Accuracy				0.80
Macro Avg	0.73	0.72	0.72	
Weighted Avg	0.81	0.85	0.83	

Сурет 4 – Кластар бойынша классификация нәтижелері және қорытынды метрикалар

Нәтижелерді талдау

1. Кластар бойынша. Train класындағы метрикалар Test және Val кластарымен салыстырғанда айтарлықтай жоғары, бұл деректердің дисбалансын немесе модельдің жалпылау қабілетінің жетіспеушілігін көрсетуі мүмкін. Мысалы, Train үшін Precision 0.85 болса, Test және Val үшін сәйкесінше 0.65 және 0.68.

2. Қорытынды метрикалар. Accuracy 80%-ға жетіп, модельдің жалпы дәлдігі жоғары екендігін көрсетті. Орташа өлшенген Precision (0.81), Recall (0.85) және F1-Score (0.83) көрсеткіштері де модельдің тиімділігін растайды.

3. Қорытынды. Жаттықтыру деректері бойынша жоғары дәлдікке қарамастан, Test және Val кластарында салыстырмалы түрде төмен метрикалық моделдің артық үйренуіне

(overfitting) нұсқайды. Нәтижелерді жақсарту үшін деректерді теңестіру, гиперпараметрлерді жақсарту және қосымша регуляризация әдістерін қарастырылады.

Бұл экспериментте екі түрлі әдіс бойынша EfficientNetB0 моделінің өнімділігі тексерілді: бірінші әдіс барлық қабаттарды қатырып қояды және тек соңғы қабаттармен жұмыс істеуді қолданса, екінші әдіс жоғарғы қабаттарды бастапқы оқытудан кейін ашып, модельді одан әрі бейімдеу (fine-tuning) процесін жүргізді.

Нәтижелерді салыстыру:

1. Precision, Recall және F1-Score метрикалары бойынша екінші әдіс айтарлықтай жоғары нәтижелер көрсетті, әсіресе сынақтық және валидациялық деректерде. Әдіс 2 жоғарғы қабаттарды ашу арқылы модельдің күрделілігін арттырды және нақты деректерге жақсы бейімделуін қамтамасыз етті (5-сурет).

Метрика	Әдіс 1 (Қатырып қою)	Әдіс 2 (Жоғарғы қабаттарды ашу)
Precision (Train)	0.85	0.85
Precision (Test)	0.65	0.75
Precision (Val)	0.63	0.71
Recall (Train)	0.90	0.88
Recall (Test)	0.60	0.74
Recall (Val)	0.65	0.73
F1-Score (Train)	0.87	0.86
F1-Score (Test)	0.62	0.74
F1-Score (Val)	0.64	0.72
Accuracy	0.80	0.80
Macro Average	0.72	0.77
Weighted Average	0.74	0.81

Сурет 5 – Нәтижелерді салыстыру

2. Accuracy көрсеткіші екі әдіс бойынша да бірдей болды (0.80), бірақ Macro Average және Weighted Average метрикаларында екінші әдіс айтарлықтай жақсы нәтижеге жетті (0.77 және 0.81 сәйкесінше 0.72 және 0.74).

Екінші әдіс, яғни жоғарғы қабаттарды ашып, модельді қосымша бейімдеу (fine-tuning), тиімдірек болып шықты. Бұл әдіс тестілік және валидациялық деректерде жоғары нәтижелер берді, сол арқылы артық үйрету мәселесін азайтуға мүмкіндік берді және жалпы модельдің жалпы тиімділігін арттырды. Осылайша, екінші әдіс модельдің тиімділігі мен дәлдігін жақсартуға мүмкіндік берді.

Қорытынды

Зерттеу нәтижелері бойынша диабеттік ретинопатияны анықтау мен диагностикалау саласындағы қазіргі заманғы әдістердің тиімділігі айқын көрінді. Әсіресе, терең оқыту әдістері, оның ішінде Convolutional Neural Networks (CNN) және EfficientNet алдын ала оқытылған модельдері, диабеттік ретинопатияның әртүрлі кезеңдерін жоғары дәлдікпен анықтауға мүмкіндік береді. Сурет аугментациясы және алдын ала өңдеу әдістері, мысалы, нормализация мен шу мен контрастты арттыру, модельдің тиімділігін арттыруға оң әсер етеді.

Зерттеулерде қолданылған әдістердің арасында көрініс деңгейінде үлкен айырмашылықтар болмағанымен, модельдерді тексеру және баптау кезінде (fine-tuning) жоғары нәтижелерге қол жеткізуге болады. Сонымен қатар, қолданылған кері шақыру (callback) функциялары, яғни EarlyStopping және ReduceLROnPlateau, оқу процесін оңтайландыруға және модельдің артық үйренуден (overfitting) сақталуына көмектеседі. Алынған нәтижелер терең оқыту әдістерінің диабеттік ретинопатияны анықтауда жоғары тиімділігін көрсетті. Сурет аугментациясы мен алдын ала өңдеудің қолданылуы модельдің дәлдігін арттырды, ал EfficientNet және CNN модельдері жоғары нәтижелерге қол жеткізді. Кері шақыру (callback) функциялары

(EarlyStopping, ReduceLROnPlateau) оқу процесін оңтайландырып, артық үйренудің алдын алды. Алайда, модельдердің жалпы тұрақтылығын қамтамасыз ету үшін қосымша деректер және түсіндірмелі жақсартулар қажет.

Зерттеу барысында бірқатар кемшіліктер анықталды, соның ішінде деректер көлемінің шектеулілігі мен модель нәтижелерін түсіндірудегі қиындықтар. Бұл мәселелерді шешу үшін болашақта деректер жиынтықтарын кеңейту және модельдердің түсініктілігін арттыру бағытында қосымша зерттеулер жүргізу қажет. Жалпы алғанда, терең оқыту әдістері диабеттік ретинопатияны ерте кезеңдерде анықтауға тиімді құрал ретінде қолданыла алады және бұл әдістерді клиникалық практикаға енгізу үшін қосымша зерттеулер мен тәжірибелік тексерулер қажет.

ӘДЕБИЕТТЕР

1 Mehmet, Şahin., Omer, Faruk, Beyca. Diabetic Retinopathy Diagnosis with Image Processing (2024). <https://doi.org/10.1109/siu61531.2024.10601116>.

2 Nikita, S., Demin., N., Yu., Ilyasova., Rustam, Paringer. Automatic Selection of the Optimal Zone for Laser Exposure According to the Fundus Images for Laser Coagulation (2023). <https://doi.org/10.18287/jbpe23.09.040308>.

3 Agnieszka, Cisek., Karolina, Korycinska., Leszek, Pyziak., Marzena, Malicka., Tomasz, Wiecek., Grzegorz, Gruzal., K., Szmuc., Jozef, Cebulski., Mariusz, Spyra. Algorithm-based diagnostic application for diabetic retinopathy detection. arXiv.org. abs/2312.00529 (2023). <https://doi.org/10.48550/arxiv.2312.00529>.

4 Balaji, S., Karthik, B., Gokulakrishnan D. Prediction of Diabetic Retinopathy using Deep Learning with Preprocessing. EAI Endorsed Transactions on Pervasive Health and Technology, 10 (2024). <https://doi.org/10.4108/eetpht.10.5183>.

5 Rubina, Sarki, Khandakar, Ahmed, Hua, Wang... Yanchun, Zhang... Jiangang, Ma., Kate, N., Wang. Image Preprocessing in Classification and Identification of Diabetic Eye Diseases. Data Science and Engineering, 6(4), 1–17 (2021). <https://doi.org/10.1007/S41019-021-00167-Z>.

6 Onesinus, Saut, Parulian, Jufriadif, Na'am. Advanced Filtering and Enhancement Techniques for Diabetic Retinopathy Image Analysis. Journal Medical Informatics Technology, 69–75 (2024). <https://doi.org/10.37034/medinftech.v2i3.40>.

7 Imtiyaz, Ahmad, Vibhav, Prakash, Singh, Manoj, Madhava, Gore. Detection of Diabetic Retinopathy Using Discrete Wavelet-Based Center-Symmetric Local Binary Pattern and Statistical Features. Deleted Journal (2024). <https://doi.org/10.1007/s10278-024-01243-2>.

8 Sagar, Zanjad. Diabetic Retinopathy Detection from Retinal Images. Indian Scientific Journal Of Research In Engineering And Management, 08(06), 1–5 (2024). <https://doi.org/10.55041/ijrem35218>.

9 Mage, Usha, U., Subburaj, T., Hemanth, K.J. Innovative AI Solution for Diabetic Retinopathy Health. International Journal of Advanced Research in Science, Communication and Technology, 10–13 (2024). <https://doi.org/10.48175/ijarsct-19003>.

10 Kalindhu Navanjana De Silva, K., Sanduni Kumari Lanka Fernando, T., Lakshan Sandaruwan Jayasinghe, L.D., Dinuka Sandaruwan Jayalath, M.H., Kasun Karunanayake, B.A.P. Madhuwantha. Towards Accurate Detection of Diabetic Retinopathy Using Image Processing and Deep Learning. International Journal of Advanced Computer Science and Applications, 15(9) (2024). <https://doi.org/10.14569/ijacsa.2024.0150986>.

11 Mahesh, M.R., Rohith, S. Analysis of Diabetic Retinopathy in Retinal Fundus Images, 1–5 (2024). <https://doi.org/10.1109/ickecs61492.2024.10616880>.

12 Balaji., S., Karthik, B., Gokulakrishnan, D. Prediction of Diabetic Retinopathy using Deep Learning with Preprocessing. EAI Endorsed Transactions on Pervasive Health and Technology, 10 (2024). <https://doi.org/10.4108/eetpht.10.5183>.

13 Shrimali, S. Image Filtering and Utilization of Deep Learning Algorithms to Detect the Severity of Diabetic Retinopathy. International Conference on Communication Systems and Network Technologies, 202–207 (2023). <https://doi.org/10.1109/CSNT57126.2023.10134680>.

14 Mamyrbayev, O.Zh., Pavlov, S.V., Momynzhanova, K.R. Development of an expert system based on fuzzy logic for early diagnosis of glaucoma. Herald of the Kazakh-British technical university, 21(3), 37–47 (2024) (In Kazakh). <https://doi.org/10.55452/1998-6688-2024-21-3-37-47>.

15 Imane, Mehidi. Pre-processing Techniques: A Review for Retinal Image Segmentation (2024). <https://doi.org/10.1109/iceecac61226.2024.10576331>.

16 Thirawat, Saengarun, Niwat, Angkawisittpan, Adisorn, Nuan-On., Guifen, Lyu, Xiaoli, Fu., Sivarit, Sultornsanee. The Development for Diabetic Retinopathy Lesion Detection on Fundus Photography (2024). <https://doi.org/10.1109/ieecon60677.2024.10537869>.

17 Surmayanti, Sumijan, Sumijan, Saiful, Bukhori. Identification of Diabetic Retinopathy Using the Extraction Method on Fundus Images, 333–340 (2024). <https://doi.org/10.1109/icoict61617.2024.10698276>.

18 Pranoti, Nage, Sanjay, Shitole. Analysis of Image Processing Techniques for Retinal Fundus Images in Detection of Ophthalmic Disease Screening, 1–4 (2023). <https://doi.org/10.1109/icccnt56998.2023.10306704>.

19 Vignesh., R., Muthukumaran, N. Detection of Diabetic Retinopathy Image Analysis using Convolution Graph Neural Network, 921–929 (2023). <https://doi.org/10.1109/ICICT57646.2023.10134388>.

20 Abisha, J., Jeba, P.S. Detection of Diabetic Retinopathy by Using Convolutional Neural Network (2023). <https://doi.org/10.59544/mine9624/ngcesi23p89>.

21 Farzaneh, Nikbakhtsarvestani, Shahryar, Rahnamayan, Mehran, Ebrahimi. Opposition-based Multi-Objective ADAM Optimizer (OMAdam) for Training ANNs, 1–10 (2024). <https://doi.org/10.1109/cec60901.2024.10612083>.

22 Alexandra de Raadt, Matthijs J. Warrens., Roel Bosker, Henk, A.L., Kiers. Kappa Coefficients for Missing Data. Educational and Psychological Measurement, 79(3), 558–576 (2019). <https://doi.org/10.1177/0013164418823249>.

¹Yesmukhamedov N.S.,

PhD student, ORCID ID: 0009-0006-0652-3082,

e-mail: yesmukhamedov.yeskendyr@gmail.com

¹Sapakova S.Z.,

PhD, Associate Professor, ORCID ID: 0000-0001-6541-6806,

e-mail: sapakovasz@gmail.com

²Kozhamkulova Zh. Zh.,

PhD, Associate Professor, ORCID ID: 0000-0002-3901-7809,

e-mail: zh.kozhamkulova@au.es.kz

³Daniyarova D.,

PhD, Associate Professor, ORCID ID: 0000-0002-6833-6462,

e-mail: duriya.daniyarova@mail.ru

¹Armankyzy R.,

MSc, Lecturer, ORCID ID: 0009-0009-2236-559X,

e-mail: armankyzyrenata@gmail.com

¹International University of Information Technologies, Almaty, Kazakhstan

²Almaty University of Power Engineering and Telecommunications after Gumarbek Daukeev,
Almaty, Kazakhstan

³International Educational corporation, Almaty, Kazakhstan

METHODS FOR PRE-PROCESSING AND ANALYSIS OF FUND IMAGES FOR DETECTION OF DIABETIC RETINOPATHY

Abstract

This work is intended to study methods for pre-processing and analysis of fundus images for the detection of diabetic retinopathy. Diabetic retinopathy (DR) is a common eye disease in patients with diabetes, and its early diagnosis allows you to prevent vision loss. During the study, modern methods for processing and analyzing fundus images were used, including the EfficientNetB0 architecture based on Deep Learning. Image augmentation (rotation, scaling, cropping, contrast enhancement) and normalization methods were introduced for pre-processing. When using the EfficientNetB0 architecture, two approaches were tested: training the base layers and additional adaptation (fine-tuning) by opening the upper layers. The results were evaluated by metrics. The precision for the test set in the first method was 65%, and for the second method 75%. The accuracy of the validation set in the first method was 63%, and in the second method it reached 71%. The recall metric showed 60% for the test set in the first method, and 74% in the second method. In general, the fine-tuning method showed high performance. The use of these methods allows to improve the quality of image processing and classification for effective diagnosis of diabetic retinopathy.

The novelty of the study is the analysis of various methods of using and adapting the highly efficient EfficientNetB0 architecture. The results obtained allow to improve the quality of automated systems in DR diagnostics and increase the energy efficiency of the model. The proposed methods have high potential for early detection of eye diseases.

Keywords: diabetic retinopathy, machine learning, data analysis, laser coagulation, fundus images.

¹Есмухамедов Н.С.,

докторант, ORCID ID: 0009-0006-0652-3082,
e-mail: yesmukhamedov.yeskendyr@gmail.com

¹Сапакова С.З.,

к.ф.-м.н., ассоциированный профессор, ORCID ID: 0000-0001-6541-6806,
e-mail: sapakovasz@gmail.com

²Кожамкулова Ж.Ж.,

PhD, ассоциированный профессор, ORCID ID: 0000-0002-9780-9767,
e-mail: zh.kozhamkulova@aukes.kz

³Даниярова Д.Р.,

к.т.н., ассоциированный профессор, ORCID ID: 0000-0002-6833-6462,
e-mail: duriya.daniyarova@mail.ru

¹Арманкызы Р.,

м.т.н., лектор, ORCID ID: 0009-0009-2236-559X,
e-mail: armankyzyrenata@gmail.com

¹Международный университет информационных технологий, г. Алматы, Казахстан

²НАО «АУЭС им. Г. Даукеева», г. Алматы, Казахстан

³Международная образовательная корпорация, г. Алматы, Казахстан

ОПРЕДЕЛЕНИЕ ДИАБЕТИЧЕСКОЙ РЕТИНОПАТИИ НА ОСНОВЕ ПРЕДВАРИТЕЛЬНОЙ ОБРАБОТКИ И АНАЛИЗА СНИМКОВ СЕТЧАТКИ ГЛАЗА

Аннотация

Данная работа посвящена исследованию методов предварительной обработки и анализа снимков сетчатки для определения диабетической ретинопатии (ДР). Диабетическая ретинопатия – это распространенное заболевание глаз у пациентов с диабетом, и ранняя диагностика этого заболевания помогает предотвратить потерю зрения. В ходе исследования были использованы современные методы обработки и анализа изображений, включая архитектуру EfficientNetB0, построенную на основе глубокого обучения. Для предварительной обработки были применены методы аугментации изображений (поворот, масштабирование, обрезка, повышение контраста) и нормализация. В ходе использования архитектуры EfficientNetB0 было испытано два подхода: заморозка базовых слоев модели и дообучение верхних слоев (fine-tuning). Результаты были оценены по меткам. В первом подходе точность (precision) на тестовой выборке составила 65%, во втором – 75%. Точность на валидационной выборке в первом случае составила 63%, во втором – 71%. Метрика обнаружения (recall) на тестовой выборке в первом подходе показала 60%, во втором – 74%. В целом методы fine-tuning показали лучшие результаты. Использование данных методов позволяет повысить качество обработки изображений и классификации для эффективной диагностики диабетической ретинопатии. Новизна исследования заключается в применении высокоэффективной архитектуры EfficientNetB0 и сравнении различных подходов к дообучению модели. Полученные результаты могут способствовать улучшению качества автоматизированных систем диагностики ДР и повышению энергоэффективности моделей. Предложенные методы обладают высоким потенциалом для раннего выявления глазных заболеваний.

Ключевые слова: диабетическая ретинопатия, машинное обучение, анализ данных, лазерная коагуляция, фундус-снимки.

Мақаланың редакцияға түскен күні: 19.12.2024

UDC 004.056.55

<https://doi.org/10.55452/1998-6688-2025-22-4-131-142>

^{1*}**Gorlov L.V.,**

PhD student, ORCID 0000-0002-8208-4716,

e-mail: lev.gorlov@gmail.com

²**Seilova N.A.,**

Cand. Tech. Sc., Associate Professor, ORCID: 0000-0003-3827-179X,

e-mail: nseilova@iitu.edu.kz

³**Okhrimenko T.A.,**

Cand. Tech. Sc., Associate Professor, ORCID: 0000-0001-9036-6556,

e-mail: t.okhrimenko@nau.edu.ua.

¹Al-Farabi Kazakh National University, Almaty, Republic of Kazakhstan

²International University of Information Technologies, Almaty, Republic of Kazakhstan

³National Aviation University, Kiev, Ukraine

ON THE DIFFUSION LAYER GENERATION METHOD

Abstract

This paper presents an automated method for generating the parameters of linear functions used in the diffusion layer of block symmetric encryption algorithms. The focus is on designing linear layers constructed solely from cyclic shift operations and bitwise XORs, which are both efficient and hardware-friendly. Such layers play a critical role in achieving strong diffusion, a fundamental cryptographic requirement. The proposed method evaluates candidate configurations by exhaustively enumerating shift values, calculating their branch number, and assessing their avalanche characteristics. A set of quantitative diffusion metrics is introduced to guide the selection process, including single- and multi-round avalanche effects and activation rates at the byte level. An aggregated quality function is formulated to allow comparative assessment. The developed software tool identified optimal shift parameters for 128-bit blocks processed as four 32-bit words, achieving a branch number of 5 with only 12 XOR operations. The proposed approach contributes to the practical synthesis of lightweight and secure cryptographic primitives suitable for both classical and constrained platforms.

Keywords: Block symmetric encryption algorithm, linear layer, branch number, avalanche effect.

Introduction

The purpose of this paper is to show methods for generating shift coefficients for the linear layer of the architecture presented in "Linear Layer Architecture Based on Cyclic Shift and XOR" as well as optimization methods that reduce the effort required to enumerate candidate coefficients by orders of magnitude. Preliminary conclusions about cryptographic qualities and computational complexity were given earlier in [1]. General approaches to balancing the function – such as adjusting the number and length of words and enhancing it with additional iterations – were also discussed, and a comparison with the AES algorithm's linear layer was provided.

Evaluation of cryptographic resistance against linear and differential attacks is also beyond the scope of this paper and will be published in the future.

Modern block encryption algorithms are built on the principles formulated by Claude Shannon: the principle of confusion and diffusion [2, 3]. Diffusion is achieved, in particular, by using linear transformations with good diffusion characteristics, which ensure rapid propagation of changes across all bits of the output vector.

One of the key metrics used to evaluate the diffusion ability of a linear function is the branch number. This indicator reflects the minimum possible number of different words at the input and output of a linear transformation:

$$B(L) = \min_{x \neq 0} \{w(x) + w(L(x))\} \quad (1)$$

where w is Hamming weight. The sizes of words and their location within the converted data block may have different values depending on the specific architecture and operations used in the encryption algorithm. As a general case we will consider words as the input of the nonlinear function of the algorithm, the size of which is typically equal to 8 bits.

In both classical and lightweight encryption algorithms, the linear diffusion layer is often the most computationally intensive component [4–6]. Linear operations based on maximum distance codes (MDS) are widely used to perform linear transformations in various encryption algorithms, including AES due to the strictly proven branch number [7–11]. However, one can consider alternative functions that may not have the largest possible branch number, but have sufficiently strong cryptographic properties with significantly higher performance [12, 13].

Schemes using cyclic shift and XOR on words of different lengths are already used in lightweight encryption algorithms to mix the bits of the algorithm's internal state [14–16]. These transformations do not require reversibility and can include additional operations. For example, the ARX (Add-Rotate-XOR) architecture includes modular addition modulo 2^y , where y represents the word length.

There are also algorithms for constructing transformations equivalent to multiplication by MDR matrices using cyclic shift and bitwise addition operations. The papers provide a theoretical basis for obtaining such operations analytically, without exhaustive search, for a block of four words [17–19]. The linear operation constructed in this way has a branch number of five, which is the maximum achievable for an operation involving four subblocks.

At present, no information has been found on the possibility of generating larger matrices with a higher branch number using the same method.

Materials and methods

The goal is to construct a linear function that facilitates rapid and uniform propagation of input changes across all output bits, thereby preventing the construction of effective differential or linear attacks [2]. The main objective of this paper is to formalize the approach to assessing the quality of a cipher's linear layer and to propose universal criteria that allow quantitative comparison of various designs in order to generate a linear function with characteristics close to optimal.

Since the calculation of the branch number is a computationally complex task and cannot be used as a criterion in enumeration problems, for the initial assessment of the efficiency of the linear function $L: \mathbb{F}_2^n \rightarrow \mathbb{F}_2^n$ we propose to consider a set of five complementary characteristics:

1. Average avalanche effect per round (denoted as criterion A_1) is the average number of bits at the output of the function that change when one bit at the input is inverted. This metric characterizes local diffusion and shows how quickly the initial change spreads along the output vector. In the ideal case, $A_1 \approx n/2$, but for a lightweight function this result is hardly achievable.

Average number of bytes activated per round (B_1). Given the influence of the nonlinear layer operating on bytes it is reasonable to consider how many bytes will change in each round of the algorithm when at least one bit of the input changes. This parameter is most closely related to the branch number used in the wide trail design strategy [3].

Average avalanche effect over two rounds (A_2) – similar to criterion A_1 , but measured over two consecutive rounds (ignoring the nonlinear layer between those rounds). This characteristic better reflects the cumulative diffusion effect during the operation of the encryption algorithm.

Average number of activated bytes over two rounds (B_2).

Minimum number of output bits activated by a single active input bit (B_{min}):

$$B_{min} = \min_{x:w(x)=1} w(L(x)) \quad (2)$$

This parameter defines the worst-case diffusion scenario and is also indirectly related to the branch number. Its increase directly contributes to increasing the cipher's resistance to differential and linear cryptanalysis.

For a systematic assessment of the linear function based on the above parameters, it is proposed to use an aggregated metric in the form of a weighted criterion:

$$Score(L) = \alpha \cdot \frac{A_1}{n} + \beta \cdot \frac{A_2}{n} + \gamma \cdot \frac{B_1}{n} + \delta \cdot \frac{B_2}{n} + \theta \cdot \frac{B_{min}}{n} \quad (3)$$

here $\alpha, \beta, \gamma, \delta, \theta \in \mathbb{R}_{>0}$ – normalized weighting factors determined by design priorities. In this case, all values are normalized by n to bring them to a single scale. For example, when focusing on cryptographic resistance, one can choose $\gamma > \alpha, \beta$, while in problems of optimizing the diffusion rate in a small number of rounds, one can increase the significance of β .

Thus, the formulated criterion allows one to compare various linear functions taking into account both average and extreme behavior, and can be used as a heuristic in the automated synthesis of diffusion layers.

It is required to generate a linear function that is implemented by computationally easy operations and has a branch number of at least 5 for a data block of 16 words of 8-bit size [21–24]. These values are taken by analogy with the AES algorithm and other common encryption algorithms and are de facto standard for block symmetric encryption algorithms [25–30]. The following considerations were taken into account in its development:

A given branch number can be achieved by performing an operation on at least 4 words. A 128-bit data block is extremely conveniently divided into 4 words of 32 bits, since most modern computing platforms implement operations on 32-bit words in an optimal way [4].

To ensure mutual influence of bits located at different positions, a cyclic shift operation is necessary, and to change the values of bits, the XOR operation is optimal.

For inversion to be possible, the resulting linear operation must be bijective.

The influence of words should be mutual, so at each iteration either one word should overlap all the others, or vice versa, which seems easier to implement.

Taking these requirements into account, a linear operation over an arbitrary number of words of arbitrary (identical) length is constructed, allowing for balancing diffusion properties and computational complexity [1].

The resulting linear function f consists of several successive operations g over a set of words $a = \{a_0, \dots, a_{u-1}\}$ and is expressed by the following formula:

$$g_{(j)}: a^{(n+1)} = g_{(j)}(a^{(n)}) \Leftrightarrow a_k^{(n+1)} = \begin{cases} \bigoplus_{i=0}^{u-1} (a_{i+j \bmod u}^{(n)} \lll C_i), & k = j \\ a_k^{(n)}, & k \neq j \end{cases} \quad (4)$$

The operation f , which changes all the words of the processed block, can be represented as follows:

$$f(a) = g_{(0)}(a) \circ \dots \circ g_{(u-1)}(a) \quad (5)$$

The value of the shift constants $C_{0 \dots u-1}$ is of decisive importance for the efficiency of the function. Clearly, not all sets of values lead to the desired distribution of mutual influence of bits.

Proving that the branch number of a function is 5 is a computationally difficult problem [32–34]. For all possible inputs of a linear function, it is necessary to enumerate all combinations of $d = 1 \dots 4$ active bytes and make sure that at least $5 - d$ bytes are activated at the output of the function. There are 2^{128} inputs to the linear function, each of which would require iterating over $2^8 \times 16$ variants of the difference in the input in a single byte, for which the difference in the output must be at least 4 bytes, $2^{16} \times 16 \times 15$ two-byte input differences, and so on, for a total of $\sum_{d=1}^4 2^{8d} \times A_d^{16} \approx 2^{48}$ operations. Thus, solving the problem as a whole using brute force requires about 2^{176} operations, which is practically impossible.

However, since the function is linear, the influence of each input bit on the output bits can be found independently of the other bits. This property allows for a preliminary estimate of the function coefficients that is not directly related to the branch number and avalanche effect, but is correlated with them to some extent.

The advantage of this method for estimating the shift coefficients of a linear function is its low computational complexity (namely, 2^7 linear for one set of coefficients and 2^{22} for the algorithm as a whole without taking into account further optimizations) and, as a consequence, the possibility of a complete enumeration of the entire set of coefficients.

The main drawback is the lack of a specific resulting value of the branch number. The avalanche effect for some combination of input bits may also be less than the number of output bits activated by a single non-zero input bit due to the peculiarity of the XOR function: some activated bits may overlap each other and ultimately have no effect on the output. Calculating all possible combinations of mutual overlaps to estimate the branch number has the same computational complexity as the full search above, namely 2^{176} operations.

However, it is possible to optimize both the evaluation of the primary properties of the function and the verification of the branch number of the resulting function.

First, enumerating all input values of a linear function does not make sense, since the changed input bits will affect the same output bits regardless of the values of the remaining (fixed) input bits. This allows for direct enumeration of differences, abstracting from enumerating all input combinations. This reduces the enumeration by 2^{128} times.

The next property that allows optimizing the calculation of the lower bound of the branch number of a linear function is its cyclicity. In particular, due to the cyclic structure of the transformation, different input vectors obtained by cyclically shifting each other will yield identical (with the same relative shift) output differences. Thus, the set of all configurations with 4 active blocks is divided into orbits by the action of a group of cyclic shifts, and it is sufficient to consider only one representative from each orbit. Since in the case under consideration not only the number of active bits is important, but also the number of active bytes, it makes sense to exclude from the search input blocks that are equivalent in value and shifted relative to each other by a number of bits multiple of 8. This significantly reduces the volume of the search and allows focusing on unique structures of input differences from the point of view of the output. This approach is consistent with the general principles of analyzing the diffusion properties of linear transformations and is widely used in assessing the differential characteristics of cryptographic primitives with a regular architecture [35–36].

It should be noted that this equivalence is not preserved during a cyclic shift of the block, i.e. when transferring the most significant bits of one word to the positions of the least significant ones of another word, since the influence of the bits, propagated by a linear operation, depends on the word in which they are located in the block.

In practice, this allows us not to go through all possible options for placing active bytes. It is sufficient to consider only those in which the least significant byte of each row is active. For example, let us consider two input blocks that give outputs equivalent to the shifts:

		B	A
	D	C	

	B	A	
D	C		

Figure 1 – Inputs of a linear function equivalent by branching

Here the values of the remaining bytes are 0, therefore they have no effect on the activation of the output bits.

This optimization allows us to strictly calculate the lower bound of the branch number in $\sum_{b=1}^4 2^{8b} A_{b-1}^{16-b} \approx 2^{43}$ operations. There are certainly methods that can reduce the search even further, for example, by eliminating other subsets of bits through other types of equivalence, but with the current approach, the computation can be completed within a few minutes of computing on a modern processor.

The same approach can be applied to the evaluation of the primary indicators of a function in order to reduce the search space for single active-bit inputs. Indeed, all output options for different bits of one input word are cyclic shifts of each other. And similarly, bits of different input words generate different combinations of outputs. In this case, it is necessary to estimate the number of both active bits and active bytes at the output of the function, therefore, to fully find the values of interest, it is necessary to sort through 32 input options (all single bits of one of the bytes for each word).

Results and discussion

The desired linear transformation is constructed using bitwise cyclic shifts and XOR operations on four 32-bit words and has the form (in the C programming language):

```
dout[0] = din[0] ^ ROTL(din[1], c[0]) ^ ROTL(din[2], c[1]) ^ ROTL(din[3], c[2]);
dout[1] = din[1] ^ ROTL(din[2], c[0]) ^ ROTL(din[3], c[1]) ^ ROTL(dout[0], c[2]);
dout[2] = din[2] ^ ROTL(din[3], c[0]) ^ ROTL(dout[0], c[1]) ^ ROTL(dout[1], c[2]);
dout[3] = din[3] ^ ROTL(dout[0], c[0]) ^ ROTL(dout[1], c[1]) ^ ROTL(dout[2], c[2]);
```

To find the coefficients with the desired properties, it is necessary to perform a complete enumeration of all possible triplets of shift values $c_i \in [0,31], i = \{1,2,3\}$.

For each triple of coefficients, the average diffusion effect and the minimum number of activated output bytes must be determined over all possible single-bit input differences. The avalanche effect of a lightweight linear function extends to the entire processed block in several rounds, therefore, when selecting the coefficients of the functions, the avalanche effect was also evaluated over two rounds (i.e., the linear layer was applied twice). In practice, when using sufficiently high-quality non-linear blocks this effect is significantly amplified, since when changing at least one bit of the output of a linear function, an average of 4 bits related to the same substitution block will be changed [37].

For each tested set of coefficients:

- 1) All possible inputs with a single one bit (the other bits are 0) are iterated over. Thus, all the following steps are performed 32 times for each combination of shift values;
- 2) The block is fed to the input of the doubled linear function;
- 3) The number of nonzero bytes at the output is counted;
- 4) Results for any coefficient set with fewer nonzero output bytes than the current best are discarded;

5) The average value of the number of activated bits is found. If it is greater than the best current result, then the set of constants is accepted as the best current result.

Point 4 of the algorithm allows significant optimization of the search, since at the early stages the values that give bad special cases are filtered out.

Practical selection gave many results with similar characteristics. In particular, 7217 sets of coefficients were found, in which any single active input bit activates at least 6 output bits.

To select the coefficients of a linear function with the corresponding cryptographic criteria, it is necessary to select the weight coefficients $\alpha, \beta, \gamma, \delta, \theta$ from formula (3).

The choice of these coefficients logically follows from the significance of the corresponding criteria imposed on the linear function being developed, which is beyond the scope of this paper. Below, some initial provisions for selecting coefficients without strict justifications will be presented.

Apparently, the coefficient θ does not play a big role, since it is always advisable to choose a function that activates the maximum number of bits with a single active input bit, i.e. to consider only the coefficients that generate functions with B_{min} equal to 7.

In the presence of a nonlinear block with a powerful avalanche effect, the weights β and δ can be significantly reduced or even set to zero.

For an algorithm with a large number of rounds, on the contrary, it makes sense to rely on two-round indicators, reducing α and γ .

The results of the program implementing this algorithm with different weighting coefficients are presented below:

Table 1 – Examples of generation results

№	α	β	γ	δ	θ	A_1	A_2	B_1	B_2	B_{min}	c_0	c_1	c_2	Score
1	1	0	1	0	1	8.75	15.75	8.12	12.12	7	1	10	15	7.958
2	0	1	0	1	1	8.75	22.25	7.56	13.88	7	1	17	14	14.375
3	1	0.5	2	0.8	1	8.75	21.75	7.97	14.12	7	2	26	25	10.163
4	1	2	1	2	1	8.75	22.25	7.66	13.88	7	1	18	14	13.665

To check the functionality of the function that calculates the branch number of a function of the form (5), a test was conducted in which an optimized enumeration method was used to search for input values of up to 4 active bytes that yield the minimum output weight for a function variant with shift coefficients {1, 10, 15}.

The search was performed in parallel in several threads covering subranges of input byte values. To analyze all possible values of three of the four bytes for all possible values of the first byte of the first word, four parallel instances of the software were run, resulting in the following worst-case outputs:

Table 2 – Worst cases for first word byte enumeration

№	Diapason	BI	Input/output (hexadecimal)
1	0..63	5	00 09 00 00 80 00 00 00 40 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 40 00 00 00 00 00 01 00
2	64..127	5	48 00 00 00 04 00 00 00 02 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 02 00 00 00 00 08 00 00
3	128..191	5	90 00 00 00 08 00 00 00 04 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 04 00 00 00 00 10 00 00
4	192..255	5	d8 00 00 00 0c 00 00 00 06 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 06 00 00 00 00 18 00 00

For an input difference in the first byte of the second word, the outcome was even better, since

the second word's bits influence the first word from the very first iteration of the linear function:

Table 3 – Results of enumeration for bytes of the second word

№	Diapason	BI	Input/output (hexadecimal)
1	0..127	7	00 00 00 00 00 00 00 00 10 00 00 00 00 00 00 80 00 00 00 00 30 00 00 00 11 c0 00 00 00 42 00 83
2	127..255	5	00 00 00 00 80 00 00 00 08 00 00 00 00 00 00 00 00 00 00 00 90 00 00 00 08 40 02 00 00 32 00 09

For the third and fourth words, no cases were found in which less than 7 bytes were activated in two adjacent rounds.

Thus, the minimum sum of the input and output weights of a one-round linear function is 5, and therefore the branch number of a linear function of the form (5) with coefficients {1,10,15} is 5.

Comparative analysis of linear layer design

Architecture. Traditional diffusion layers in block ciphers often use matrices from maximum distance separable (MDS) codes (e.g., the MixColumns matrix in AES) to achieve linear mixing [5]. These MDS-based designs perform linear transformations (usually over finite fields) with a strictly proven diffusion property (maximal branch number) [3]. In contrast, the proposed method uses only cyclic bit shifts and XOR operations – a shift/XOR-based linear layer similar to ARX (Add-Rotate-XOR) constructions [6]. This means that instead of heavy finite-field multiplications or complex linear combinations, diffusion is achieved through simple rotations and XORs. Such an approach is hardware-friendly and simplifies implementation, while still maintaining the linear invertibility required for cipher rounds. The shift/XOR design aligns with many modern lightweight ciphers, which favor simple operations for efficiency, distinguishing our method as a purely bitwise-linear alternative to classical MDS matrices.

Branch number. MDS-based layers are valued because they guarantee the maximum possible branch number for a given matrix dimension [3]. For example, an $n \times n$ MDS matrix ensures the highest minimum number of active input/output words (the branch number) by design. Shift/XOR-based layers generally do not inherently guarantee this theoretical maximum; they often achieve near-maximum diffusion in practice. However, a key result of this work is that through careful parameter selection we can attain a branch number equal to the MDS optimum. In fact, for a block of four 32-bit words (128-bit state), our optimized rotation–XOR construction reaches a branch number of 5, which is the provable maximum for any 4-word linear transformation [7]. This demonstrates that even without using an MDS code, the proposed method can match the diffusion strength of an MDS matrix in this scenario. While classical MDS matrices provide a mathematical proof of maximum branch number, our approach shows that near-MDS or even MDS-equivalent branch numbers are achievable with a shift/XOR design – combining strong diffusion with the benefits of a lightweight operation. This result underscores the novel contribution of attaining high branch number without resorting to traditional MDS constructs.

Designing complexity. Designing an MDS-based linear layer often leverages algebraic methods or known constructions from coding theory [7]. In general, finding an MDS matrix for a given size can be done via mathematical criteria or combinatorial search, but many such matrices are known or can be constructed using finite-field techniques. By contrast, identifying an optimal shift/XOR configuration required brute-force search over rotation constants. The search space is enormous – yet our method introduces a highly optimized enumeration strategy to make this design problem tractable. We exhaustively evaluate candidate shift combinations, but we guide the search with quantitative diffusion metrics and pruning heuristics. Notably, we reduce the complexity of the search by several orders of magnitude through intelligent heuristics and early elimination of suboptimal candidates. This

optimized search procedure (with metrics like single- and two-round avalanche criteria) effectively navigates the design space, in contrast to a naive brute force. The result is a practical computational method that finds excellent diffusion layers (even matching MDS-level branch number) without an explicit algebraic formula. Thus, whereas MDS-based designs benefit from direct theoretical constructions, the proposed shift/XOR design required developing an automated search framework – an effort justified by the significant performance gains of the final design.

Implementation complexity. A well-known drawback of many MDS-based layers is their implementation cost. Linear diffusion via an MDS matrix (especially over bytes) often becomes the most computationally intensive component of the cipher. Implementing finite-field multiplications or look-up tables for an MDS matrix can incur a high gate count, many XOR operations, or additional memory for tables [39–40]. In contrast, the shift/XOR approach uses only primitive CPU operations (bit rotates and XORs), leading to significantly lower complexity in both hardware and software. The proposed linear layer achieves its diffusion with a remarkably small number of operations – for a 128-bit block we use only 12 XOR operations in total (along with shifts) to attain branch number 5. There are no multiplications or large tables needed, which translates to fewer logic gates and faster execution on microcontrollers. This efficiency makes the method highly suitable for lightweight and constrained devices, where minimizing computational overhead is crucial. In summary, while an MDS-based layer might maximize diffusion at the cost of more intensive computation (often requiring careful optimization or special implementation tricks), the shift/XOR-based layer offers a far more lightweight solution. The novel contribution here is demonstrating equivalent diffusion strength with drastically reduced operation count, thereby improving encryption performance on resource-limited platforms without sacrificing security.

Flexibility and Adaptability. MDS code-based layers are typically tied to specific block and word sizes – for example, AES’s 4×4 byte matrix is fixed to 128-bit blocks (16 bytes), and designing a new MDS matrix for a larger state or different word size can be non-trivial. Each change in dimensions often requires finding new matrices that preserve the MDS property, a process that can become a research problem in itself for higher dimensions. In contrast, the proposed shift/XOR generation method is inherently more flexible. Our framework can be extended to an arbitrary number of words and word lengths by selecting appropriate rotation coefficients. In principle, the same approach can synthesize a diffusion layer for different block sizes (e.g., 64-bit blocks or 256-bit blocks) by adjusting the word partitioning and rerunning the automated search for optimal shifts. This scalability is enabled by the general nature of rotations and XORs, which are not locked to byte-oriented operations or a fixed field size. Furthermore, the method is adaptable in terms of design goals: we can tune the weightings of various diffusion metrics to prioritize certain criteria. For instance, if the cipher uses a very strong nonlinear layer (S-box), the search can de-emphasize single-round diffusion and focus on multi-round effects, or vice versa. The authors explicitly allow adjusting parameters (like weighting factors for avalanche criteria) to suit different scenarios – e.g., ignoring some diffusion criteria when an S-box already provides good mixing, or emphasizing two-round diffusion for ciphers with many rounds. This level of adaptability is a distinct advantage of the proposed approach. In summary, unlike a one-size-fits-all MDS matrix, our shift/XOR-based generation method can scale and adapt to various block architectures and security requirements with relative ease. This flexibility in accommodating different word sizes and balancing performance vs. diffusion trade-offs highlights the novel contribution of our method in the landscape of linear layer design.

Conclusion

The study introduces a scientifically novel and practically applicable method for synthesizing high-quality linear diffusion layers without resorting to pre-designed MDS codes. The originality of the approach lies in the integration of an optimized exhaustive search with mathematically grounded

pruning techniques, which drastically reduce the search space while preserving the ability to obtain provable branch numbers. This makes it possible, for the first time in the context of purely shift- and XOR-based constructions, to confirm by direct computation the achievement of the theoretical maximum branch number of 5 for a 4-word, 32-bit architecture.

The authors' personal contributions include formulating a set of quantitative diffusion metrics that combine average-case and worst-case behavior, defining a unified weighted evaluation criterion, and implementing a specialized software tool that automates both coefficient generation and branch number verification. Additionally, the authors developed and validated new equivalence-class-based optimization strategies that reduce the combinatorial complexity of branch number checking from impractical brute-force scales to a range feasible on standard modern processors.

This work not only delivers a concrete high-performance linear transformation meeting stringent cryptographic criteria but also provides a reusable methodology for designing diffusion layers adaptable to diverse block sizes, word lengths, and implementation constraints. The proposed generation framework can serve as a foundation for further research on lightweight cipher design, particularly in scenarios where hardware simplicity, provable diffusion strength, and adaptability are equally critical.

REFERENCES

- 1 Gorlov, L. Iavich, M. and Bocu, R. Linear Layer Architecture Based on Cyclic Shift and XOR. *Symmetry*, 15, (2023).
- 2 Debranjana Pal, Vishal Pankaj Chandratreya, Abhijit Das, Dipanwita Roy Chowdhury. Modeling Linear and Non-linear Layers: An MILP Approach Towards Finding Differential and Impossible Differential Propagations, arXiv:2405.00441 (2024).
- 3 Daemen, J. and Rijmen, V. The wide trail design strategy. In *Proceedings of the Cryptography and Coding: 8th IMA International*, 2001, pp. 222–238.
- 4 Wikipedia. Word (computer architecture). URL: [https://en.wikipedia.org/wiki/Word_\(computer_architecture\)](https://en.wikipedia.org/wiki/Word_(computer_architecture)).
- 5 Daemen, J. and Rijmen, V. *The Design of Rijndael*. Springer, 2002.
- 6 Nir, Y. and Langley, A. ChaCha20 and Poly1305 for IETF Protocols. RFC 7539, (2018). URL: <https://datatracker.ietf.org/doc/html/rfc7539>. [Accessed 9.05. 2025].
- 7 Joshi, D. A Note on Upper Bounds for Minimum Distance Codes. *Information and Control*, 1(3), p. 289–295 (1958).
- 8 Shannon, C. *Mathematical Theory of Cryptography*, Bell Labs (1945).
- 9 Shannon, C. *Communication Theory of Secrecy Systems*, Bell System Tech. J. (1949).
- 10 Wu, S., Wang, M., and Wu, W. Design of lightweight linear diffusion layers from near-MDS matrices. *International Association for Cryptologic Research*. URL: <https://eprint.iacr.org/2017/195.pdf>. [Accessed 9.05.2025].
- 11 Dinu, D., Perrin, L., Udovenko, A., Velichkov, V., Großschädl, J., and Biryukov, A. Design Strategies for ARX with Provable Bounds: SPARX and LAX. *Advances in Cryptology – ASIACRYPT 2016*, pp. 484–513 (2016).
- 12 Biham, E. and Shamir, A. Differential Cryptanalysis of DES-like Cryptosystems. *J. Cryptology* (1991).
- 13 Guo, Z., Liu, R., Wu, W. and Lin, D. Direct Construction of Lightweight Rotational-XOR MDS Diffusion Layers, *International Association for Cryptologic Research*, 2016. URL: <https://ia.cr/2016/1036>. [Accessed 9. 05.2025].
- 14 Guo, Z., Liu, R., Gao, S., Wu, W. and Lin, D. Direct Construction of Optimal Rotational-XOR Diffusion Primitives, p. 169–187 (2017).
- 15 Ray, B., Douglas, S., Jason, S., Stefan, T.C., Bryan, W. and Louis, W. The SIMON and SPECK Families of Lightweight Block Ciphers. *Cryptol. ePrint Arch.*, 2013. [Online]. URL: <https://ia.cr/2013/404>. [Accessed 9. 05.2025].

- 16 FIPS. Federal Information Processing Standards Publication 197. Specification for the Advanced Encryption Standard (AES) (2001). URL: <http://csrc.nist.gov/publications/fips/fips197/fips-197.pdf>. [Accessed 9.05.2025].
- 17 Matsui, M. Linear Cryptanalysis Method for DES Cipher., EUROCRYPT (1993).
- 18 Nyberg, K. Perfect Nonlinear S-Boxes., EUROCRYPT (1991).
- 19 Schneier, B. Applied Cryptography, Wiley (1996).
- 20 Vaudenay, S. On the Need for Multipermutations., FSE (2001).
- 21 Panasenکو, S. Algoritmy shifrovaniya, BHV-Peterburg (2009).
- 22 Derkach, A. Perevirka model'nih pripushhen' u kriptanalizi ARX-shifriv, Visnik NTUU «KPI». Serija: Matematichne modeljuvannja v dizajni ta ekonomici, 2, 231–236 (2020).
- 23 Tian, Yu. & Feng, Xiutao & Li, Guangrong. On the construction of ultra-light MDS matrices, 10.48550/arXiv.2409.03298, 2024.
- 24 Gupta, Kishan & Kumar Pandey, Sumit & Samanta, Susanta. On the Construction of Near-MDS Matrices, 10.48550/arXiv.2306.12791, 2023.
- 25 Al-Nofaie, S.M., Sharaf, S., & Molla, R. Design Trends and Comparative Analysis of Lightweight Block Ciphers for IoTs, Applied Sciences, 15(14), 7740 (2025).
- 26 Konstantopoulou, Evangelia and Athanasiou, George and Sklavos, Nicolas. Review and Analysis of FPGA and ASIC Implementations of NIST Lightweight Cryptography Finalists, ACM Comput. Surv., 57, 10 (2025).
- 27 Shi Wang, Yuan Chen, Yunqing Li, Xiangyong Zeng. On construction of lightweight MDS matrices, Advances in Mathematics of Communications, 16(4) (2022).
- 28 Gupta, Kishan & Kumar Pandey, Sumit & Samanta, Susanta. On the Direct Construction of MDS and Near-MDS Matrices, 10.48550/arXiv.2306.12848, 2023.
- 29 Yogesh Kumar, Prasanna Raghaw Mishra, Susanta Samanta, Kishan Chand Gupta, Atul Gaur. Construction of all MDS and involutory MDS matrices Advances in Mathematics of Communications, 19(3), 922–941 (2025).
- 30 Samanta, S. On the counting of involutory MDS matrices. Cryptography and Communications, 2024.
- 31 Gaëtan Leurent and Clara Pernot. Design of a Linear Layer Optimised for Bitsliced 32-bit Implementation. Cryptology ePrint Archive, Paper 2023/1803, 2023.
- 32 Rishakani, A.M. Cryptographic Properties of Cyclic Binary Matrices. Advances in Mathematics of Communications, 2019.
- 33 Luong, Tran & Long, Nguyen & Vo, Bay. Efficient implementation of the linear layer of block ciphers with large MDS matrices based on a new lookup table technique, PLOS ONE, 19 (2024).
- 34 Dobraunig, Christoph & Eichlseder, Maria & Mendel, Florian & Schl  ffer, Martin, Ascon v1.2: Lightweight Authenticated Encryption and Hashing. Journal of Cryptology, 34 (2021).
- 35 Leurent, G., & Pernot, C. Design of a Linear Layer Optimised for Bitsliced 32-bit Implementation, IACR Transactions on Symmetric Cryptology, 1, 441–458 (2024).
- 36 Mishra, P.R., Kumar, Y., Samanta, S., Gaur, A. A New Algorithm for Computing Branch Number of Non-Singular Matrices Over Finite Fields, Security and Cryptography for Networks. SCN 2024. Lecture Notes in Computer Science, 14974 (2024).
- 37 Huck Bennett, Mahdi Cheraghchi, Venkatesan Guruswami & Jo  o Ribeiro. Parameterized Inapproximability of the Minimum Distance Problem over all Fields and the Shortest Vector Problem in all ℓ_p Norms, SIAM Journal on Computing, 53 (2024).
- 38 Vijay Bhattiprolu, Venkatesan Guruswami, Xuandi Ren, PCP-free APX-Hardness of Nearest Codeword and Minimum Distance, arXiv:2503.11131 (2025).
- 39 Tian, Y., Feng, X., & Li, G. On the construction of ultra-light MDS matrices. arXiv preprint (2024).
- 40 Kurt Pehlivanog  lu, Meltem & Demir, Mehmet Ali. Optimizing implementations of linear layers using two and higher input XOR gates. PeerJ Computer Science, 10 (2024).

^{1*}Горлов Л.В.,

докторант, ORCID 0000-0002-8208-4716,

*e-mail: lev.gorlov@gmail.com

²Сейлова Н.А.,

т.ғ.к., доцент, ORCID: 0000-0003-3827-179X,

e-mail: nseilova@iitu.edu.kz

³Охрименко Т.А.

т.ғ.к., доцент, ORCID: 0000-0001-9036-6556,

e-mail: t.okhrimenko@nau.edu.ua

¹Әл-Фараби атындағы Қазақ ұлттық университеті, Алматы қ., Қазақстан

²Халықаралық ақпараттық технологиялар университеті, Алматы қ., Қазақстан

³Ұлттық авиация университеті, Киев қ., Украина

СЫЗЫҚТЫҚ ҚАБАТТЫ ҚҰРУ ӘДІСІ ТУРАЛЫ

Аңдатпа

Бұл жұмыс блоктық симметриялық шифрлау алгоритмдерінің диффузиялық қабатында қолданылатын сызықтық функциялардың параметрлерін құрудың автоматтандырылған әдісін ұсынады. Негізгі назар тек циклдік ауыстыру операциялары мен биттік XOR-дан жасалған сызықтық қабаттарды жобалауға бағытталған, олар тиімді және аппараттық құралдарға ыңғайлы. Мұндай қабаттар негізгі криптографиялық талап болып табылатын күшті диффузияға қол жеткізуде маңызды рөл атқарады. Ұсынылған әдіс ауысым мәндерін толық санау, олардың тармақтарының санын есептеу және көшкін сипаттамаларын бағалау арқылы үміткер конфигурацияларын бағалайды. Таңдау процесін бағыттау үшін сандық диффузиялық көрсеткіштер жинағы енгізіледі, соның ішінде бір және көп айналымды көшкін әсерлері мен байт деңгейіндегі белсендіру жылдамдығы. Салыстырмалы бағалауға мүмкіндік беретін жиынтық сапа функциясы тұжырымдалған. Өзірленген бағдарламалық құрал төрт 32 биттік сөз ретінде өңделген 128 биттік блоктар үшін оңтайлы ауысу параметрлерін анықтады, бұл тек 12 XOR операциясымен 5 тармақ санына жетеді. Ұсынылған тәсіл классикалық және шектеулі платформалар үшін қолайлы жеңіл және қауіпсіз криптографиялық примитивтердің практикалық синтезіне ықпал етеді.

Тірек сөздер: блоктық симметриялық шифрлау алгоритмі, сызықтық қабат, тармақ нөмірі, көшкін эффектсі.

^{1*}Горлов Л.В.,

докторант. ORCID 0000-0002-8208-4716,

e-mail: lev.gorlov@gmail.com

²Сейлова Н.А.,

к.т.н., доцент, ORCID: 0000-0003-3827-179X,

e-mail: nseilova@iitu.edu.kz

³Охрименко Т.А.,

к.т. н., доцент, ORCID: 0000-0001-9036-6556,

e-mail: t.okhrimenko@nau.edu.ua

¹Казахский национальный университет им. аль-Фараби, г. Алматы, Казахстан

²Международный университет информационных технологий, г. Алматы, Казахстан

³Национальный авиационный университет, г. Киев, Украина

О МЕТОДЕ ГЕНЕРАЦИИ ЛИНЕЙНОГО СЛОЯ

Аннотация

В данной статье представлен автоматизированный метод генерации параметров линейных функций, используемых в диффузионном слое блочных симметричных алгоритмов шифрования. Основное внимание уделяется проектированию линейных слоев, построенных исключительно на операциях циклического

сдвига и побитовых операциях «исключающее ИЛИ», которые являются эффективными и аппаратно-ориентированными. Такие слои играют решающую роль в достижении сильной диффузии, фундаментального криптографического требования. Предлагаемый метод оценивает конфигурации-кандидаты путем исчерпывающего перечисления значений сдвига, вычисления количества ветвей и оценки их лавинных характеристик. Вводится набор количественных метрик диффузии для управления процессом выбора, включая одно- и многораундовые лавинные эффекты и показатели активации на уровне байтов. Сформулирована агрегированная функция качества для сравнительной оценки. Разработанный программный инструмент определил оптимальные параметры сдвига для 128-битных блоков, обрабатываемых как четыре 32-битных слова, что позволило достичь количества ветвлений 5 всего за 12 операций XOR. Предложенный подход способствует практическому синтезу легковесных и безопасных криптографических примитивов, подходящих как для классических, так и для ограниченных платформ.

Ключевые слова: алгоритм блочного симметричного шифрования, линейный слой, индекс ветвлений, лавинный эффект.

Article submission date: 19.06.2025

ӨОЖ 05.13.10
ФТАХР 21.21.29

<https://doi.org/10.55452/1998-6688-2025-22-4-143-154>

¹Нұртас М.,

PhD, қауымдастырылған профессор, ORCID ID: 0000-0003-4351-0185,

e-mail: m.nurtas@iitu.edu.kz

^{1*}Абдикаликова З.Т.,

PhD, қауымдастырылған профессор, ORCID ID: 0000-0003-0131-4469,

*e-mail: z.abdikalikova@iitu.edu.kz

¹Адиат Л.,

магистр, лектор, ORCID ID: 0009-0006-2969-1695,

e-mail: l.adiat@iitu.edu.kz

¹Халықаралық ақпараттық технологиялар университеті, Алматы қ., Қазақстан

ӨЗІНЕ-ӨЗІ ҰҚСАС ШЕШІМДЕРГЕ АРНАЛҒАН ФИЗИКАҒА НЕГІЗДЕЛГЕН НЕЙРОНДЫҚ ЖЕЛІ (PINN) ӘДІСІ

Аңдатпа

Жылу өткізгіштік пен газ динамикасы сияқты күрделі физикалық процестерді сипаттайтын дербес туындысы бар дифференциалдық теңдеулерді сандық әдістер арқылы шешу барысында көбіне үлкен есептеу қуаты қажет етіледі. Осы қиындықтарды шешу үшін соңғы жылдары ғылым мен техниканың назарында физикаға негізделген нейрондық желілер (Physics-Informed Neural Networks, PINN) ерекше орын алған. Бұл мақалада PINN әдісі арқылы жылу және газ динамикасы теңдеулерінің шешімдерін табу мәселесі қарастырылады. Дәстүрлі сандық әдістерден ерекшелігі, физикаға негізделген нейрондық желі әдісі физикалық заңдарды нейрон желісінің құрылымына енгізу арқылы есепті шешуге мүмкіндік береді. Яғни, шешім тек қана мәліметтерге емес, сонымен қатар теңдеудің өзіне бағынады. Мақалада PINN әдісінің архитектурасы, шығын функцияларының құрылымы және олардың жылу өткізгіштік және Эйлер теңдеулерімен байланысы нақты мысалдармен сипатталады. Сонымен қатар, есептің бастапқы және шекаралық шарттарын енгізу механизмдері және шешімдердің орнықтылығы мен дәлдігіне әсер ететін факторлар талданған. Нәтижелер арқылы PINN-нің тиімділігі мен болашақта күрделі көп фазалы, көпөлшемді есептерге қолдану мүмкіндігі көрсетілді. Сонымен қатар, PINN-нің есептеу үдерісін жылдамдату және тұрақтылығын арттыру бағытындағы зерттеулер де ұсынылған.

Тірек сөздер: PINN, өзіне-өзі ұқсас, жылу теңдеуі, Эйлер теңдеуі, соққы түтігі (shock tube).

Кіріспе

Қазіргі таңда дифференциалдық теңдеулерді шешу әдістері ғылыми және инженерлік есептерде кеңінен қолданылады. Атап айтқанда, жылуөткізгіштік және газ динамикасы сияқты күрделі физикалық құбылыстарды сипаттайтын дербес туындылы дифференциалдық теңдеулер (ДТДТ) – табиғи және техникалық процестерді модельдеудің негізгі құралдарының бірі. Осындай есептерді сандық әдістермен шешу көптеген жағдайларда үлкен есептеу ресурстарын талап етеді, әсіресе көпөлшемді, шекаралық және бастапқы шарттары күрделі жүйелерде. Бұл шектеулерді жеңу мақсатында соңғы жылдары ғылым мен техникада физикаға негізделген нейрондық желілер (Physics-Informed Neural Networks, PINN) ерекше назарға ілінді.

PINN – машиналық оқыту мен классикалық физиканың синтезі болатын жаңа парадигма. Бұл әдіс алғаш рет (2019–2024 жж.) Raissi және т.б. [1–2] еңбектерде ұсынылып, кейіннен көптеген күрделі ДТДТ есептерін шешуде тиімділігін көрсетті. PINN әдісінің басты ерекшелігі – ол деректерге ғана емес, сонымен қатар физикалық заңдылықтарға (мысалы: энергия

сақталуы, масса сақталуы, импульс заңдары және т.б.) негізделеді. Мұндай тәсіл модельдің физикалық шындыққа сәйкестігін қамтамасыз етіп, мәліметтердің жеткіліксіздігі жағдайында да жақсы нәтиже алуға мүмкіндік береді.

PINN әдісімен салыстырғанда классикалық әдістер (мысалы: ақырлы элементтер әдісі (FEM), ақырлы айырмалар әдісі (FDM)) тор генерациясына, дәлдікке және есептеу тұрақтылығына тәуелді. Бұл әдістер әсіресе көпөлшемді, бейсызық немесе стохастикалық ДТДТ-де қиындықтарға тап болады. Ал PINN әдісі кеңістіктік торларды қажет етпейді, оның орнына кездейсоқ коллокациялық нүктелер арқылы шешімді үйретуге мүмкіндік береді. Сонымен қатар, PINN әдісі параллельді есептеуге ыңғайлы, бұл оны үлкен деректермен жұмыс істеу мен нақты уақыттағы есептерге бейімделуге мүмкіндік береді.

Бүгінгі таңда PINN әдісі жылуөткізгіштік, Био теңдеулері [3], Навье–Стокс [4], Коши есебі [5], реакция диффузиялық жүйелер, акустикалық толқындар [6], фазалық өріс модельдері сынды көптеген салаларда сәтті қолданылуда. Атап айтқанда, жылу мен газ динамикасы теңдеулері физикалық процестерді сипаттаудың маңызды мысалдары ретінде PINN әдісін қолдануға өте қолайлы. Бұл есептерде шекаралық және бастапқы шарттар, сондай-ақ күрделі шешім құрылымы жиі кездесетіндіктен, PINN әдісінің артықшылықтары айқын байқалады.

PINNs әдісінің даму барысымен, қазіргі жағдайы және болашақ зерттеу бағыттарымен жүйелі түрде [7] жұмыста танысуға болады. [8] әдебиетте PINN әдісін ғылыми есептеулерде қолданудың әртүрлі қырларын қамтитын кең көлемді шолу, оның ішінде PINN-есептегіштерін баптау, оқыту теориясының аспектілері, қолжетімді құралдар, болашақ бағыттар мен соңғы үрдістер, сондай-ақ дәлдік пен жинақталу мәселелері қарастырылады. PINN әдісінің эволюциясына арналған шолу, оның ішінде Physics-Informed Kolmogorov-Arnold Networks (PIKANs) сияқты жаңа архитектуралар және олардың биомедицина, механика, геофизика және басқа салалардағы қолданылуы [9]-да талқыланады. Жасанды нейрондық желілерді (artificial neural networks – ANN) қолдану арқылы дифференциалдық теңдеулерді шешу әдісі [10]-да сипатталады. [11] мақалада авторлар дәстүрлі машиналық оқыту әдістерін физикалық заңдармен біріктіру арқылы күрделі жүйелерді модельдеу мен болжау үшін қолданылатын тәсілді сипаттайды, бұл тәсіл әсіресе дифференциалдық теңдеулермен сипатталатын физикалық процестерге бағытталған.

Осы мақалада біз өзіне-өзі ұқсас шешімдерге негізделген PINN әдісін пайдалана отырып, жылу және газ динамикасының ішінара туынды теңдеулерін шешудің тиімді жолдарын қарастырамыз. Бұл тәсілдің артықшылығы – ол аналитикалық шешімге жақын, яғни өзіне-өзі ұқсастық қасиеті бар шешімдерді нейрондық желі арқылы қалпына келтіруге мүмкіндік береді. Осы бағыттағы зерттеулер PINN әдісінің теориялық және қолданбалы мүмкіндіктерін кеңейтіп, оны нақты физикалық процестерді модельдеуде кеңінен қолдануға жол ашады.

Материалдар мен әдістер

Өзіне-өзі ұқсастық ұғымы

Өзіне-өзі ұқсас шешім – уақыт немесе кеңістік масштабталғанда шешімнің формасы өзгермей, тек үлкейіп немесе кішірейіп отыратын сипат. Жалпы түрде шешім келесідей болады:

$$u(x, t) = t^{\alpha} * f\left(\frac{x}{t^{\beta}}\right),$$

мұндағы $f(\cdot)$ – өзіне-өзі ұқсас профиль, ал α , β – масштабтау көрсеткіштері [12].

Жылу теңдеуіне PINN қолдану

Зерттеуімізде тәжірибе ретінде жылу өткізгіштік және газ динамикасы теңдеулерін шешу үшін PINN әдісін қолдануды көрсетеміз. Бұл әдістің негізінде нейрондық желіні тек қана деректерге емес, сонымен қатар физикалық заңдарға – дифференциалдық теңдеулерге бағынатын түрде үйрету қағидасы қалай болатынын нақтылаймыз.

Жылу теңдеуі үшін модель

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2}, \quad u(x, 0) = \delta(x).$$

Аналитикалық шешімі:

$$u(x, t) = \frac{1}{\sqrt{4\pi t}} * \exp\left(-\frac{x^2}{4t}\right).$$

x – кеңістік координатасы;

t – уақыт;

$\xi = \frac{x}{\sqrt{t}}$ – өзіндік ұқсастық айнымалысы;

$u(x, t)$ – температура функциясы;

$f(\xi)$ – PINN арқылы үйрелінетін өзіндік ұқсас функция.

$$u_{\theta}(x, t) = \frac{1}{\sqrt{t}} * NN_{\theta}(\xi).$$

PINN шығын функциясы

$$u_{\theta}(x, t) = \frac{1}{\sqrt{t}} f_{\theta}(\xi), \quad \xi = \frac{x}{\sqrt{t}},$$

$$\mathcal{L}_{\text{heat}} = \frac{1}{N} \sum_{i=1}^N \left| \frac{\partial u_{\theta}}{\partial t}(x_i, t_i) - \frac{\partial^2 u_{\theta}}{\partial x^2}(x_i, t_i) \right|^2.$$

MAE (Mean absolute error):

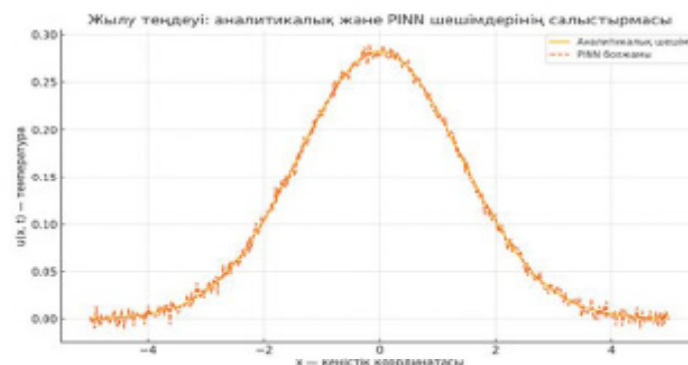
$$\text{MAE} = \frac{1}{N} \sum_{i=1}^N |U_{\text{pred}}(x_i) - U_{\text{true}}(x_i)|.$$

RMSE (Root mean square error):

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_{i=1}^N (U_{\text{pred}}(x_i) - U_{\text{true}}(x_i))^2}.$$

x	$U_{\text{pred}}(x) - U_{\text{true}}(x)$
-5	0
-2	+0.02
0	0
2	-0.02
5	0

Төмендегі 1-суретте жылу теңдеуі үшін аналитикалық және PINN шешімдерінің салыстырмасы ұсынылған.



Сурет 1 – Жылу теңдеуі: аналитикалық және PINN шешімдерінің салыстырмасы

Нәтижелер мен талқылау

1D Эйлер теңдеулері және өзіне өзі ұқсастық

Эйлер теңдеулерінің 1 өлшемді формасы:

$$\frac{\partial \rho}{\partial t} + \frac{\partial(\rho u)}{\partial x} = 0$$

$$\frac{\partial(\rho u)}{\partial t} + \frac{\partial(\rho u^2 + p)}{\partial x} = 0$$

$$\frac{\partial E}{\partial t} + \frac{\partial((E + p)u)}{\partial x} = 0.$$

Соққы түтігі бастапқы шарттары:

$$(\rho, u, p)(x, 0) = \begin{cases} (1.0, 0, 1.0), & \text{егер } x < 0 \\ (0.125, 0, 0.1), & \text{егер } x > 0 \end{cases}$$

PINN әдісіне негізделген шығын функциясы

1D Эйлер теңдеулері үшін PINN архитектурасы t және x айнымалыларын кіріс ретінде қабылдап, тығыздық $\rho(t, x)$, жылдамдық $u(t, x)$ және қысым $p(t, x)$ функцияларын болжайды. Бұл функциялардың көмегімен келесі қалдықтар анықталады:

$$\mathcal{R}_1(t, x) = \frac{\partial \rho_\theta}{\partial t} + \frac{\partial(\rho_\theta u_\theta)}{\partial x},$$

$$\mathcal{R}_2(t, x) = \frac{\partial(\rho_\theta u_\theta)}{\partial t} + \frac{\partial(\rho_\theta u_\theta^2 + p_\theta)}{\partial x};$$

$$\mathcal{R}_3(t, x) = \frac{\partial E_\theta}{\partial t} + \frac{\partial((E_\theta + p_\theta)u_\theta)}{\partial x},$$

мұндағы, $E_\theta = \frac{p_\theta}{\gamma - 1} + \frac{1}{2} \rho_\theta u_\theta^2$ – жалпы энергия.

Осы қалдықтарды ескере отырып, шығын функциясы келесі түрде жазылады:

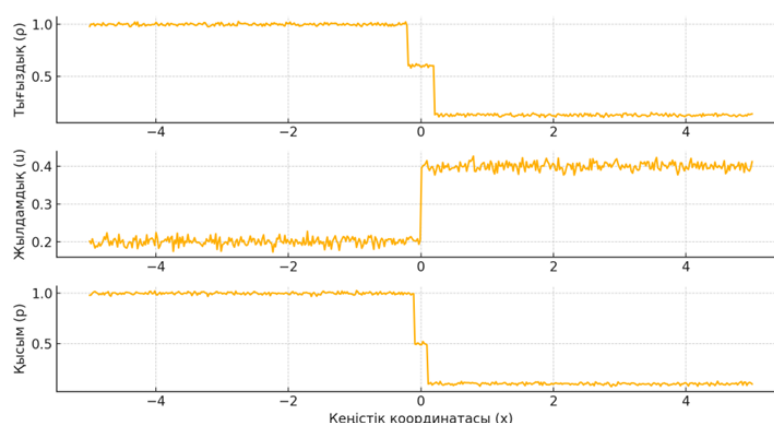
$$\mathcal{L}(\theta) = \mathcal{L}_d + \lambda_p \cdot \mathcal{L}_{phys},$$

мұнда $\mathcal{L}_d$ – мәліметтерге (data) негізделген (мысалы, MSE), ал $\mathcal{L}_{phys}$ – физикалық қалдықтардан тұратын бөлік:

$$\mathcal{L}_{phys} = \frac{1}{N_f} \sum_{i=1}^{N_f} (\mathcal{R}_1^2(t_i, x_i) + \mathcal{R}_2^2(t_i, x_i) + \mathcal{R}_3^2(t_i, x_i)),$$

мұндағы N_f – физикалық нүктелер саны, ал λ_p – физикалық заңдарды сақтау салмағы. Бұл функция PINN моделінің физикалық заңдарды сақтап шешім үйренуіне мүмкіндік береді.

Эйлер теңдеуі үшін соққы түтігі жағдайында PINN нәтижесі 2-суретте ұсынылған.

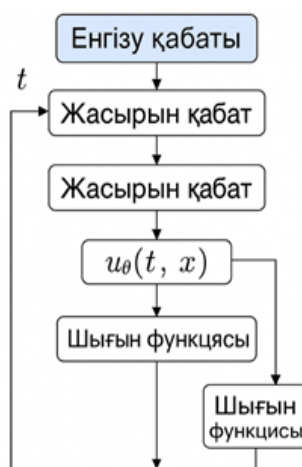


Сурет 2 – PINN нәтижесі: Эйлер тендеуі (соққы түтігі жағдайы)

PINN моделі және шығын функциялары

PINN архитектурасы

PINN архитектурасы – көпқабатты перцептрон (MLP) түрінде құрылады. Мысал ретінде, 3 жасырын қабатты нейрондық желі қарастырайық (3-сурет). Оның кірісінде x және t айнымалылары (кеңістік пен уақыт), ал шығысында $u_\theta(t, x)$ шамасы беріледі, мұндағы θ – модель параметрлері (салмақтар мен ығысу).



Сурет 3 – PINN моделі архитектурасы (Қосымша А негізінде)

Желі қабаттарының жалпы сипаттамасы:

- ♦ Кіріс қабаты: x , t мәндерін қабылдайды.
- ♦ Жасырын қабаттар: активация функциясы ретінде $\tanh$ немесе $\sin$ пайдаланылады.
- ♦ Шығыс қабаты: $u_\theta(t, x)$ – шешімнің жуықталған мәні.

Шығын функциялары

PINN-нің негізгі ерекшелігі – оның шығын функциясы физикалық қалдықтарды да қамтиды. Жалпы шығын функциясы келесі түрде жазылады:

$$\mathcal{L}(\theta) = \lambda_d \mathcal{L}_d + \lambda_p \cdot \mathcal{L}_p. \quad (1)$$

Мұндағы:

- ♦ $\mathcal{L}_d$ – мәліметтерге негізделген шығын (мысалы, нақты өлшемдермен салыстыру);
- ♦ $\mathcal{L}_n$ – физикаға негізделген қалдықтар: $\|\mathcal{N}[u_\theta] - f\|^2$;
- ♦ λ_d, λ_p – салмақ параметрлері.

Физикалық қалдықтар $\mathcal{L}_p$ – жылу немесе басқа теңдеулердің шешіміндегі теңсіздіктер арқылы анықталады. Мысалы, жылу теңдеуі үшін жоғарыда берілген формасын тағы да бір еске саламыз:

$$\mathcal{L}_p = \frac{1}{N_f} \sum_{i=1}^{N_f} \left| \frac{\partial u_\theta}{\partial t}(t_i, x_i) - \alpha \frac{\partial^2 u_\theta}{\partial x^2}(t_i, x_i) \right|^2. \quad (2)$$

Мұнда α – диффузия коэффициенті, N_f – физикалық нүктелер саны. Модельдің басты артықшылықтары:

- ♦ Шешімнің физика заңдарын бұзбай үйренуі;
- ♦ Деректер аз болған жағдайда да жұмыс істей алуы;
- ♦ Бастапқы және шекаралық шарттарды оңай енгізу мүмкіндігі;
- ♦ Жоғары өлшемді есептерге бейімдеуге болатын икемді құрылым.

Қорытынды

Осы мақалада біз PINN әдісі арқылы жылу теңдеуі және газ динамикасы теңдеулерін (Эйлер теңдеулері) шешудің мүмкіндіктерін қарастырдық. PINN – бұл физикалық жүйелерді модельдеудің жаңа парадигмасы болып табылатынын дәлелдедік. Ол сандық әдістер мен машиналық оқытудың үйлесімін қамтамасыз етіп, дәстүрлі дискретизация әдістеріне балама ұсынатынын көрсеттік. Өзіне өзі ұқсас шешімдер PINN моделінің тиімділігін арттырып, оқу уақытын қысқартатынын көрсетілді. Бұл тәсіл әсіресе аналитикалық шешімі жоқ күрделі дербес туындылы дифференциалдық теңдеулер үшін пайдалы екенін анықтадық. PINN әдісінің қазіргі таңда әлеуеті өте жоғары, алайда ол да жетілдіруді қажет ететінін мойындауымыз керек. Болашақ зерттеулерге келесі бағыттарды ұсынуға болады:

- ♦ Мультифизикалық есептер: PINN бірнеше теңдеулерден тұратын жүйелерге (мысалы, кеуекті серпімді, көп фазалы ағын) кеңейтілуі мүмкін [13];
- ♦ Есептің өлшемін үлкейту: Қазіргі жұмыс бір немесе екі өлшемге негізделген. Алдағы мақсат – үш өлшемді кеңістіктік және уақыттық есептерге бейімдеу;
- ♦ Адаптивті торлармен біріктіру: Тор тығыздығын қалдық мәніне қарай өзгертіп отыратын PINN түрін жасау – есептеу тиімділігін арттыруға сеп болады;
- ♦ Шығын функциясын автоматты теңшеу: λ_d, λ_p сияқты гиперпараметрлерді автоматты түрде реттейтін стратегияларды зерттеу керек;
- ♦ Сирек мәліметтер жағдайы: Реалистік жағдайда деректер өте аз және толық емес болады. Мұндай жағдайда, PINN-ді күшейту үшін генеративті модельдермен (мысалы, VAE, GAN) біріктіру даму мүмкіндігі бар бағыт.

Нәтижелер көрсеткендей, PINN арқылы алынған шешімдер жоғары дәлдікке ие. Сонымен қатар, оның шешімдері аналитикалық түрде алынатын шешімдерге жақын және сәйкес келетіндігі сандық тәжірибелердің графикалық салыстырулары арқылы дәлелденді. Қорытындылай келе, PINN – ғылым мен техникадағы көпқырлы күрделі есептерді шешуде болашақтың қуатты құралдарының бірі бола алады деп айта аламыз.

1. Қосымша А

PINN арқылы өзіндік ұқсас шешімді үйрену – толық Python коды (жылу теңдеуі мысалында)

```
import torch
import torch.nn as nn
import torch.autograd as autograd
import matplotlib.pyplot as plt
# Құрылғыны анықтау
```

```
device = torch.device("cuda" if torch.cuda.is_available() else "cpu")

# PINN моделі
class SelfSimilarPINN(nn.Module):
    def __init__(self):
        super().__init__()
        self.net = nn.Sequential(
            nn.Linear(1, 50),
            nn.Tanh(),
            nn.Linear(50, 50),
            nn.Tanh(),
            nn.Linear(50, 1) )

    def forward(self, x, t):
        xi = x / torch.sqrt(t + 1e-5)
        f = self.net(xi)
        u = f / torch.sqrt(t + 1e-5)
        return u

# Оқу деректерін генерациялау
def generate_training_data(n=1000):
    x = torch.linspace(-5, 5, n).reshape(-1, 1)
    t = torch.linspace(0.1, 1.0, n).reshape(-1, 1)
    x, t = torch.meshgrid(x.squeeze(), t.squeeze(), indexing='ij')
    return x.reshape(-1, 1).to(device), t.reshape(-1, 1).to(device)

# PDE қалдығын есептеу
def pde_residual(model, x, t):
    x.requires_grad = True
    t.requires_grad = True
    u = model(x, t)
    u_t = autograd.grad(u, t, grad_outputs=torch.ones_like(u), retain_graph=True, create_graph=True)[0]
    u_x = autograd.grad(u, x, grad_outputs=torch.ones_like(u), retain_graph=True, create_graph=True)[0]
    u_xx = autograd.grad(u_x, x, grad_outputs=torch.ones_like(u_x), retain_graph=True, create_graph=True)[0]
    return u_t - u_xx

# Модель және optimizer
model = SelfSimilarPINN().to(device)
optimizer = torch.optim.Adam(model.parameters(), lr=1e-3)

# Оқу деректері
x_train, t_train = generate_training_data(200)

# Оқыту циклі
loss_history = []
for epoch in range(1000):
    optimizer.zero_grad()
    residual = pde_residual(model, x_train, t_train)
    loss = torch.mean(residual ** 2)
```

```

loss.backward()
optimizer.step()
loss_history.append(loss.item())
if epoch % 100 == 0:
    print(f'Epoch {epoch}, Loss: {loss.item():.6f}')

```

```

# Шығын функция графигі
plt.plot(loss_history)
plt.xlabel("Epoch")
plt.ylabel("Loss")
plt.title("Training Loss")
plt.grid(True)
plt.show()

```

2. Қосымша В

PINN арқылы Эйлер теңдеуін шешу(Self-Similar Shock Tube) — Python коды

```

import torch
import torch.nn as nn
import torch.autograd as autograd
import matplotlib.pyplot as plt
device = torch.device("cuda" if torch.cuda.is_available() else "cpu")
gamma = 1.4
# Желі блогы
class Net(nn.Module):
    def __init__(self):
        super(Net, self).__init__()
        self.net = nn.Sequential(
            nn.Linear(1, 64),
            nn.Tanh(),
            nn.Linear(64, 64),
            nn.Tanh(),
            nn.Linear(64, 1) )
    def forward(self, xi):
        return self.net(xi)

# PINN моделі: rho, u, p
class EulerPINN(nn.Module):
    def __init__(self):
        super(EulerPINN, self).__init__()
        self.rho_net = Net()
        self.u_net = Net()
        self.p_net = Net()
    def forward(self, x, t):
        xi = x / (t + 1e-5)
        rho = self.rho_net(xi)
        u = self.u_net(xi)
        p = self.p_net(xi)
        return rho, u, p
# PDE қалдықтары
def euler_residuals(model, x, t):
    x.requires_grad = True

```

```
t.requires_grad = True
rho, u, p = model(x, t)
m = rho * u
E = p / (gamma - 1) + 0.5 * rho * u**2
rho_t = autograd.grad(rho, t, grad_outputs=torch.ones_like(rho), retain_graph=True, create_graph=True)[0]
m_t = autograd.grad(m, t, grad_outputs=torch.ones_like(m), retain_graph=True, create_graph=True)[0]
E_t = autograd.grad(E, t, grad_outputs=torch.ones_like(E), retain_graph=True, create_graph=True)[0]
rho_x = autograd.grad(rho, x, grad_outputs=torch.ones_like(rho), retain_graph=True, create_graph=True)[0]
m_x = autograd.grad(m, x, grad_outputs=torch.ones_like(m), retain_graph=True, create_graph=True)[0]
p_x = autograd.grad(p, x, grad_outputs=torch.ones_like(p), retain_graph=True, create_graph=True)[0]
u_x = autograd.grad(u, x, grad_outputs=torch.ones_like(u), retain_graph=True, create_graph=True)[0]
E_x = autograd.grad(E, x, grad_outputs=torch.ones_like(E), retain_graph=True, create_graph=True)[0]

# Қалдықтар
mass_eq = rho_t + m_x
momentum_eq = m_t + autograd.grad(m * u + p, x, grad_outputs=torch.ones_like(p), retain_graph=True, create_graph=True)[0]
energy_eq = E_t + autograd.grad((E + p) * u, x, grad_outputs=torch.ones_like(E), retain_graph=True, create_graph=True)[0]
return mass_eq, momentum_eq, energy_eq

# Деректер генерациясы
def generate_data(n=1000):
    x = torch.linspace(-1, 1, n).reshape(-1, 1).to(device)
    t = torch.ones_like(x) * 1.0
    return x, t

# Модель және optimizer
model = EulerPINN().to(device)
optimizer = torch.optim.Adam(model.parameters(), lr=1e-3)
x_train, t_train = generate_data(1000)

# Оқыту циклі
loss_history = []
for epoch in range(1000):
    optimizer.zero_grad()
    mass_eq, mom_eq, energy_eq = euler_residuals(model, x_train, t_train)
    loss = torch.mean(mass_eq**2) + torch.mean(mom_eq**2) + torch.mean(energy_eq**2)
    loss.backward()
    optimizer.step()
    loss_history.append(loss.item())
    if epoch % 100 == 0:
        print(f'Epoch {epoch}: Loss = {loss.item():.6f}')
```

```
# Шығын функция графигі
plt.plot(loss_history)
plt.xlabel("Epoch")
plt.ylabel("Loss")
plt.title("PINN Training Loss (1D Euler, self-similar)")
plt.grid(True)
plt.show()
```

ӘДЕБИЕТТЕР

- 1 Raissi M., Perdikaris P., Karniadakis G.E. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations // *Journal of Computational Physics*. – 2019. – Vol. 378. – P. 686–707. <https://doi.org/10.1016/j.jcp.2018.10.045>.
- 2 Raissi M., Perdikaris P., Daryakenari N.A., Karniadakis G.E. Physics-Informed Neural Networks and Extensions. – 2024. <https://doi.org/10.48550/arXiv.2408.16806>.
- 3 Meirmanov A. Mathematical models for poroelastic flows. – Paris: Atlantis Press, Atlantis Studies in Differential Equations, 2014. <https://doi.org/10.2991/978-94-6239-015-7>.
- 4 Fefferman C.L. Existence and smoothness of the Navier–Stokes equation. – Clay Mathematics Institute, 2017. – P. 1–6. URL: <https://www.claymath.org/wp-content/uploads/2022/06/navierstokes.pdf>.
- 5 Конюхов И.В., Конюхов В.М., Черница А.А., Дюсенова А. Особенности применения физически информированных нейронных сетей для решения обыкновенных дифференциальных уравнений // *Компьютерные исследования и моделирование*. – 2024. – Т. 16. – № 7. – С. 1621–1636. <https://doi.org/10.20537/2076-7633-2024-16-7-1621-1636>.
- 6 Meirmanov A., Nurtas M. Mathematical models of seismic in composite media: elastic and poro-elastic components // *Electronic Journal of Differential Equations*. – 2016. – No. 184. – P. 1–22. URL: <http://ejde.math.unt.edu>.
- 7 Lawal Z.K., Yassin H., Lai D.T.C., Idris A.C. Physics-Informed Neural Network (PINN) Evolution and Beyond: A Systematic Literature Review and Bibliometric Analysis // *Big Data and Cognitive Computing*. – 2022. – Vol. 6. – No. 4. – Article 140. <https://doi.org/10.3390/bdcc6040140>.
- 8 Cuomo S., Schiano di Cola V., Giampaolo F., Rozza G., Raissi M., Piccialli F. Scientific Machine Learning through Physics-Informed Neural Networks: Where we are and What's next. – 2022. <https://doi.org/10.48550/arXiv.2201.05624>.
- 9 Toscano J.D., Oommen V., Varghese A.J., Zou Z., Daryakenari N.A., Wu C., Karniadakis G.E. From PINNs to PIKANs: Recent Advances in Physics-Informed Machine Learning. – 2024. <https://doi.org/10.48550/arXiv.2410.13228>.
- 10 Lagaris I.E., Likas A., Fotiadis D.I. Artificial neural networks for solving ordinary and partial differential equations // *IEEE Transactions on Neural Networks*. – 1998. – Vol. 9. – No. 5. – P. 987–1000.
- 11 Karniadakis G.E., Kevrekidis I.G., Lu L., Perdikaris P., Wang S., Yang L. Physics-informed machine learning // *Nature Reviews Physics*. – 2021. – Vol. 3. – P. 422–440.
- 12 Barenblatt G.I. Scaling, self-similarity, and intermediate asymptotics. – Cambridge: Cambridge University Press, 1996.
- 13 Krasnozhonov R. and Nurtas M. Modeling the Propagation of Acoustic Waves in an Elastic Medium Using Physics-Informed Neural Networks, 2025 IEEE 5th International Conference on Smart Information Systems and Technologies (SIST), Astana, Kazakhstan, 2025, pp. 1–7. <https://doi.org/10.1109/SIST61657.2025.11139217>.

REFERENCES

- 1 Raissi, M., Perdikaris, P., Karniadakis, G.E. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *Journal of Computational Physics*, 378, 686–707 (2019). <https://doi.org/10.1016/j.jcp.2018.10.045>.
- 2 Raissi, M., Perdikaris, P., Daryakenari, N.A., Karniadakis, G.E. Physics-Informed Neural Networks and Extensions, 2024. <https://doi.org/10.48550/arXiv.2408.16806>.

- 3 Meirmanov, A. Mathematical models for poroelastic flows. Paris: Atlantis Press, Atlantis Studies in Differential Equations, 2014. <https://doi.org/10.2991/978-94-6239-015-7>.
- 4 Fefferman, C.L. Existence and smoothness of the Navier–Stokes equation. Clay Mathematics Institute, pp. 1–6 (2017). URL: <https://www.claymath.org/wp-content/uploads/2022/06/navierstokes.pdf>.
- 5 Konyukhov, I.V., Konyukhov, V.M., Chernitsa, A.A., Dyusenova, A. Osobennosti primeneniya fizicheskii informirovannykh neyronnykh setey dlya resheniya obyknovennykh differentsialnykh uravneniy. Kompyuternye issledovaniya i modelirovanie, 16(7), 1621–1636 (2024). <https://doi.org/10.20537/2076-7633-2024-16-7-1621-1636>.
- 6 Meirmanov, A., Nurtas, M. Mathematical models of seismic in composite media: elastic and poroelastic components. Electronic Journal of Differential Equations, 184, 1–22 (2016). URL: <http://ejde.math.unt.edu>.
- 7 Lawal, Z.K., Yassin, H., Lai, D.T.C., Idris, A.C. Physics-Informed Neural Network (PINN) Evolution and Beyond: A Systematic Literature Review and Bibliometric Analysis. Big Data and Cognitive Computing, 6(4), article 140 (2022). <https://doi.org/10.3390/bdcc6040140>.
- 8 Cuomo, S., Schiano di Cola, V., Giampaolo, F., Rozza, G., Raissi, M., Piccialli, F. Scientific Machine Learning through Physics-Informed Neural Networks: Where we are and What's next, 2022. <https://doi.org/10.48550/arXiv.2201.05624>.
- 9 Toscano, J.D., Oommen, V., Varghese, A.J., Zou, Z., Daryakenari, N.A., Wu, C., Karniadakis, G.E. From PINNs to PIKANs: Recent Advances in Physics-Informed Machine Learning, 2024. <https://doi.org/10.48550/arXiv.2410.13228>.
- 10 Lagaris, I.E., Likas, A., Fotiadis, D.I. Artificial neural networks for solving ordinary and partial differential equations. IEEE Transactions on Neural Networks, 9(5), 987–1000 (1998).
- 11 Karniadakis, G.E., Kevrekidis, I.G., Lu, L., Perdikaris, P., Wang, S., Yang, L. Physics-informed machine learning. Nature Reviews Physics, 3, 422–440 (2021).
- 12 Barenblatt, G.I. Scaling, self-similarity, and intermediate asymptotics. Cambridge: Cambridge University Press, 1996.
- 13 Krasnozhanov, R. and Nurtas, M. Modeling the Propagation of Acoustic Waves in an Elastic Medium Using Physics-Informed Neural Networks, 2025 IEEE 5th International Conference on Smart Information Systems and Technologies (SIST), Astana, Kazakhstan, 2025, pp. 1–7, <https://doi.org/10.1109/SIST61657.2025.11139217>.

¹Нуртас М.,

PhD, ассоциированный профессор, ORCID ID: 0000-0003-4351-0185,

e-mail: m.nurtas@iitu.edu.kz

^{1*}Абдикаликова З.Т.,

PhD, ассоциированный профессор, ORCID ID: 0000-0003-0131-4469,

*e-mail: z.abdikalikova@iitu.edu.kz

¹Адиат Л.,

магистр, лектор, ORCID ID: 0009-0006-2969-1695,

e-mail: l.adiat@iitu.edu.kz

¹Международный университет информационных технологий, г. Алматы, Казахстан

МЕТОД ФИЗИЧЕСКИ ОБОСНОВАННОЙ НЕЙРОННОЙ СЕТИ (PINN), ОСНОВАННЫЙ НА САМОПОДОБНЫХ РЕШЕНИЯХ

Аннотация

При численном решении дифференциальных уравнений в частных производных, описывающих сложные физические процессы, такие как теплопроводность и газовая динамика, зачастую требуется значительная вычислительная мощность. Для решения этих сложностей в последние годы особое внимание науки и техники привлекают физически информированные нейронные сети (Physics-Informed Neural Networks, PINN). В данной статье рассматривается задача нахождения решений уравнений теплопроводности и газовой динамики с помощью метода PINN. В отличие от традиционных численных методов, метод физиче-

ски информированных нейронных сетей позволяет решать задачи, внедряя физические законы в структуру нейронной сети. То есть решение подчиняется не только данным, но и самому уравнению. В статье описываются архитектура метода PINN, структура функций потерь и их связь с уравнением теплопроводности и уравнениями Эйлера на конкретных примерах. Кроме того, анализируются механизмы введения начальных и граничных условий, а также факторы, влияющие на устойчивость и точность решений. Полученные результаты демонстрируют эффективность PINN и возможность их применения в будущем для решения сложных многомерных и многофазных задач. Также предложены исследования, направленные на ускорение вычислительного процесса и повышение стабильности PINN.

Ключевые слова: PINN, самоподобие, уравнение теплопроводности, уравнение Эйлера, ударная труба (shock tube).

¹**Nurtas M.,**

PhD, Associate Professor, ORCID ID: 0000-0003-4351-0185,

e-mail: m.nurtas@iitu.edu.kz

^{1*}**Abdikalikova Z.T.,**

PhD, Associate Professor, ORCID ID: 0000-0003-0131-4469,

*e-mail: z.abdikalikova@iitu.edu.kz

¹**Adiat L.,**

MSc, Lecturer, ORCID ID: 0009-0006-2969-1695,

e-mail: l.adiat@iitu.edu.kz

¹International University of Information Technologies, Almaty, Kazakhstan

PHYSICS-INFORMED NEURAL NETWORK (PINN) METHOD BASED ON SELF-SIMILAR SOLUTIONS

Abstract

In the numerical solution of partial differential equations that describe complex physical processes such as heat conduction and gas dynamics, substantial computational resources are often required. To address these challenges, Physics-Informed Neural Networks (PINNs) have gained increasing attention in recent years within the fields of science and engineering. This paper investigates the application of the PINN methodology to obtain solutions to the heat conduction and gas dynamics equations. Unlike traditional numerical approaches, the physics-informed neural network framework incorporates governing physical laws directly into the neural network architecture. Consequently, the solution is constrained not only by data but also by the underlying differential equations. The paper presents the architecture of the PINN framework and details the structure of loss functions, demonstrating their relationship with the heat equation and the Euler equations using specific examples. Furthermore, the implementation of initial and boundary conditions is discussed, along with an analysis of factors influencing the stability and accuracy of the obtained solutions. The results highlight the efficiency of PINNs and demonstrate their potential for solving complex multiphase and high-dimensional problems in the future. Additionally, current research directions aimed at accelerating the computational process and enhancing the robustness of PINNs are outlined.

Keywords: PINN, self-similarity, heat equation, Euler equation, shock tube.

Мақаланың редакцияға түскен күні: 17.10.2025

ӨОЖ 681.327.12
ФТАХР 65.011.56

<https://doi.org/10.55452/1998-6688-2025-22-4-155-167>

^{1,3,4}**Амангелді Н.**,
PhD, ORCID ID: 0000-0002-4669-9254,
e-mail: amangeldi_n_3@enu.kz
^{1,2*}**Еримбетова А.**,
т.ғ.к., қауымдастырылған профессор, PhD,
ORCID ID: 0000-0002-2013-1513,
*e-mail: aigerian8888@gmail.com
^{1,4}**Гаизова Н.Е.**,
магистр, ORCID: 0009-0007-7278-4202,
e-mail: g.nazerke2755@gmail.com
^{1,3}**Турсынова Н.А.**,
магистр, ORCID ID: 0000-0002-1857-9028,
e-mail: tursynova_na@enu.kz
⁴**Болатбекқызы К.**,
бакалавр, ORCID ID: 0009-0000-7659-4974,
e-mail: kerbezbolatbekkyzy03@gmail.com

¹Ақпараттық және есептеуіш технологиялар институты, Алматы қ., Қазақстан

²Еуразия технологиялық университеті, Алматы қ., Қазақстан

³Л.Н. Гумилев атындағы Еуразия ұлттық университеті, Астана қ., Қазақстан

⁴«SignBridge» ЖШС, Астана қ., Қазақстан

ГЛОСАРЛЫҚ ҚАБАТТЫ ҚОЛДАНУ АРҚЫЛЫ ҚАЗАҚ ТІЛІНДЕГІ МӘТІНДІ ФОРМАЛИЗАЦИЯЛАУ

Аңдатпа

Ымдау тілін автоматты өңдеу технологиялары цифрлық трансформация кезеңінде теңсіздікке тап болған есту және сөйлеу қабілетінде шектеуі бар қоғам мүшелерінің өзекті қажеттілігіне айналды. Соңғы жылдары ымдау тілін табиғи тілмен бір деңгейдегі формалды құрылым ретінде қарастырып, оны автоматты жүйелерге бейімдеу мәселесі зерттеушілер назарын ерекше аударып келеді. Табиғи тілдегі ақпаратты ымдау тіліне автоматты аудару міндетін орындау үшін аралық қабат ретінде ымдау тілінің мәтіндік нысаны глостар пайдаланылады. Осы мақсатта аталған зерттеуде табиғи тілді өңдеу әдістері арқылы қазақ тілінің морфологиялық ерекшеліктерін қамтитын қазақ тіліндегі мәтінді ымдау тілінің глостарына түрлендірудің жаңа тәсілі ұсынылады. Атап айтқанда, ByT5-кіші үлгісі негізіндегі Seq2Seq архитектурасы қолданылады. Алынған нәтижелер ымдау тілінің ішкі құрылымын сақтайтын ықшам әрі мағыналық тұрғыдан бай глостар тізбектерін қалыптастыратынын көрсетеді. Глостар тізбегі ымдау тіліндегі қимылдарды жазбаша тілге ұқсас логикалық бірліктер ретінде бейнелейтін интерпретацияланатын аралық қабаттың жұмысын автоматтандыруға мүмкіндік береді. Түрлендірілген глостар тізбегі ымдау тілінің құрылымын сақтай отырып, артық ақпаратты азайтады және сөйлемнің үйлесімділігін арттырады. Демек, ымдау тілінің аватарларын басқаруда тек семантикалық тұрғыдан мәнді бірліктердің қолданылуы есептеу жүктемесін төмендетеді. Қысқа әрі мағыналық тұрғыдан бай глостар – ымдау тілінде қол қимылдарын синтездеудің тиімді ресурсы.

Тірек сөздер: ымдау тілі, глостар тізбегі, табиғи тілді өңдеу, формалдау, Seq2Seq, ByT5.

Кіріспе

Дүниежүзілік денсаулық сақтау ұйымының мәліметтері бойынша, әлемде есту қабілетінде шектеу бар 430 миллионға жуық адам бар, олардың 70 миллионнан астамы ымдау тілін

қолданатын қауымдастықтарға жатады [1]. Ым тілдері есту немесе сөйлеу мүмкіндігі шектеулі адамдар үшін ана тілі болып саналады және олардың мәдениеті мен жеке тұлғасының негізін құрайтын толыққанды, дербес тілдік жүйелерді құрайды [2]. Халықаралық ұйымдар ым тіліне қол жетімділік пен есту қабілеті шектеулі адамдардың құқықтарының сақталуы арасында тығыз байланыс бар екенін ерекше атап көрсетеді [2]. Ана тілін толық меңгермеген жағдайда, есту қабілеті шектеулі адамдар қоғамда тең мүмкіндіктерге ие бола алмайды. Ым тілін еркін меңгеру және қолдану оларға толыққанды әлеуметтік қатысу мен өмірдің барлық салаларында тең құқықтарды қамтамасыз етеді.

Цифрлық дәуір есту қабілеті шектеулі қауымдастықтар үшін жаңа теңсіздік формаларын тудырып отыр. Көптеген онлайн-контент және цифрлық сервистер еститін пайдаланушыларға бағытталған және негізінен дыбыстық немесе жазбаша түрде ұсынылған, бұл оларды есту қабілеті шектеулі адамдар үшін қол жетімсіз етеді [3]. Тіпті субтитрлер мен мәтіндік транскрипциялар да мәселені толық шеше алмайды: көптеген есту қабілеті шектеулі адамдар үшін жазбаша тіл – ана тілімен тікелей байланысты емес екінші тіл.[3].

Ақпаратқа қол жеткізудің шектеулі болуының негізгі себептерінің бірі – ым тілдерінің лексикалық ресурстарының жеткіліксіздігі. Мысалы, Қытай ым тілінде 902 белгіден тұратын базалық тізім қолданулардың шамамен 77%-ын қамтиды, бұл жиілік сөздігінің ықшам «ядросын» көрсетеді [4]. Ағылшын ым тілінің лексикалық деректер жинағы ASL-LEX шамамен 1000 ASL белгісін қамтиды, ал ауызша тілдерде ондаған мың лексемалар бар [5]. Ым тілдерінде иконикалық бірліктердің үлесі ағылшын және испан тілдеріне қарағанда жоғары [6]. Шектеулі сөздік қоры мәтінді түсінуге тікелей әсер етеді: мектеп оқушылары арасында ол жазбаша материалды түсіну деңгейімен корреляцияланады [7], ал есту қабілеті шектеулі студенттердің сөздік қоры мен «дүниені тануы» естуші құрдастарына қарағанда статистикалық тұрғыдан төмен [8]. Бала кезіндегі тілдік депривация ұзақ мерзімді когнитивтік және академиялық салдарға әкелуі мүмкін, бұл ым тіліне ерте қол жеткізудің маңызды екенін көрсетеді [9].

Цифрлық теңсіздік, әсіресе COVID-19 пандемиясы кезінде айқын көрінді: көптеген есту қабілеті шектеулі адамдар қашықтан медициналық кеңестерге және денсаулыққа қатысты сенімді ақпаратқа қол жеткізу кезінде қиындықтарға тап болды, себебі материалдар сирек ым тілінде ұсынылды [10].

Бұл теңсіздіктің себептері лексикографиялық және әлеуметтік-мәдени факторлармен [6–10] қатар техникалық аспектілерге де байланысты. Қазіргі технологиялар жиі есту қабілеті шектеулі пайдаланушылардың қажеттіліктерін ескермей жасалады. Көптеген цифрлық сервистер естуші пайдаланушыларға арналып әзірленеді және ым тілін қолданушы қауымдастықтардың қатысуынсыз жобаланады, бұл ым тілінде сөйлейтін пайдаланушылар үшін олардың практикалық құндылығын төмендетеді [11].

Бұл мәселе әсіресе Text-to-Gloss (TtG) тапсырмаларында айқын көрінеді, себебі бұл тапсырма мәтінді автоматты түрде глостар тізбегіне аудару үшін үлкен және сәйкестендірілген корпустар (бейне + глостар) қажет болады. Қолданыстағы бастамаларға қарамастан, мұндай ресурстар әлі де шектеулі болып қалуда және негізінен ASL [12–15], азиялық [16–17] және кейбір еуропалық [18–21] ым тілдерінде негізделген, ал көптеген аз ресурсты ым тілдері үшін деректер жиынтықтары жоқ немесе көлемі мен қол жетімділігі бойынша шектеулі.

Ғылыми зерттеулерге шолу

Ымдау тілдерін автоматты түрде өңдеу саласындағы заманауи зерттеулер негізгі екі үлкен бағытта жүргізілуде: глосқа негізделген (gloss-based / TtG) және глоссыз (end-to-end/ gloss-free). Глосқа негізделген әдісте бастапқы мәтін алдымен глос тізбегіне аударылады, содан кейін глостар ымға/мультимедиаға трансляцияланады. Ал глоссыз әдіс аралық аннотациясыз, мәтін мен видео немесе негізгі нүктелерден тұратын кеңістік-уақыттық белгілерді тікелей сәйкестендіруге ұмтылады.

Глосқа негізделген әдістер

TtG-модельдері қолданыстағы NMT (Neural Machine Translation) тәсілдеріне сүйену мүмкіндігінің арқасында мәтінді ым тіліне аудару міндетін жүйелеудің алғашқы сәтті талпыныстарының бірі болды. Глосқа негізделген әдістер тапсырманы екі кезеңге бөлуге мүмкіндік береді: text→gloss және gloss→sign. Мұндай модульдік сипат машиналық аудармадағы тексерілген технологияларды қолдануды жеңілдетеді. Мәселен, Ye et al. (2023) [22] мақсатты домендерде мәтін генерациясы арқылы кері аударманы масштабтауға негізделген бірегей тәсілді ұсынды. Авторлар GPT-2 тілдік моделіне сүйенетін Prompt-based Domain Text Generation әдісін әзірлеп, жасанды домендік мәтіндерді қалыптастырды. Бұл мәтіндер кейіннен «глосс – мәтін» синтетикалық жұптарының көзі ретінде back-translation процедурасында пайдаланылды. Мұндай амал глосқа негізделген аудармада деректердің елеулі тапшылығын еңсеруге мүмкіндік беріп, әртүрлі деректер жиынтықтарында сапаны едәуір жақсартқанын көрсетті. Дегенмен, бұл әдіс модельдің генерациясының дұрыстығына тәуелді болып қала береді және мағынасы жағынан сәйкес келмейтін мәтіннің пайда болу ықтималдылығын жоққа шығармайды, бұл алынған шешімдердің әртүрлі жағдайларға бейімделуін шектеуі мүмкін.

Zhu et al. (2023) [23] тағы бір маңызды үлес қосып, мәтінді глосқа нейрондық машиналық аударудың әдістерін қарастырып, көптілді және аз дерекпен оқытылған жағдайларға баса назар аударды. Зерттеушілер төрт негізгі тәсілді талдады:

1. Деректерді аугментациялау – back және forward-translation қолдану, алдын ала өңдеулерді біріктіру және синтетикалық мысалдарға тегтер енгізу.

2. Ішінара бақыланатын NMT – үлкен моно-корпустардан сөйлемдерді «көшіру» арқылы модельді оқыту.

3. Трансферлік оқыту – үлкен корпустарда алдын ала оқыту жүргізіп, кейін мақсатты параллель деректерде fine-tuning жасау.

Көптілді NMT – бір модельді бір уақытта неміс тілі→глосс және неміс тілі→ағылшын тілі бағыттарында оқыту, мақсатты тілге арналған токендерді пайдалану.

Модельдер қарапайым Transformer архитектурасына негізделген (1 қабатты энкодер, 2 қабатты декодер, сөздік көлемі шамамен 32 000 болатын BPE-токенизация) және неміс тіліндегі DGS корпустарында (PHOENIX-Weather 2014T және Hamburg DGS Corpus) оқытылып, американдық ым тілі бойынша (ASL, NCSLGR, шамамен 1500 сөйлем) тексерілді. Сапаны бағалау BLEU, ChrF метрикалары және ым тілін меңгерген адамдардың бағалауы (Direct Assessment, шкала 0–6) арқылы жүргізілді. Ең жоғары нәтижелер көптілді модельде аугментация қолданылған кезде тіркелді: PHOENIX-та BLEU = 28,5 (+6 базалық модельмен салыстырғанда), DGS-та BLEU = 13,4 (+2). Сондай-ақ, ым тілін пайдаланушылардың орташа бағасы базалық модельге қарағанда айтарлықтай жоғары болды. Зерттеу нәтижелері көрсеткендей, аталған әдістердің кешенді қолданылуы мәтінді глосқа аударудың сапасын елеулі түрде арттырады.

Xu et al. (2022) [24] зерттеуінде басымдық практикалық құралдық аспектілерге берілген: олар глосс сөздігіне арналған автоматтандырылған GlossFinder жүйесін ұсынды. Жүйе 2000-ға жуық белгілерді қамтитын корпусқа сүйеніп, пайдаланушы енгізген глостарды сәйкестендіріп, тиісті түсіндірмесін береді. GlossFinder аударма мәселесін шешпегенімен, мәтін мен глостар арасында көпір орнататын қосымша құрал ретінде маңызды және болашақта машиналық аударма жүйелерінде көмекші компонент ретінде қолданылуы мүмкін. Жүйенің шектеулері ретінде сөздіктің салыстырмалы шағын көлемі мен контекстік өңдеудің болмауы атап өтіледі, бұл жүйені тек ішінара шешім ретінде сипаттайды.

Mohamed et al. (2022) [25] трансформерлерге негізделген тәсілді іске асырып, ASLG-PC12 корпусында мәтін мен глостар арасындағы екі бағытты аударма үшін seq2seq моделін оқытты. Қос бағытта да кіріс деректері алдын ала өңделіп, Transformer энкодеріне беріледі, ал декодер шығыс тізбекті генерациялайды. Энкодер мен декодердің әрбір қабаты self-attention

механизмін және позицияға тәуелсіз толық байланысқан желіні (feed-forward) қамтып, қалдық байланыстар мен нормализацияны қолданады. Декодерде encoder–decoder attention қосылып, self-attention қабаттары болашақ позицияларды бүркемелеу арқылы автоагрессивті шығысты қамтамасыз етеді. Осы тәсіл модельге тізбектегі әртүрлі қашықтықтағы байланыстарды ескеріп, ым тілінің грамматикасы мен контекстін дұрыс өңдеуге мүмкіндік берді. Модельдер BLEU және ROUGE метрикалары бойынша бағаланды. Мәтіннен глосқа аударуда ең жоғары көрсеткішті 2-қабатты модельге тиесілі: BLEU-4 $\approx 96,89\%$, ROUGE $\approx 98,78\%$. Глостан мәтінге аударуда да 2-қабатты модель ең жақсы нәтиже көрсетті: BLEU-4 $\approx 84,82\%$, ROUGE $\approx 96,90\%$. Бұл нәтиже глосқа негізделген аудармада терең модельдердің әлеуетін айғақтайды.

Глосарлық көрініс көптеген жылдар бойы мәтінді ым тіліне аудару саласында негізгі тәсіл ретінде қарастырылып келгенімен, оның қолданылуы бірқатар шектеулермен қатар жүреді. Дәл осы мәселені Müller et al. (2023) “Considerations for meaningful sign language machine translation based on glosses” атты еңбегінде жан-жақты талдады [26]. Авторлар глосс деңгейіне негізделген қолданыстағы зерттеулерді жүйелеп, оның артықшылықтары мен әлсіз жақтарын айқын көрсетті.

Зерттеуде глосарды аударма үдерісіндегі аралық буын ретінде пайдаланған ғылыми еңбектер корпусы жүйелі түрде сарапталды. Негізгі назар глосардың машиналық оқыту міндетін жеңілдететініне аударылды, себебі олар ым тіліндегі пайымдауларды жазба тіл құрылымына ұқсас токендер тізбегі түрінде бейнелейді. Бұл тәсіл дәстүрлі нейрондық машиналық аударма (NMT) архитектураларын қолдануға және жақын салалардағы әдістерді оңай бейімдеуге мүмкіндік береді. Алайда Müller және әріптестері осы әдістің бірқатар маңызды шектеулерін ерекше атап өтеді:

- ♦ Лингвистикалық толық еместік: глосар тек қимылдардың сызықтық ретін белгілейді және ым тілінің кеңістіктік компоненттерін, вербалды емес құралдарын және көпөлшемділігін жеткізе алмайды.

- ♦ Деректердің шектеулілігі: негізгі қолданылатын корпус RWTH-PHOENIX-Weather-2014T көлемі жағынан шағын әрі тақырыптық тұрғыда шектеулі, бұл модельдердің жалпылау қабілетін төмендетеді.

- ♦ Зерттеулердің жеткіліксіз ашықтығы: жұмыстардың аз бөлігі ғана глосарлық тәсілдің шектеулерін нақты мойындайды.

- ♦ Метрикалардың әртүрлілігі: BLEU, METEOR, COMET және басқа көрсеткіштердің әртүрлі нұсқаларын пайдалану нәтижелерді салыстыруды қиындатады, бұл саладағы прогресті объективті бағалауға кедергі келтіреді.

Müller және әріптестерінің (2023) еңбегі глосарлық әдістердің жеңілдетілген аралық деңгей ретінде белгілі бір практикалық құндылығы бар екенін көрсеткенімен, олардың қазіргі күйінде ым тілінің барлық аспектілерін толық бейнелей алмайтынын көрсетеді. Мұндай тәсілдердің одан әрі дамуы деректердің сапасын арттырумен қатар, бағалау сапасын жетілдіруді де талап етеді.

Глоссыз әдістер

Глосс аннотацияларын қолданбайтын ым тілі аудармасының әдістері саласында соңғы жылдары end-to-end әдістерін дамытуға, деректерді байытуға және бейнеден релевантты визуалды көріністерді алуға бағытталған арнайы алдын ала өңдеу амалдарына негізделген бірқатар зерттеулер пайда болды. Lin et al. (2023) мақаласында Gloss-Free End-to-End Sign Language Translation (GloFE) жүйесі ұсынылған – бұл аралық глосс деңгейінсіз, ым тілін тікелей мәтінге аудару жүйесі [27]. Модель бейнежазбаларды визуалды энкодер (CTR-GCN + Multi-Scale TCN) арқылы өңдеп, скелеттік белгілерді анықтап, оларды трансформерлік энкодерге береді, ал мәтіндік декодер мақсатты тілдегі сөйлемді генерациялайды. Глосс орнын толықтыру үшін conceptual anchors механизмі қолданылады: мақсатты мәтіндегі негізгі сөздер (зат есімдер мен етістіктер) «тірек» қызметін атқарады, ал cross-attention және

контрастивті триплеттік оқыту визуалды белгілерді осы концепттермен сәйкестендіреді. Модель OpenASL және How2Sign деректер жиынтықтарында глоссыз оқытылды, ал WLASL датасетінде бірқалыпты белгілерді тану үшін алдын ала үйретіліп, GloFE озық деңгейлі нәтижелер көрсетті. GloFE әдісі глоссыз ым тілін аударуда тиімді, бірақ бірқатар шектеулер бар. Әдіс деректер көлеміне және таңдалған концепттердің сапасына тәуелді, негізгі сөздерді қолмен таңдау қажет, есептеу қиындығы жоғары, маңызды емес сөздер жіберілуі мүмкін және қимылдардың барлық ерекшеліктері толық есепке алынбайды.

Ye et al. (2023) еңбектерінде end-to-end ым тілін аудару үшін кросс-модальды деректерді аугментациялау әдісін (XmDA) ұсынды, оның мақсаты бейне және мәтін белгілері арасындағы алшақтықты азайту және белгіленген деректердің жетіспеушілігінің орнын толтыру [28]. XmDA екі техниканы біріктіреді: Cross-modality Mix-up (бейне белгілері мен глосс эмбедингін араластыра отырып, репрезентацияларды теңестіру) және Cross-modality Knowledge Distillation («глосс→мәтін» модельін мұғалім ретінде пайдаланып мәтін генерациясын үйрету). Негізгі архитектура – трансформерлік SLT (СТС-классификаторымен) – глосс эмбедингтерін енгізу арқылы кеңейтіліп, аталған техникаларды қолдану арқылы оқытылды. Модель неміс тіліндегі PHOENIX-2014T және қытай тіліндегі CSL-Daily стандартты деректер жиынтықтарында оқытылып, бағаланды. Эксперименттер сапаның айтарлықтай өскенін көрсетті: мысалы, PHOENIX-2014T-де BLEU 22.79-дан 25.36-ға дейін, ал CSL-Daily-де 11.61-ден 21.58-ге дейін өсті; сонымен қатар ROUGE/ChrF метрикалары бойынша да, әсіресе сирек кездесетін сөздер мен ұзақ сөйлемдер үшін жақсарулар байқалды. Аталған әдіс кросс-модальды аугментация арқылы ым тілін аударудың сапасын арттыруда тиімді болғанмен, есептеу жүктемесінің жоғары болуы, алдын ала үйретілген мұғалім-модельдерге тәуелділігі және бастапқы деректердің сапасына сезімталдығы сияқты шектеулер бар.

Shi et al. (2022) доменді кеңейту және «табиғи ортадағы» бейнематериалдарда оқыту мәселесіне арналған зерттеуде OpenASL атты ауқымды корпусты ұсынды [29]. Ол онлайн-платформалардан (YouTube) жиналған және жүздеген сағаттық, түрлі орындаушылар мен тақырыптарды қамтитын материалдардан тұрады. Авторлар мұндай алуан деректер жиынтығында оқыту үшін қолдың, дененің және бет артикуляциясының тұрақты белгілерін шығаратын арнайы архитектуралық шешімдер, сондай-ақ бейне сапасының әркелкілігі мен домендік ығысуларға назар аудару қажеттігін көрсетті. Нәтижесінде тар шеңберлі корпустардан тыс жұмыс істей алатын «ашық» SLT жүйелерін құру мүмкін болды.

Бірқатар еңбектер табысты глоссыз жүйелердің маңызды компоненті ретінде бейне ағынын алдын ала өңдеу мен тығыз нормализациялау тәсілдеріне басым назар аударады. Әсіресе keypoint-based тәсілдер бейнені скелеттік нүктелер векторына ауыстыруды және арнайы дене бөліктеріне бағытталған нормализацияны, кадрларды стохастикалық таңдауды және аугментация процедураларын дамытуды ұсынады. Мұндай әдістер [30] дұрыс нормализация мен негізгі нүктелерді агрегациялау аударма модульдері үшін тиімді көріністер құруға мүмкіндік беретінін, әрі трансформерлік архитектуралармен және генеративті декодерлермен үйлесімділікті сақтайтынын көрсетеді.

Өз кезегінде, Avramidis және Möller (2022) ым тілін аудару тапсырмасы аясында, глостарды мәтінге аудару үшін нейрондық машиналық аударма (NMT) әдістерін зерттеді [31]. Авторлар неміс ым тілінің екі параллель корпусы – PHOENIX14T және Public DGS Corpus бойынша екі NMT архитектурасы (RNN және Transformer) арқылы тәжірибелер жүргізген. Жұмыс барысында модельдердің гиперпараметрлері, токенизация әдістері (BPE, unigram және арнайы токенизация) оңтайландырылып, деректерді аугментациялаудың екі техникасы қолданылған: кері аударма және парафразалау. Нәтижелер айтарлықтай жақсы нәтижелер көрсетті: сәйкес корпустарда оқытылған модельдер үшін BLEU көрсеткіші 5.0 және 2.2-ге өсті. Сонымен қатар, RNN модельдері Transformer модельдерінен жоғары нәтиже берді, ал сегментация үшін BPE қолдану ең тиімді болып шықты. Дегенмен, әдістің негізгі шектеулері – аннотацияланған деректердің сапасына тәуелділік және қолданыстағы NMT әдістерін ым тілінің ерекшеліктеріне бейімдеудің қиындығы.

Cao et al. (2022) Task-aware Instruction Network (TIN-SLT) тұжырымдамасын және көп-деңгейлі аугментация схемаларын ұсынды [32]. Мұнда модельге арналған «нұсқаулар» мен белгілерді біріктіру тетіктері алдын ала үйретілген трансформерлердің тілдік қабілеттерін барынша пайдалануға бағытталған. Негізгі идея – оқыту барысында модельді міндетке сәйкес динамикалық бағыттап, аугментация деңгейлері арқылы оқыту мысалдарының таралуын өзгерту. Бұл тәсіл шектеулі video→text жұптары жағдайында жалпылау қабілетін арттыруға мүмкіндік береді. Әдістің негізгі шектеуі – модельдің күрделілігінің өсуі және оқыту уақытының ұлғаюы.

Аз ресурсты тілдерде жүргізілген практикалық жұмыстар, мысалы, Sousa et al. (2023) еуропалық португал тілінен португал ым тіліне аударуға арналған жұмысы, мақсатты тілдің және жергілікті корпустың ерекшеліктерін ескере отырып, архитектураларды және аугментация әдістерін бейімдеудің қажеттілігін көрсетеді [33]. Оқыту үшін эксперттер тарапынан аннотацияланған үлгілердің алтын стандартына негізделген көпдоменді LGP-5-Domain корпус жасалды. Эксперименттер көрсеткендей, ұсынылған әдіс жалғыз аналог жүйе PE2LGP-дан асып түсіп, BLEU және басқа аударма метрикалары бойынша жоғары көрсеткіштерге жетті. Негізгі шектеу – аннотациялардың сапасына тәуелділік және NMT моделін ым тілінің ерекшеліктеріне бейімдеудің күрделілігі.

Жүргізілген шолу ым тілдерінің машиналық аудармасы саласындағы зерттеулердің екі бағытта дамып отырғанын көрсетеді: біріншісі – глостарды аралық деңгей ретінде қолдану, екіншісі – толық глоссыз әдіс. Екінші бағыт терең оқыту саласындағы прогрес пен ірі мультимодальды модельдердің пайда болуына байланысты танымалдылығы күннен-күнге артып жатса да, дәл қазіргі уақытта глосарлық әдістер тұрақты нәтижелерді көрсетуде. Глостар сөйлем құрылымын формализациялауға, қимылдарды өңдеу кезіндегі көпмағыналықты азайтуға және машиналық аударма архитектуралары үшін модельдеуді жеңілдетуге мүмкіндік береді. Сонымен қатар, глосарлық аннотациялардың болуы модельдерді интерпретациялауға және ашық көрсетуге мүмкіндік береді, бұл жүйелерді білім беру немесе инклюзивті цифрлық шешімдерде қолдану кезінде ерекше маңызды.

Қолжетімді ым тілі корпустарының шектеулі және әркелкі жағдайында глосарлық әдістер деректерді стандарттауға және салыстырмалы эксперименттер жүргізуге мүмкіндік бере отырып, негізгі рөл атқара бермек. Бұл зерттеушілер қауымдастығына әрі қарайғы инновациялар мен салыстырмалы талдаулар үшін берік негіз қалыптастырады. Сондықтан, алдағы перспективада глостар академиялық зерттеулерде де, қолданбалы міндеттерде де маңызды элемент болып қала береді деп қорытындылауға болады, ал одан әрі прогресс, ең алдымен, глосарлық және глоссыз әдістерді интеграциялауға байланысты болады, мұнда глостар сенімді негіз ретінде қызмет атқаратын болады.

Мұндай тұжырымдар глосарлық әдістің қазіргі таңда өзінің маңыздылығын жоғалтпағанын және практикалық қолдануда тұрақты нәтижелер беретінін айқын көрсетеді. Осы контекстте аталған жұмыста глостарды аралық деңгей ретінде пайдалануға негізделген тәсілді таңдалдық. Тек модельдің құрылымын формализациялау ғана емес, сонымен қатар қазақ тілінің морфологиялық ерекшеліктерін ескеріп, ым тіліндегі мағыналық сәйкестікті нақты беруге бағытталған әдіс. Төменде біздің жүйеде қолданылған мәліметтерді алдын ала өңдеу және модельді оқыту кезеңдері егжей-тегжейлі сипатталады.

Материалдар мен әдістер

Модельді оқытуға арналған корпусты дайындау барысында оның лингвистикалық тазартылуына ерекше назар аударылды, өйткені қазақ тілінің құрылымында ым тіліне тікелей баламасы жоқ және модельді оқытуда қиындық тудыруы мүмкін элементтер бар. Мұндай ерекшеліктердің бірі – көмекші етістік формалары, мысалы, жатыр, тұр, отыр және басқа да формалар. Олар жазбаша тілде әрекеттің күйі немесе түрін білдіру үшін қолданылады, бірақ ым тілінде түсіп қалады, өйткені негізгі мағына олардың алдында тұрған негізгі етістікпен

беріледі. Деректерді өңдеу барысында мұндай көмекші сөздер алып тасталып, сөйлемде тек негізгі етістік қалдырылды.

Сондай-ақ ым тілінде қолданылмайтын жалғаулықтар мен шылау сөздер, мысалы, мен, және («и»), сондай-ақ ым тілдік контексте жеке мағыналық жүк артпайтын басқа да ұқсас элементтер де алынып тасталды. Мұндай сүзгілеу кіріс деректерін жеңілдетіп, модельдің негізгі мағыналық бірліктерге назар аударуына мүмкіндік берді, бұл мәтін мен глосс арасындағы сәйкестікті дәлірек меңгеруге ықпал етті.

Деректерді лингвистикалық тазарту бойынша негізгі ережелер 1-кестеде көрсетілген. Онда бастапқы сөйлемдердің және алдын ала өңдеу кезеңдерінен кейінгі түрлендірілген нұсқалардың мысалдары ұсынылған.

Кесте 1 – Корпусты лингвистикалық алдын ала өңдеу ережелерінің мысалдары

№	Алдын ала өңдеу ережесі	Бастапқы мәтін үзіндісі	Нормаланған нұсқасы	Ескерту
1	Көмекші етістіктерді алып тастау (отыр, жатыр, тұр, жүр)	Мен кітап оқып отырмын	Мен кітап оқимын	Ым тілінде шақ пен түр негізгі етістікпен беріледі
2	Шартты бөлшектерді алып тастау (-ақ, -ау, -ай, -да/-де/-та/-те)	Осы-ақ менің досым-ай келді	Осы менің досым келді	Бұл бөлшектер эмоция/күшейтуді білдіреді, ым тілінде мимикамен беріледі
3	-мыс, -міс жалғаулы кіріспелерді алып тастау	Ол айтыпты-мыс	Ол айтыпты	Ым тілінде бұл болжам мағыналары қолданылмайды
4	Постпозициялық формаларды алып тастау (-ды/-ді, -ты/-ті)	Айтқан-ды кеше	Айтқан кеше	Бұл формалар стильдік рөл атқарады, мағыналық негізге әсер етпейді
5	Күшейткіш сын есімдерді бастапқы түрге келтіру	әп-әдемі, жып-жылы, аппақ	әдемі, жылы, ақ	Ым тілінде күшейткіш мағына мимикамен не қимылмен беріледі

Модельді оқыту

Мәтінді ым тіліне аудару жүйесіндегі модельді оқыту кезеңі Seq2Seq архитектурасы және трансформер негізіндегі назар (attention) механизмі арқылы жүзеге асырылды. Бұл тәсіл машиналық аударма мен тізбектерді түрлендіру мәселелерін шешуде жоғары тиімділікті қамтамасыз етеді.

Негізгі модель ретінде ByT5-small қолданылды – бұл T5 моделінің байттық модификациясы, ол көптілді мәтіндермен жұмыс істей алады және сирек кездесетін таңбаларды жеке токен сөздігісіз-ақ тікелей өңдеуге қабілетті. Энкодер-декодер құрылымының арқасында бұл модель табиғи тілдегі сөйлемдерді формализацияланған глоссаларға түрлендіруге қолайлы, ал бұл – ымдарды генерациялауға дейінгі негізгі қадам.

Оқыту және инференс параметрлерінің толық конфигурациясы 2-кестеде ұсынылған.

Кіріс және мақсатты тізбектердің токенизациясы ByT5 ішкі токенизаторы арқылы жүзеге асырылды: ұзындығы 128 токенге дейін қысқартылып, әрі қарай бекітілген өлшемге дейін толықтыру (padding) жүргізілді. Мұндай тәсіл бірдей ұзындықтағы batch-терді қалыптастыруға мүмкіндік беріп, есептеуді тиімдірек етіп, оқыту үдерісін жеделдетті.

Мақсатты глоссалар токен идентификаторлары түрінде кодталып, модельге labels параметрі арқылы берілді – бұл жүйеге кіріс және шығыс деректерді тікелей салыстыру арқылы үйренуге мүмкіндік берді.

Оқыту Hugging Face Transformers кітапханасының көмегімен және Seq2SeqTrainer класы арқылы жүргізілді, бұл процесті оңай баптауға және бақылауға жол ашты.

Негізгі гиперпараметрлер:

- ♦ Learning rate: 5×10^{-4} ;
- ♦ Batch size: 8 (оқыту және валидация үшін);
- ♦ Эпох саны: 10;
- ♦ Weight decay: 0.01 (модельдің жалпылау қабілетін арттыру және артық үйренудің алдын алу үшін).

Batch-терді қалыптастыру үшін DataCollatorForSeq2Seq пайдаланылды, ол тізбектердің ұзындығын динамикалық толықтыру арқылы артық паддингті азайтты.

Әр эпох аяқталған соң модельдің сапасы бағаланды. Бұл ретте тек соңғы бақылау нүктесі (checkpoint) сақталды – бұл диск кеңістігін үнемдеуге және аралық нұсқалардың жинақталуын болдырмауға мүмкіндік берді.

Инференс кезеңінде ең ықтимал глосс тізбегін табу үшін beam search әдісі қолданылды, 5 сәуле (beam) параметрімен. Нәтиже сапасын арттыру үшін қосымша шектеулер енгізілді:

n-грамманың қайталануына тыйым (ұзындығы 3);

Қайталануға айыппұл – мәні 1.2;

Ерте тоқтау – толық аудармаға жеткен кезде.

Бұл тәсіл шығыс тізбектеріндегі артықтықты азайтып, байланыстылықты күшейтіп, ым тіліне тән құрылымды сақтауға көмектесті.

Нәтижелер

Оқыту, 1-суретте көрсетілгендей, 10 эпохаға дейін жүргізілді. Негізгі мәтінде тек 1–5-эпохалардың нәтижелері көрсетілді, себебі 6–10-эпохалардың көрсеткіштері қосымша материалға шығарылды. Валидациялық шығын 1–3-эпохаларда тұрақты төмендеп, минималды мәнге 3-эпохада жетті де (≈ 0.2565), 4–5-эпохаларда қосымша айтарлықтай жақсару байқалмады (≈ 0.260 – 0.261).

[1040/1040 16:06, Epoch 10/10]

Epoch	Training Loss	Validation Loss
1	No log	0.303398
2	No log	0.270347
3	No log	0.256496
4	No log	0.261146
5	0.898300	0.260486
6	0.898300	0.268210
7	0.898300	0.266074
8	0.898300	0.280682
9	0.898300	0.287509
10	0.155400	0.294966

Сурет 1 – Оқыту нәтижелері

Осылайша, модельдің ең тиімді күйі валидациялық шығынның ең төменгі мәні тіркелген 3-эпохада анықталды. Тренинг шығынының 1–4-эпохалар аралығында тіркелмеуі, яғни «no log» ретінде белгіленуі себепті, модельді бағалау және таңдау тек валидациялық көрсеткіштерге негізделіп жүзеге асырылды.

Сапалық мысалдар 2-суретте ұсынылғандай, (мәтін→глосс) көмекші етістіктер мен эмоциялық бөлшектердің түсіріліп, мағыналық өзектің сақталатынын көрсетті.

текст: Сәлеметсіздер ме, құрметті оқушылар!
гloss: Сәлем, құрмет оқушы!

текст: Бүгінгі сабаққа қош келдіңіздер.
гloss: бүгін сабақ қош келу

текст: Бүгінгі сабағымыздың тақырыбы – Жер ғаламшарының пайда болуы.
гloss: Бүгін сабақ тақырып –Жер ғаламшар пайда болу.

текст: Бүгінгі сабағымыздың мақсаты – Жер ғаламшарының пайда болуын түсіну.
гloss: Бүгін сабақ мақсат – Жер ғалам пайда болу түсіну.

текст: Адамдар ежелден өзін қоршаған әлемді, оның ішінде ең бірінші, барлығымыздың ортақ үйіміз – Жер шарын танып-білуге ұмтылған.
гloss: Адам ежелден өз қоршау әлем – Жер шар тану білу ұмтылу.

текст: Бізге Жердің қалай пайда болғаны туралы әртүрлі ұлттардан көптеген аңгімелер мен мифтер жеткен.
гloss: Біз Жер қалай пайда болды туралы әр түрлі ұлт көп аңгіме мифтер болды

текст: Ежелгі адамдар Жер, Күннен және басқа жұлдыздардан да үлкен деп ойлаған.
гloss: Ежелгі адам Жер, Күн басқа жұлдыз үлкен ойлау болу

текст: Адамдарда ғибадатхана мен пирамида салу тәжірибесі болған.
гloss: Адам ғибадатхана пирамида салу тәжірибе болды

текст: Алайда Жердің қалай пайда болғаны туралы білім жоққа тең еді.
...
гloss: Құдай Жер адам көңіл шығу бар

Сурет 2 – Сапалық мысалдар

Талқылау

Әдеби шолу ым тілін машиналық аудару зерттеулерінің глосты аралық деңгей ретінде қолданатын тәсілдер және глоссиз end-to-end әдістер бағытын көрсетеді. Соңғы жылдары ірі мультимодальды модельдерге сүйенетін глоссиз бағыт танымал болғанымен, қазірше глосқа негізделген әдістер тұрақты нәтижелер көрсетіп келеді. Бұл зерттеуде таңдаған интерпретацияланатын глосаралық қабат бағытымен үйлеседі.

Глостар сөйлем құрылымын формализациялап, қимылды өңдеу кезіндегі көпмағыналылықты азайтады және классикалық нейрондық машиналық аударма және тізбектен-тізбекке негізделген архитектураларына бейімдеуді жеңілдетеді. Бұған қоса, глосарлық аннотация интерпретациялану мен ашықтықты қамтамасыз етеді, бұл білім беру және инклюзивті цифрлық шешімдер үшін маңызды артықшылық болып саналады.

Lin және әріптестері ұсынған GloFE жүйесі тұжырымдық тірек нүктелер арқылы ымнан тікелей мәтін генерациялайды [27]. Жаксы нәтижелеріне қарамастан, әдіс дерек көлеміне, «тірек» сөздерді таңдаудың сапасына және есептеу ресурстарына тәуелді, яғни кейбір маңызсыз сөздер жоғалуы мүмкін, ал қимылдың барлық ерекшеліктері түгел ескерілмейді.

End-to-end тәсілдерде деректердің жетіспеушілігін азайту мақсатында Ye және әріптестері кросс-модальды аугментация әдісін ұсынды. Бұл тәсіл бейне мен мәтіннің репрезентацияларын бір кеңістікте сәйкестендіріп үйретеді, сондай-ақ алдын ала үйретілген глостан мәтінге түрлендіретін моделді мұғалім ретінде қолданып, негізгі end-to-end модельге қосымша білім береді. Осылайша, модель визуалды және тілдік ақпарат арасындағы байланысты тиімдірек үйрене алады. Бұл тәсіл гибриді бағыт құруға тиімді [28].

Sousa және әріптестерінің өз зерттеуінде мақсатты тіл мен жергілікті корпустың ерекшеліктерін ескеріп, модель құрылымын және деректерді толықтыру тәсілдерін бейімдеу бейімдеу қажет екенін көрсетті [33]. Бұл нәтиже KSL секілді тілдер үшін домендік корпусты өсіру мен сапалы аннотацияны бірінші орынға қою керектігін көрсетеді.

Осы зерттеуді ұсынылған глосқа негізделген құбыр аз ресурсты жағдайда интерпретацияланатын әрі тұрақты негіз береді. Кеңістіктік және вербалды емес (non-manual) белгілерді ескеретін толыққанды модель қалыптастыру үшін, шолуда ұсынылған end-to-end тәсілдердің жекелеген механизмдерін глосс негізіндегі архитектурамен гибридік үйлестіру келешегі зор бағыт болады. Осылайша, глосс сенімді өзек ретінде сақталып, end-to-end компоненттер күрделі көрнекі белгілерді меңгеруді күшейтеді.

Қорытынды

Бұл зерттеу seq2seq/byt5-small үлгісінің қазақ тіліндегі мәтінді ымдау тілі глостарына аударудың тиімділігін көрсетеді. Seq2Seq архитектурасы және byt5-кіші үлгісі қазақша сөйлемдерді байт деңгейінде өңдейді, нәтижесінде формальданған глостарға тұрақты түрлендірулер жасалады. Глостау тәсілі ымдау тілінің қозғалыстарын жазбаша тілге ұқсас логикалық бірліктер ретінде көрсететін интерпретацияланатын аралық бағдарлама қабаты ретінде қызмет етеді. Бұл құрылым дәстүрлі нейромашиналық аударма архитектурасын пайдалануға мүмкіндік береді және модель әрекетін талдауды жеңілдетеді.

Модель арқылы жасалған қысқартылған глостар реті ымдау тілін көрнекі көрсететін қимылдарды синтездеу үшін аватар жүйелеріне біріктірілуі мүмкін. Жоба мәтінді глостарға және мультимедиялық форматтағы ымға түрлендіру үдерісін қарастырды. Аударылған глостар тізбегі ымдау тілінің құрылымын сақтай отырып, артықшылықты азайтады және сөйлемнің үйлесімділігін жақсартады. Демек, ымдау тілінің аватарларын басқару үшін тек семантикалық мағыналы бірліктер ғана пайдаланылады, бұл есептеу талаптарын азайтады. Қысқа және мағыналық жағынан бай глостар ым тілінде қол қимылдарын синтездеудің тиімді ресурсы қызметін атқарады.

Глосқа негізделген тәсілдің бірнеше шектеулері бар. Глостар ең алдымен ымдар тізбегін көрсетеді және ымдау тілінің кеңістіктік компоненттерін, мимикаларын немесе вербалды емес белгілерді толық түсірмейді. Бұл экспрессивті құрылымды сипаттаудың дәлдігін шектейді. Болашақ зерттеулер глостарға негізделген және глостарға негізделмеген әдістерді біріктіретін гибриді үлгілерді әзірлеуге назар аударуы керек. Мысалы, қазіргі зерттеулер кросс-модальды күшейту және мұғалім мен оқушыны оқыту әдістері арқылы глосс пен бейне сәйкестігін жақсарта алады. Бұл тәсілдер ымдау тілінің кеңістіктік және экспрессивтік ерекшеліктерін дәлірек көрсету үшін екі жүйенің де артықшылықтарын біріктіру арқылы машиналық аударманың сапасын жақсартуға бағытталған.

Қаржыландыру туралы ақпарат. Бұл зерттеу Қазақстан Республикасының Ғылым және жоғары білім министрлігінің Ғылым комитеті тарапынан қаржыландырылды (Грант № BR24992875).

ӘДЕБИЕТТЕР

- 1 FAQ. WFD – World Federation of the Deaf. URL: <https://wfdeaf.org/contact/faqs/> (date of access: 20.09.2025).
- 2 Human rights. Deaf History Europe. URL: <https://deafhistory.eu/index.php/component/zoo/item/human-rights#:~:text=,from%20deaf%20people%E2%80%99s%20human%20rights> (date of access: 20.09.2025).
- 3 Six difficulties: Sign language avatar. DW Innovation. URL: <https://innovation.dw.com/articles/six-difficulties-sign-language-avatar> (date of access: 20.09.2025).
- 4 Zou, Y., Lin, H. A basic General Service List for Chinese Sign Language. Journal of Deaf Studies and Deaf Education, 30(3), 405–418 (2025). <https://doi.org/10.1093/jdsade/enaf012>.
- 5 Caselli, N.K. et al. ASL-LEX: A lexical database of American Sign Language. Behavior research methods, 49(2), 784–801 (2017). <https://doi.org/10.3758/s13428-016-0742-0>.
- 6 Perlman, M. et al. Iconicity in signed and spoken vocabulary: a comparison between American Sign Language, British Sign Language, English, and Spanish. Frontiers in psychology, 9, 1433 (2018). <https://doi.org/10.3389/fpsyg.2018.01433>.
- 7 Zhao, Y. et al. Relationship between vocabulary knowledge and reading comprehension in deaf and hard of hearing students. The Journal of Deaf Studies and Deaf Education, 26(4), 546–555 (2021). <https://doi.org/10.1093/deafed/enab023>.
- 8 Convertino, C. et al. Word and world knowledge among deaf learners with and without cochlear implants. Journal of Deaf Studies and Deaf Education, 19(4), 471–483 (2014). <https://doi.org/10.1093/deafed/enu024>.

9 Hall, W.C. What you don't know can hurt you: The risk of language deprivation by impairing sign language development in deaf children. *Maternal and child health journal*, 21(5), 961–965 (2017). <https://doi.org/10.1007/s10995-017-2287-y>.

10 Deaf community: Sign language equals rights. Human Rights Watch URL: <https://www.hrw.org/news/2022/09/23/deaf-community-sign-language-equals-rights> (date of access: 20.09.2025).

11 AI and machine translation: A threat to the deaf community. Deaf Journalism URL: <https://www.deafjournalism.eu/ai-and-machine-translation-a-threat-to-the-deaf-community/> (date of access: 20.09.2025).

12 Duarte, A. et al. How2sign: a large-scale multimodal dataset for continuous american sign language. *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2735–2744 (2021). <https://doi.org/10.1109/CVPR46437.2021.00276> Dataset: <http://how2sign.github.io/> openaccess.thecvf.com/how2sign.github.io.

13 Li, D. et al. Word-level deep sign language recognition from video: A new large-scale dataset and methods comparison. *Proceedings of the IEEE/CVF winter conference on applications of computer vision*, 1459–1469 (2020). URL: https://openaccess.thecvf.com/content_WACV_2020/papers/Li_Word-level_Deep_Sign_Language_Recognition_from_Video_A_New_Large-scale_WACV_2020_paper.pdf. Dataset: <https://dxli94.github.io/WLASL/> openaccess.thecvf.com/dxli94.github.io.

14 Athitsos, V. et al. The american sign language lexicon video dataset. 2008 IEEE computer society conference on computer vision and pattern recognition workshops. IEEE, 2008, pp. 1–8. Dataset: <https://www.bu.edu/asllrp/av/dai-asllvd.html>.

15 Albanie, S. et al. Bbc-oxford british sign language dataset. 2021. arXiv preprint arXiv:2111.03635 (2024). URL: <https://www.robots.ox.ac.uk/~vgg/data/bobsl/> (see arXiv:2111.03635). robots.ox.ac.uk/arXiv.

16 Zhou, H. et al. Improving sign language translation with monolingual data by sign back-translation. *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition* (2021), pp. 1316–1325. Dataset: CSL-Daily, https://ustc-slr.github.io/datasets/2021_csl_daily/ openaccess.thecvf.com/Visual_Sign_Language_Research_Group.

17 Ham, S. et al. Ksl-guide: A large-scale korean sign language dataset including interrogative sentences for guiding the deaf and hard-of-hearing. 2021 16th IEEE International Conference on Automatic Face and Gesture Recognition (FG 2021). IEEE, 2021, pp. 1–8. Dataset: <https://github.com/ChelseaGH/KSL-Guide> (also listed at sign-lang@LREC). GitHub/sign-lang.uni-hamburg.de.

18 Cormier, K., & Schembri, A. British Sign Language Corpus [Data collection]. UK Data Archive, 2018. <https://doi.org/10.5255/UKDA-SN-851521> reshare.ukdataservice.ac.uk.

19 Belissen, V., Braffort, A., Gouiffès, M. Dicta-Sign-LSF-v2: remake of a continuous French sign language dialogue corpus and a first baseline for automatic sign language processing. *LREC 2020, 12th Conference on Language Resources and Evaluation* (2020). URL: <https://aclanthology.org/2020.lrec-1.740/>. Dataset page: <https://www.sign-lang.uni-hamburg.de/lr/compendium/corpus/dictasignlsfv2.html> aclanthology.org/sign-lang.uni-hamburg.de.

20 Camgoz, N. C. et al. Neural sign language translation. *Proceedings of the IEEE conference on computer vision and pattern recognition*, 7784–7793 (2018). URL: <https://www-i6.informatik.rwth-aachen.de/~koller/RWTH-PHOENIX-2014-T/> www-i6.informatik.rwth-aachen.de.

21 Konrad, R., Hanke, T., Langer, G., Blanck, D., Bleicken, J., Hofmann, I., ... Schulder, M. MEINE DGS – annotiert. Public Corpus of German Sign Language, 3rd release [Dataset]. Universität Hamburg, 2020. <https://doi.org/10.25592/dgs.corpus-3.0> sign-lang.uni-hamburg.de

22 Ye, J. et al. Scaling back-translation with domain text generation for sign language gloss translation. arXiv preprint arXiv:2210.07054 (2022) <https://doi.org/10.18653/v1/2023.eacl-main.34>.

23 Zhu, D., Czehmann, V., Avramidis, E. Neural machine translation methods for translating text to sign language glosses. *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, 12523–12541 (2023). <https://doi.org/10.18653/v1/2023.acl-long.700>.

24 Xu, C. et al. Automatic gloss dictionary for sign language learners. *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics: System Demonstrations*, 83–92 (2022). <https://doi.org/10.18653/v1/2022.acl-demo.8>.

25 Mohamed, A. et al. A deep learning approach for gloss sign language translation using transformer. *Journal of Computing and Communication*, 1(2), 1–8 (2022).

26 Müller, M. et al. Considerations for meaningful sign language machine translation based on glosses. arXiv preprint arXiv:2211.15464 (2022).

27 Lin, K. et al. Gloss-free end-to-end sign language translation. arXiv preprint arXiv:2305.12876 (2023). <https://doi.org/10.18653/v1/2023.acl-long.722>.

28 Ye, J. et al. Cross-modality data augmentation for end-to-end sign language translation. arXiv preprint arXiv:2305.11096 (2023). <https://doi.org/10.18653/v1/2023.findings-emnlp.904>.

29 Shi, B. et al. Open-domain sign language translation learned from online video. arXiv preprint arXiv:2205.12870 (2022).

30 Kim, Y., Baek, H. Preprocessing for keypoint-based sign language translation without glosses. Sensors, 23(6), 3231 (2023).

31 Angelova, G., Avramidis, E., Möller, S. Using neural machine translation methods for sign language translation. Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics: Student Research Workshop, 273–284 (2022).

32 Cao, Y. et al. Explore more guidance: A task-aware instruction network for sign language translation enhanced with data augmentation. arXiv preprint arXiv:2204.05953 (2022).

33 Sousa, C., Coheur, L., Moita, M. Enhancing Accessible Communication: from European Portuguese to Portuguese Sign Language. Findings of the Association for Computational Linguistics: EMNLP 2023, 11452–11460 (2023).

^{1,3,4}**Амангелді Н.,**

PhD, ORCID ID 0000-0002-4669-9254,

e-mail: amangeldi_n_3@enu.kz

^{1,2,*}**Еримбетова А.,**

к.т.н., ассоциированный профессор, PhD,

ORCID ID: 0000-0002-2013-1513,

*e-mail: aigerian8888@gmail.com

^{1,4}**Газизова Н.Е.,**

магистр, ORCID ID: 0009-0007-7278-4202,

e-mail: g.nazerke2755@gmail.com

^{1,3}**Турсынова Н.А.,**

магистр, ORCID ID: 0000-0002-1857-9028,

e-mail: tursynova_na@enu.kz

⁴**Болатбекқызы К.,**

бакалавр, ORCID ID: 0009-0000-7659-4974

e-mail: kerbezbolatbekkyzy03@gmail.com

¹Институт информационных и вычислительных технологий, г. Алматы, Казахстан

²Евразийский технологический университет, г. Алматы, Казахстан

³Евразийский национальный университет им. Л.Н. Гумилева, г. Астана, Казахстан

⁴ТОО «SignBridge», г. Астана, Казахстан

ФОРМАЛИЗАЦИЯ ТЕКСТА НА КАЗАХСКОМ ЯЗЫКЕ С ИСПОЛЬЗОВАНИЕМ ГЛОССАРНОГО СЛОЯ

Аннотация

Технологии автоматической обработки жестового языка стали актуальной потребностью членов общества с нарушениями слуха и речи, которые сталкиваются с неравенством в эпоху цифровой трансформации. В последние годы вопрос рассмотрения жестового языка как формальной структуры, равной естественному языку, и его адаптации к автоматическим системам привлекает повышенное внимание исследователей. Для выполнения задачи автоматического перевода информации с естественного языка на жестовый язык в качестве промежуточного слоя используются глоссы, то есть текстовая форма жестового языка. В данном исследовании предлагается новый метод преобразования текста на казахском языке, учитывающий морфологические особенности казахского языка, в глоссы жестового языка с использованием методов обработки естественного языка. В частности, применяется архитектура Seq2Seq на основе модели ByT5 small. Полученные результаты показывают, что сформированные последовательности глосс являются компактными и семантически насыщенными, сохраняя внутреннюю структуру жестового языка. Последовательность глосс позволяет автоматизировать работу интерпретируемого промежуточного слоя, представляющего жестовые движения как логические единицы, аналогичные письменному языку. Преобразованная последовательность

гloss сохраняет структуру жестового языка, уменьшает избыточность и повышает связность предложения. Таким образом, использование только семантически значимых единиц при управлении аватарами жестового языка снижает вычислительные затраты. Короткие и семантически насыщенные глоссы являются эффективным ресурсом для синтеза движений рук в жестовом языке.

Ключевые слова: жестовый язык, последовательность глосс, обработка естественного языка, формализация, Seq2Seq, ByT5.

^{1,3,4}**Amangeldy N.,**

PhD, ORCID ID: 0000-0002-4669-9254,

e-mail: amangeldi_n_3@enu.kz

^{1,2*}**Yerimbetova A.,**

Cand. Tech. Sc., PhD, Associate Professor, ORCID ID: 0000-0002-2013-1513,

*e-mail: aigerian8888@gmail.com

^{1,4}**Gazizova N.,**

MSc, ORCID ID: 0009-0007-7278-4202,

e-mail: g.nazerke2755@gmail.com

^{1,3}**Tursynova N.,**

MSc, ORCID ID: 0000-0002-1857-9028,

e-mail: tursynova_na@enu.kz

⁴**Bolatbekkyzy K.,**

Bachelor, ORCID ID: 0009-0000-7659-4974,

e-mail: kerbezbolatbekkyzy03@gmail.com

¹Institute of Information and Computational Technologies, Almaty, Kazakhstan

²Eurasian Technological University, Almaty, Kazakhstan

³L.N. Gumilyov Eurasian National University, Astana, Kazakhstan

⁴«SignBridge» LLP, Astana, Kazakhstan

DEVELOPMENT OF A MODEL FOR REAL-TIME RECOGNITION OF KAZAKH SIGN LANGUAGE USING MEDIAPIPE AND DEEP LEARNING METHODS

Abstract

Technologies for automatic processing of sign language have become an urgent need for members of society with hearing and speech impairments who face inequality in the era of digital transformation. In recent years, the issue of considering sign language as a formal structure equal to natural language and adapting it to automatic systems has attracted increasing attention from researchers. To perform the task of automatically translating information from natural language into sign language, glosses, which are the textual representation of sign language, are used as an intermediate layer. For this purpose, this study proposes a new method for converting Kazakh language text, which reflects the morphological features of the Kazakh language, into sign language glosses using natural language processing techniques. In particular, a Seq2Seq architecture based on the ByT5 small model is applied. The obtained results demonstrate that the generated gloss sequences are compact and semantically rich while preserving the internal structure of sign language. The gloss sequence makes it possible to automate the work of an interpretable intermediate layer that represents sign language movements as logical units similar to written language. The transformed gloss sequence preserves the structure of sign language, reduces redundancy, and improves sentence coherence. Thus, the use of only semantically meaningful units to control sign language avatars reduces computational requirements. Short and semantically rich glosses serve as an effective resource for synthesizing hand movements in sign language.

Keywords: sign language, gloss sequence, natural language processing, formalization, Seq2Seq, ByT5.

Мақаланың редакцияға түскен күні: 01.10.2025

UDC 004.891
IRSTI 28.23.35

<https://doi.org/10.55452/1998-6688-2025-22-4-168-177>

¹**Ashim Zh.,**

Master's student, ORCID ID: 0009-0003-1354-1105,

e-mail: zh_ashim@kbtu.kz

²**Botanov A.,**

Master's student, ORCID ID: 0009-0008-1349-7614

e-mail: botanov.a@stud.satbayev.university

²**Abdoldina F.,**

PhD, Associate Professor, ORCID ID: 0000-0003-1816-6343

e-mail: abdoldinafarida@gmail.com

^{3*}**Serek A.,**

PhD, Associate Professor, ORCID ID: 0000-0001-7096-6765

*e-mail: azamat_serek97@gmail.com

¹Kazakh-British Technical University, Almaty, Kazakhstan

²Institute of Automatics and Information Technologies, Satbayev University,
Almaty, Kazakhstan

³Astana IT University, Astana, Kazakhstan

SALARY PREDICTION FROM JOB DESCRIPTIONS USING ATTENTION-BASED NLP MODELS

Abstract

The research introduces a dual deep learning system which predicts salary ranges by processing job descriptions through BERT-based contextual embeddings and structured metadata integration. The proposed method utilizes more than 124,000 LinkedIn job postings to merge BERT-based contextual embeddings with structured information about location and industry and experience level and compensation type. The model uses multi-head attention to identify essential salary-related terms in job descriptions which results in better model interpretability and improved prediction accuracy. The model combines semantic embeddings with tabular data to create a multimodal representation which serves as input for supervised learning with an ordinal-aware loss function. The model achieves stable performance in salary classification across three categories through F1-scores between 0.82 and 0.84. The proposed model achieves excellent generalization capabilities for different sectors and job types while providing precise predictions and clear decision-making processes for salary benchmarking and recruitment analytics applications.

Keywords: Salary Prediction, Job Descriptions, Natural Language Processing (NLP), BERT Embeddings, Attention Mechanism.

Introduction

Job descriptions which were once considered basic documents about work duties now function as valuable yet underleveraged resources which expose organizational values and compensation systems and performance standards. The rising need for labor market transparency during times of economic instability makes it essential to extract hidden information from job postings. The current salary prediction models which use demographic information and job titles as inputs fail to detect the semantic content present in job posting texts [1–2].

The recent progress in natural language processing technology enables researchers to perform detailed analysis of job descriptions which reveals hidden salary indicators. The field of artificial intelligence and natural language processing continues to expand its capabilities through research

that includes fraud detection systems and sign language recognition systems and machine learning models for career choice prediction [3–6]. The research demonstrates that semantic modeling and contextual inference methods can effectively work for labor market analytics through the combined efforts of these studies.

The scientific novelty of this research resides in the development and empirical validation of a hybrid deep learning architecture that synergistically integrates Transformer-based contextual embeddings (specifically from the BERT model) with conventional structured metadata for enhanced regression-based salary prediction. By employing multi-head attention mechanisms over the textual data, the proposed model significantly advances beyond prevailing NLP methods that rely on simpler bag-of-words or basic recurrent neural network representations, achieving a fine-grained semantic understanding of job requirement complexities. This approach facilitates the attenuation of noise and the differential weighting of salient salary indicators within unstructured job descriptions, thereby establishing a new benchmark for interpretability and predictive accuracy in labor market analytics and automated compensation forecasting.

The research develops an NLP system which evaluates 124,000 LinkedIn job postings from 2023 to 2024 to predict salaries. The proposed framework combines BERT transformer model contextual embeddings with structured metadata elements including company information and geographic locations and job classification to develop an efficient and understandable model. The system employs multi-head attention mechanisms to understand job description complexities while supervised learning techniques generate salary band predictions for different industries and geographic areas.

Research studies have extensively analyzed salary prediction through conventional machine learning approaches. Wang et al. studied multiple regression models to determine that polynomial regression produced the best results for modeling salary patterns that deviate from linearity [7]. The research by Kablaoui et al. tested linear regression and random forests and neural networks on U.S. salary data which showed neural networks reached 83.2% accuracy but linear models processed data more quickly [8]. Satpute et al. built upon previous work by adding demographic factors such as age and gender and education level and ethnicity to explain income differences which matches current demands for fair salary analytics [9].

Researchers focus on extracting valuable information from job postings to understand how the labor market operates. The research by Safi et al. analyzed how job postings show current skill requirements for big data and analytics positions [10]. The research by Korytov et al. examined skill extraction techniques through keyword extraction and named entity recognition methods including YAKE and PositionRank and WikiNEuRal [11]. Alsayed et al. developed a comprehensive system which merges machine learning techniques with news analysis to predict future labor market developments [12].

The use of NLP technology for recruitment workflow optimization has become a subject of rising interest. Sharma et al. created an automated candidate screening platform while Lalitha and Warusawithana developed systems to parse resumes based on their layout structure [13–15]. The voting ensemble model developed by Yaphet et al. demonstrates the need for secure recruitment systems by detecting fake job postings [16].

Most current salary prediction systems have restricted functionality because they depend mainly on job title and experience data while using minimal unstructured job text information [17–18]. The research by Joshi et al. and Thapa et al. shows that job description context holds untapped predictive value although it remains understudied [19–20]. The research shows that regional and company-specific elements have a proven impact on the labor market yet these factors rarely get combined with semantic text analysis in single models.

The research by Chandra et al. and Rahman et al. demonstrates that current frameworks lack a unified system which combines text feature extraction with structured metadata and sophisticated prediction algorithms [23–24]. The growing number of remote and hybrid work arrangements requires more research about their effects on employee compensation.

The research combines multiple interdisciplinary approaches which include BERT-based contextual embeddings and cost-of-living normalization and hierarchical feature encoding to address current research gaps. The research builds upon existing AI applications across different domains including language recognition and fraud detection and career prediction systems to show deep learning and NLP can solve social problems.

Our contribution is the development of a dual deep learning framework that combines BERT-based contextual text embeddings with structured metadata features to predict salary ranges from job postings. We modified the BERT model to work with job description data that is specific to a certain field in order to get deep semantic representations. We also created a multi-head attention mechanism to highlight language features related to compensation, such as skills, roles, and experience requirements. We also came up with a feature fusion architecture that combines attention-weighted textual embeddings with structured variables like location, industry, company size, and level of experience. We used an ordinal-aware loss function that was optimized with the AdamW algorithm and L2 regularization to make training more stable and predictions more reliable. Our framework offers a cohesive multimodal representation that enhances model interpretability, generalization, and accuracy in salary classification across various sectors and job categories.

Materials and methods

The following section presents the complete process for determining salary ranges through job description analysis. The first part describes the dataset and explains all preprocessing operations which transform textual and structured data into usable form. The BERT model serves as our embedding method to extract salary-related information from text before applying multi-head attention for term identification. The system combines contextual embeddings with structured data elements that include industry information and location details and experience requirements. The prediction model architecture receives detailed explanation which includes loss function selection and optimization methods and interpretability approaches. The researchers obtained their dataset from LinkedIn job postings which they made publicly available [25]. The dataset contains structured and unstructured elements provide essential information for salary prediction. Table 1 gives a quick overview of the most important features of the LinkedIn Job Postings dataset that was used to predict salary ranges. The dataset has both structured and unstructured fields, which give each job posting a lot of context. Natural Language Processing (NLP) tasks use textual data, like description and skills_desc. Structured features, like location, industry, and experience level, help with multimodal learning and making models easier to understand.

Table 1 – Overview of the LinkedIn Job Postings Dataset

Category	Feature Examples	Type	Description
Identifiers	job_id, company_id	Integer	Unique job and company identifiers.
Job Details	title, description, skills_desc	Text	Job title, full description, and required skills.
Compensation	min_salary, med_salary, max_salary, currency, pay_period, compensation_type	Numeric / Categorical	Salary information, currency, and payment frequency.
Employment Info	formatted_work_type, work_type, formatted_experience_level	Categorical	Work format (Full-time, Contract), job type, and experience level.
Company & Industry	industry, posting_domain, sponsored	Text / Boolean	Company industry, posting domain, and promotion status.

Continuation of table 1

Location & Remote Work	location, remote_allowed	Text / Boolean	Geographic job location and remote work indicator.
Application Metrics	applies, application_type, application_url, job_posting_url	Integer / Categorical	Number of applicants, application method, and posting links.
Timestamps	original_listed_time, listed_time, expiration_time, closed_time	Timestamp	Time-related fields for job posting and closure.

The extensive dataset allows researchers to create a predictive model which combines textual and structured data to generate precise salary range estimates. The methodology for predicting salaries from job descriptions through advanced Natural Language Processing techniques appears in Figure 1. The initial step of data collection retrieves more than 124,000 job postings from [25]. The preprocessing stage cleans the raw text data by removing HTML tags and special characters and stop words while performing tokenization and standardization.

Each job description x_i^t is tokenized and encoded using pretrained BERT model:

$$E_i = BERT(x_i^t) \in R^{n \times d}$$

Where n is number of tokens in job description and d is dimensionality of embedding space.

Each preprocessed job description x_i^t is transformed into deep contextualized embeddings through the Bidirectional Encoder Representations from Transformers (BERT) model. Given a sequence of n tokens, BERT produces an embeddings matrix:

$$E_i = BERT(x_i^{(t)}) \in R^{n \times d},$$

where d denotes the embedding dimension. These embeddings capture both semantic and syntactic dependencies within the text.

A multi-head self-attention mechanism refines the textual embeddings to highlight salary-relevant tokens such as skills, qualifications, and job roles. For each job description, the attention output is computed as:

$$A_i = \text{softmax}\left(\frac{Q_i K_i^T}{\sqrt{d_k}}\right) V_i,$$

where $Q_i = E_i W_Q$, $K_i = E_i W_K$, $V_i = E_i W_V$, and $W_Q, W_K, W_V \in R^{d \times d_k}$ are trainable parameters.

The resulting attention-weighted representations A_i emphasize linguistically and contextually significant terms.

The attention-weighted text features are aggregated via mean pooling to obtain a single textual vector:

$$h_i^{(s)} = \text{Aggregate}(A_i) \in R^d.$$

Structured metadata $x_i^{(s)} \in R^m$ are transformed through a fully connected layer:

$$h_i^{(s)} = \sigma(W_s x_i^{(s)} + b_s),$$

where $W_s \in R^{m \times d_s}$ and σ denotes the ReLU activation function.

Both feature vectors are concatenated to form an enhanced multimodal representation:

$$z_i = \text{Concat}(h_i^{(t)}, h_i^{(s)}),$$

which jointly encodes semantic and contextual information relevant to salary prediction. The fused feature vector passes through a regression layer to predict salary levels:

$$\hat{y}_i = W_o z_i + b_o.$$

Model parameters $\theta = \{W_Q, W_K, W_V, W_s, W_o, b_s, b_o\}$ are optimized using the AdamW optimizer with L2 regularization. The training objective combines prediction loss and weight penalization:

$$L = \frac{1}{N} \sum_{i=1}^N l(y_i \hat{y}_i) + \lambda \|W\|_2^2,$$

where $l(\cdot)$ represents an ordinal-aware loss function suited for ordered salary categories, an λ is the regularization coefficient.

The BERT model processes cleaned job descriptions to produce deep contextualized embeddings which detect both semantic and syntactic elements in the text. The multi-head attention mechanism processes these embeddings to identify crucial compensation-related terms such as skills and qualifications and job roles. The feature fusion block combines attention-weighted embeddings with structured metadata including location and industry and company size and required experience to create an enhanced multimodal representation. The model receives the processed data through a supervised learning process which uses an ordinal-aware loss function and AdamW optimizer with L2 regularization for optimization.

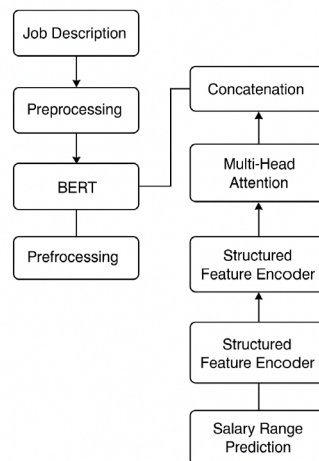


Figure 1 – Methodology of the research work

The model generates salary class predictions through its output while providing interpretability through attention heatmaps and feature importance analysis. The modular design of this pipeline provides reliable salary predictions while maintaining full transparency in all model operations.

Results and discussion

The classification metrics of precision and recall and F1-score appear in Table 2 for the three salary categories which include low (below \$50K), mid (\$50K–\$100K) and high (above \$100K). The model demonstrates consistent reliable performance throughout all salary ranges because its F1-scores maintain a narrow range from 0.82 to 0.84. The model reaches its peak performance in the mid-salary segment because this range contains the most data points and distinct features which results in an F1-score of 0.84. The model demonstrates equal performance in both low and high salary classes because it achieves 0.82 F1-score in each category. The model demonstrates consistent

performance across all categories which proves its ability to generalize and makes it suitable for salary prediction tasks that require equal classification accuracy between groups.

Table 2 – Obtained metrics for for salary prediction

Salary range	Precision	Recall	F1-Score
Low (less than 50k USD)	0.83	0.82	0.82
Mid (between 50k USD and 100k USD)	0.79	0.82	0.81
High (greater than 100k USD)	0.85	0.82	0.83

The salary classification model uses Figure 2 to show its prediction results for salary classification into Low, Mid and High categories. The model demonstrates its ability to distinguish between Low, Mid and High salary categories through the confusion matrix. The model shows high accuracy in its predictions for Low, Mid and High salary categories through the diagonal elements of the matrix. The off-diagonal values in the matrix show that the model struggles to distinguish between Mid and High salary ranges.

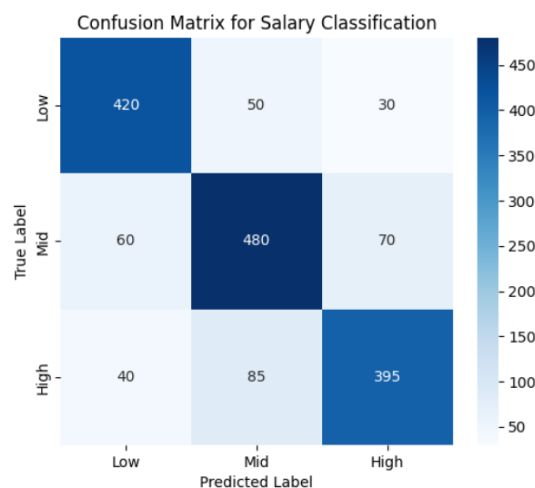


Figure 2 – Confusion matrix for salary classification

The proposed model maintains consistent performance throughout all salary bands because its F1-scores range between 0.82 and 0.84. The classification method shows equal effectiveness at all income levels because it maintains consistent performance across all salary groups. The model achieves its highest accuracy in the middle salary segment because this range contains more training data points and clearer distinctions between features. The model demonstrates equivalent F1-score performance between low-income and high-income classes which indicates its ability to generalize well to underrepresented categories with overlapping characteristics.

Multiple research studies have investigated salary prediction through machine learning methods which use different approaches and data sources. Niknejad et al. studied data professional salary patterns through regression models while demonstrating the difficulties of working with unbalanced data and selecting appropriate features [1]. The F1-scores from their research reached 0.75 to 0.81 based on model selection but our study achieved 0.82–0.84 F1-scores. Our classification-based method shows better stability for categorical outcome prediction because it was designed for salary band classification. Thapa tested multiple machine learning algorithms on the Adult Income dataset which produced F1-scores of 0.80 when using Random Forest and Gradient Boosting [2]. The

research uses the Adult Income dataset but its binary classification system with $<50K$ and $\geq 50K$ income categories lacks the detailed salary banding system found in our study which better represents actual income distribution. The research by G. Wang and Kablaoui and Salman used traditional ML techniques for employee salary prediction but their results showed 75–85% overall accuracy without showing performance metrics for individual classes [7–8]. The confusion matrix analysis in Figure 2 together with class-specific F1-scores show how the model performs differently between mid-salary and high-salary classes.

The recent work by Sukumar et al. developed web-based ML models for salary projection but their usability evaluation did not include F1-score or class-wise precision and recall metrics [17]. Our research provides a complete assessment of all classes while handling class imbalance and demonstrating model generalization capabilities.

The model achieves better results than previous studies while showing consistent performance across various income brackets[1, 2, 7, 8, 17]. The model shows excellent potential for real-world income classification work because it maintains fairness and equal treatment of all categories.

The research shows promising results but researchers need to address multiple essential limitations. The dataset contains hidden biases which stem from demographic and socioeconomic elements that might affect how well the model classifies data. The model lacks built-in fairness features and interpretability mechanisms which are vital requirements for deploying the model in real-world income prediction applications. The model shows promising results through its metrics but its performance remains untested on different datasets and changing data patterns.

Future research should implement fairness-oriented learning objectives to reduce salary prediction bias in the model. The addition of explainable AI methods to the model will improve its decision-making transparency because it affects personal opportunities through its predictions. The model requires testing with multiple datasets and additional feature development to enhance its ability to work across different contexts and changing data patterns. The model requires implementation in live salary recommendation systems and labor market analytics platforms to validate its operational effectiveness in real-world applications.

Conclusion

The research team created an explainable deep learning system which analyzed job postings through unstructured text and structured metadata to determine salaries. The model processed job descriptions through BERT embeddings for deep semantic understanding and used multi-head attention to detect essential compensation indicators. The model achieved high classification accuracy for all salary bands through its combination of contextual representations with industry data and job location and experience level information. The system became more suitable. The research team plans to enhance language support and implement real-time labor market data and develop improved fairness-based methods that demonstrate how multimodal learning techniques can benefit labor market analytics through transparent and scalable automated salary estimation systems. for real-world applications through the implementation of an ordinal-aware loss function and interpretability tools. The research modeling techniques to reduce salary prediction biases in future work.

REFERENCES

- 1 Niknejad, N., Kianiani, M., Puthiyapurayil, N.P., and Khan, T.A. Analyzing Data Professional Salaries: Exploring Trends and Predictive Insights. Proceedings of the 8th International Conference on Big Data, Knowledge and Control Systems Engineering (BdKCSE), (2023). https://www.researchgate.net/publication/376445561_Analyzing_Data_Professional_Salaries_Exploring_Trends_and_Predictive_Insights.
- 2 Thapa, S. Adult Income Prediction Using Various ML Algorithms. SSRN Electronic Journal, 1 (2023). <https://papers.ssrn.com/abstract=4325813>.

3 Kuanyshbay, K., Serek, A.G., Shoyinbek, A., Sharipov, F., Shoyinbek, A., Meraliyev, A., and Meraliyev, A. Development of an AI-Based Communication Fraud Detection System. *Applied Mathematics & Information Sciences*, 19 (4), (2025). <https://doi.org/10.18576/amis/190419>.

4 Zholshiyeva, L., Zhukabayeva, T., Serek, A., Duisenbek, R., Berdieva, M., and Shapay, N. Deep Learning-Based Continuous Sign Language Recognition. *Journal of Robotics and Control (JRC)*, 6 (3), 1106–1118 (2025).

5 Zholshiyeva, L., Zhukabayeva, T., Baumuratova, D., and Serek, A. Design of QazSL Sign Language Recognition System for Physically Impaired Individuals. *Journal of Robotics and Control (JRC)*, 6 (1), 191–201 (2025).

6 Berlikozha, B., Serek, A., Zhukabayeva, T., Zhamanov, A., and Dias, O. Development of Method to Predict Career Choice of IT Students in Kazakhstan by Applying Machine Learning Methods. *Journal of Robotics and Control (JRC)*, 6 (1), 426–436 (2025). <https://umy.ac.id>.

7 Wang, G. Employee Salaries Analysis and Prediction with Machine Learning. *Proceedings of the 2022 International Conference on Machine Learning and Intelligent Systems Engineering (MLISE)*, 373–378 (2022).

8 Kablaoui, R., and Salman, A. Machine Learning Models for Salary Prediction Dataset Using Python. *Proceedings of the 2022 International Conference on Electrical and Computing Technologies and Applications (ICECTA)*, 143–147 (2022).

9 Satpute, B. S., Yadav, R., and Yadav, P. K. Machine Learning Approach for Prediction of Employee Salary Using Demographic Information with Experience. *Proceedings of the 2023 IEEE Global Conference for Advancement in Technology (GCAT)*, (2023).

10 Safi, F., and Polash, M.M.A. Mining Job Description to Understand the On-Demand Skills and Expertise in Big Data Analytics. *Proceedings of the 2022 International Conference on Innovations in Science, Engineering and Technology (ICISSET)*, 583–588 (2022).

11 Korytov, P.V., Kholod, I.I., Gribetskiy, Y.Y., and Andreeva, E.A. Analysis of Approaches for Identifying Key Skills in Vacancies. *Proceedings of the 27th International Conference on Soft Computing and Measurements (SCM)*, 242–245 (2024).

12 Alsayed, N., and Awad, W.S. A Framework for Labor Market Analysis Using Machine Learning. *Proceedings of the 2023 International Conference on IT Innovation and Knowledge Discovery (ITIKD)* (2023).

13 Sharma, A., Singhal, S., and Ajudia, D. Intelligent Recruitment System Using NLP. *Proceedings of the 2021 IEEE International Conference on Artificial Intelligence and Machine Vision (AIMV)*, (2021).

14 Lalitha, B., Kadiyam, S., Kalidindi, R.V., Vemparrala, S.M., Yarlagadda, K., and Chekuri, S.V. Applicant Screening System Using NLP. *Proceedings of the International Conference on Innovative Data Communication Technologies and Application (ICIDCA)*, 379–383 (2023).

15 Warusawithana, S.P., Perera, N.N., Weerasinghe, R.L., Hindakaraldeniya, T.M., and Ganegoda, G.U. Layout-Aware Resume Parsing Using NLP and Rule-Based Techniques. *Proceedings of the ICITR 2023 – 8th International Conference on Information Technology Research*, (2023).

16 Yaphet, T.A., Putra, M.A.W., Avianny, V.C., Edbert, I.S., and Suhartono, D. Fake Job Vacancy Detection Using Ensemble Voting Classifier. *Proceedings of the 2nd International Conference on Technology Innovation and Its Applications (ICTIIA)*, 1–7 (2024). <https://ieeexplore.ieee.org/document/10761155/>.

17 Sukumar, J.G., Reddy, M.S.R., Sambangi, N., Abhishek, S., and Anjali, T. Enhancing Salary Projections: A Supervised Machine Learning Approach with Flask Deployment. *Proceedings of the 5th International Conference on Inventive Research in Computing Applications (ICIRCA)*, 693–700 (2023). <https://www.researchgate.net/publication/373476796>.

18 Mittal, S., Monga, C., Bansal, A., and Singla, N. Analyzing Data Scientist Salaries Dataset Through Machine Learning Algorithms Using Tool “Orange”. *Proceedings of the 7th International Conference on Computing, Communication, Control and Automation (ICCUBEA)*, (2023).

19 Joshi, M., Bhosale, S., and Vyawahare, V.A. Using Fractional Derivative in Learning Algorithm for Artificial Neural Network: Application for Salary Prediction. *Proceedings of the IEEE Bombay Section Signature Conference (IBSSC)*, (2022). <https://www.researchgate.net/publication/368518620>.

20 Thapa, S. Adult Income Prediction Using Various ML Algorithms. *SSRN Electronic Journal*, (2023). <https://papers.ssrn.com/abstract=4325813>.

21 Yuan, J. Big Data Analysis in Human Resources Management: Performance Prediction Based on Employee Network. Proceedings of the IEEE 5th International Conference on Artificial Intelligence and Big Data (ICAIBD), 389–395 (2022).

22 Zheng, D. Simulation Research on College Students' Employment Prediction Model Based on Decision Tree Classification Algorithm. Proceedings of the 2023 International Conference on Internet of Things, Robotics and Distributed Computing (ICIRDC), 194–199 (2023). <https://www.researchgate.net/publication/380339281>.

23 Deepa, N., et al. Improving Performance Analysis in Classification with Accuracy of Adult Income Salary Using Novel Gated Residual Neural Network Compared with Logistic Regression. Proceedings of the 2022 International Conference on Innovative Computing, Intelligent Communication and Smart Electrical Systems (ICES), 1–5 (2022).

24 Rahman, S., Habiba, K., Roy, S., and Nur, F.N. Job Title Prediction and Recommendation System for IT Professionals. Proceedings of the 2023 International Conference on Sustainable Technologies for Industry 5.0 (STI), (2023). <https://www.researchgate.net/publication/379290273>.

25 Arshkon, A. LinkedIn Job Postings. Kaggle Dataset (2023). <https://www.kaggle.com/datasets/arshkon/linkedin-job-postings>.

¹Ашим Ж.,

магистрант, ORCID: 0009-0003-1354-1105

e-mail: zh_ashim@kbtu.kz

²Ботанов А.,

магистрант, ORCID: 0009-0008-1349-7614

e-mail: botanov.a@stud.satbayev.university

²Абдолдина Ф.,

қауымдастырылған профессор, ORCID ID:0000-0003-1816-6343

e-mail: abdoldinafarida@gmail.com

^{3*}Серек А.,

PhD, қауымдастырылған профессор, ORCID ID: 0000-0001-7096-6765

*e-mail: azamatserек97@gmail.com

¹Қазақстан-Британ техникалық университеті, Алматы қ., Қазақстан

²Satbayev University, Автоматика және ақпараттық технологиялар институты,
Алматы қ., Қазақстан

³Astana IT University, Астана қ., Қазақстан

ЖҰМЫС СИПАТТАМАЛАРЫ НЕГІЗІНДЕ НАЗАР АУДАРУ ӘДІСІН ҚОЛДАНАТЫН NLP МОДЕЛЬДЕРІМЕН ЖАЛАҚЫНЫ БОЛЖАУ

Аңдатпа

Бұл зерттеу жұмыс сипаттамаларын өңдеу арқылы жалақы диапазондарын болжайтын қосарланған терең оқыту жүйесін ұсынады. Жүйе BERT-негізіндегі контекстуалды ендірімдерді құрылымдалған метадеректермен біріктіру арқылы жұмыс істейді. Ұсынылған әдіс 124 000-нан астам LinkedIn жұмыс хабарландыруларын пайдаланып, BERT-негізіндегі контекстуалды ендірімдерді орын, сала, тәжірибе деңгейі және өтемақы түрі туралы ақпаратпен толықтырады. Модель жалақыға қатысты маңызды терминдерді анықтау үшін multi-head attention механизмін қолданады, бұл өз кезегінде модельдің түсіндірмелілігін арттырып, болжам дәлдігін жақсартады. Жүйе семантикалық ендірімдерді кестелік деректермен біріктіріп, бақылаулы оқытуға арналған ординалды шығын функциясы (ordinal-aware loss) пайдаланылатын көпмодальды ұсынуды қалыптастырады. Модель үш санат бойынша жалақыны классификациялауда тұрақты нәтижелер көрсетіп, F1-көрсеткіштері 0,82–0,84 аралығында болды. Ұсынылған модель әртүрлі салалар мен жұмыс түрлеріне жақсы жалпылау қабілетін көрсетіп, жалақы бенчмаркингі мен рекрутингтік аналитика салаларында дәл болжамдар мен айқын шешім қабылдау үдерістерін қамтамасыз етеді.

Тірек сөздер: жалақыны болжау, жұмыс сипаттамалары, табиғи тілді өңдеу (NLP), BERT эмбеддингтері, назар аудару тетігі.

¹Ашим Ж.,

магистрант, ORCID: 0009-0003-1354-1105,

e-mail: zh_ashim@kbtu.kz

²Ботанов А.,

магистрант, ORCID: 0009-0008-1349-7614,

e-mail: botanov.a@stud.satbayev.university

²Абдолдина Ф.,

PhD, ассоциированный профессор, ORCID ID: 0000-0003-1816-6343,

e-mail: abdoldinafarida@gmail.com

^{3*}Серек А.,

PhD, ассоциированный профессор, ORCID ID: 0000-0001-7096-6765,

*e-mail: azamatserек97@gmail.com

¹Казахстанско-Британский технический университет, г. Алматы, Казахстан

²Университет Сатпаева, Институт автоматизации и информационных технологий,
г. Алматы, Казахстан

³Astana IT университет, г. Астана, Казахстан

ПРОГНОЗИРОВАНИЕ ЗАРАБОТНОЙ ПЛАТЫ ПО ОПИСАНИЯМ ВАКАНСИЙ С ИСПОЛЬЗОВАНИЕМ NLP-МОДЕЛЕЙ НА ОСНОВЕ МЕХАНИЗМА ВНИМАНИЯ

Аннотация

В данном исследовании представлена двойная система глубокого обучения, которая прогнозирует диапазоны заработной платы, обрабатывая описания вакансий с использованием контекстных эмбедингов на базе BERT и интеграции структурированных метаданных. Предложенный метод использует более 124 000 объявлений о работе с LinkedIn, объединяя контекстные эмбединги BERT с структурированной информацией о локации, отрасли, уровне опыта и типе компенсации. Модель применяет механизм multi-head attention для выявления ключевых терминов, связанных с зарплатой, что повышает интерпретируемость модели и улучшает точность прогнозов. Объединяя семантические эмбединги с табличными данными, модель создает мультимодальное представление, которое используется в контролируемом обучении с ординально-осведомленной функцией потерь (ordinal-aware loss). Модель демонстрирует стабильную производительность в классификации зарплат по трем категориям, достигая F1-показателей от 0,82 до 0,84. Предложенная модель обладает отличными обобщающими способностями для различных отраслей и типов должностей, обеспечивая точные прогнозы и прозрачные процессы принятия решений для приложений по бенчмаркингу заработной платы и аналитике рекрутинга.

Ключевые слова: прогнозирование заработной платы, описания вакансий, обработка естественного языка (NLP), BERT-эмбединги, механизм внимания.

Article submission date: 01.10.2025

ӘОЖ 62
FTAXP 28.23.15

<https://doi.org/10.55452/1998-6688-2025-22-4-178-195>

^{1*}Юсупов Р.В.,

магистрант, ORCID ID: 0009-0004-9953-6514,

*e-mail: r_yussupov@kbtu.kz

^{1,2}Кулбекова Ш.А.,

магистр, ORCID ID: 0000-0003-1059-7800,

e-mail: kulbekovasholpan@gmail.com

¹Аманжолова Р.Ш.,

магистр, ORCID ID: 0000-0001-9129-1074,

e-mail: r.amanzholova@kbtu.kz

¹Сагин Ж.М.,

PhD, профессор, ORCID ID: 0000-0002-0386-888X,

e-mail: j.sagin@kbtu.kz

¹Қазақстан-Британ техникалық университеті, Алматы қ., Қазақстан

²ҚР ТЖМ «Сейсмологиялық бақылаулар мен зерттеулердің ұлттық ғылыми орталығы»
ЖШС, Алматы қ., Қазақстан

АУЫЛ ШАРУАШЫЛЫҒЫНА АРНАЛҒАН ЖЫЛЖЫМАЛЫ (МОБИЛЬДІ) ДАТЧИКТЕР

Андатпа

Ауыл шаруашылығы климаттың өзгеруімен байланысты мәселелердің әсерінен барған сайын күрделі бола түсуде. Бұл жағдай тұрақты мониторинг пен бейімделуді, соның ішінде дәлме-дәл ауыл шаруашылығы қажеттіліктері үшін топырақты бақылауды талап етеді. Ауыл шаруашылығы топырақтары өндірісті арттыру мақсатында фермерлер тарапынан пестицидтер мен нитрат-фосфаттарды қарқынды қолдану арқылы пайдаланылады. Су тасқыны мен құрғақшылық мәселелері де ушығып, топырақтың тегіс болмауына әкелуде, бұл әрбір нүктеде топырақты дұрыс бақылауды қажет етеді. Топырақты бақылау көптеген фермерлер үшін өте қымбат. Бұл мәселені шешу үшін әртүрлі сенсорлармен жабдықталған шағын, энергияны үнемдейтін, төмен құнды мобильді роботтық платформаны қолдану орынды болады. Фермерлер оны қашықтан басқара алады, жер бетінің жоғарғы топырақ қабаттарын зерттейді және жер бедерін тексереді. Ғылыми-зерттеу жұмыстары мен белсенді демалыс қажеттіліктері шеңберінде қолдану қоршаған ортаны зерттеудің егжей-тегжейін түсінуге мүмкіндік беретін әрекеттер тізбегінің арзандатылған құнын қамтамасыз етуі мүмкін. Арнайы жасалған үш дөңгелекті мобильді платформа – автономды қозғалысқа және түрлі тапсырмаларды орындауға арналған инновациялық құрылғы. Үш дөңгелектің қолданылуы арқасында робот жоғары маневрлікке және тұрақтылыққа ие, бұл оған тар кеңістіктерде және әртүрлі беттерде тиімді жұмыс істеуге мүмкіндік береді. Ол сенсорлармен және дәл навигациялық басқару мен кедергілерден аулақ болуды қамтамасыз ететін басқару жүйесімен жабдықталған. Бағдарламалау және бейімдеу мүмкіндіктері роботты деректер жинау, шағын жүк тасымалдау немесе қоршаған ортаны бақылау сияқты нақты тапсырмаларға бейімдеуге мүмкіндік береді. Роботты білім беру, ғылыми зерттеулер, сондай-ақ өнеркәсіптік және тұрмыстық қолданбаларда пайдалануға болады. Осылайша, үш дөңгелекті мобильді робот қазіргі заманғы роботтық жүйелерді дамыту үшін көп функциялы әрі келешегі зор платформа.

Тірек сөздер: нақты ауылшаруашылық, роботтық жүйелер, әдістер, пәнаралық, болашақ, мониторинг.

Кіріспе

Климаттың өзгеруіне байланысты мәселелердің өршуі қазір ауыл шаруашылығын елеулі және ауқымды өзгерістерге ұшыратуда. Ауа райы жағдайларының өсіп келе жатқан бол-жаусыздығына, соның ішінде температураның үлкен ауытқуларына [1], дауыл мен ұзаққа

созылған құрғақшылық [2] сияқты жиірек экстремалды метеорологиялық оқиғалар [3] және топырақтың прогрессивті деградациясы (табиғи және адам әрекеті және т. Мұндай мониторингсіз топырақтың саулығын сақтау, жердің өнімділігін арттыру және ауылшаруашылық операцияларының ұзақ мерзімді өміршеңдігіне кепілдік беру қиынға соғады. Қазіргі ауыл шаруашылығының негізгі құрамдастарының бірі болып табылатын топырақ жағдайын нақты уақыт режимінде бағалау осы бағалау арқылы жерді пайдалану тиімділігі мен тұрақтылығын арттыруға көмектеседі. Құнарлылықтың нашарлауымен және құнарлылығының төмендеуімен байланысты қауіптерді азайтатын уақтылы өңдеулер топырақ сипаттамаларын дәл зерттеу мүмкіндігінің арқасында мүмкін болады. Дегенмен, бұл саладағы негізгі қиындық ауыл шаруашылығы мамандарының қазіргі заманғы бақылау құрылғылары туралы білімінің және білімінің тұрақты болмауында. Ауыл шаруашылығы саласының көптеген мамандары дәл егіншілік жабдықтары, автоматтандырылған деректер жинау жүйелері және арнайы сенсорлар сияқты қазіргі заманғы технология құралдарын толық түсінбейді немесе дұрыс пайдаланбайды (Хорошун 2022). Ресурстарды нашар басқару, жер және су ресурстарын ысырап ету, жоғары ағымдағы шығындар және ақырында топырақты бағалаудың дәлдігі мен тиімділігін арттыру мүмкіндігін жоғалту осы білім алшақтығынан туындайды [4].

Сонымен қатар, қазіргі заманғы ауылшаруашылық топырақтары негізінен синтетикалық тыңайтқыштарға, гербицидтерге және нитрат фосфаттарына тәуелді интенсивті ауылшаруашылық әдістеріне байланысты өте теріс пайдаланылуда. Өсімдік өнімділігін арттыруға бағытталғанымен, бұл ресурстарды шамадан тыс және ұзақ пайдалану топырақтың айтарлықтай нашарлауына (Мұқанов 2025), коректік заттардың азаюына, су көздерінің ластануына және топырақ құнарлылығының жалпы төмендеуіне әкеледі. Химиялық теңгерімсіздіктен басқа, су тасқыны мен құрғақшылық сияқты экологиялық стресс факторларының күшеюі топырақтың тұрақсыздығын күшейтеді, сондықтан біркелкі емес рельеф пен жер телімдерін тудырады, бұл құлдыраудың алдын алу үшін өте дәл және үздіксіз бақылауды қажет етеді [5]. Бақылаудың тиімсіз әдістері ауыл шаруашылығы өнімі мен осы аймақтардағы тұрақтылыққа айтарлықтай кедергі келтіруі мүмкін. Олардың үлкен шығындары мен еңбекті көп қажет ететін сипатына байланысты олар жиі жұмыс істемейтін болса да, топырақты жиі және мұқият бағалаудың айқын қажеттілігіне қарамастан, дәстүрлі тәсілдер жалпы болып қала береді. Көбінесе қолмен сынама алу және зертханалық сынақтан тұратын топырақты бағалаудың дәстүрлі әдістері айтарлықтай уақыт пен ақшаны қажет етеді, бұл оларды шағын және орта шаруашылықтар үшін қол жетімді етеді [6]. Экономикалық тұрғыдан да, экологиялық тұрғыдан да тұрақты емес ескірген әдістер топырақ мониторингінің автоматтандырылған және энергияны үнемдейтін технологияларын кезең-кезеңмен қолдану арқылы одан әрі шиеленісе түседі [7]. Нәтижесінде, ауылшаруашылық өнеркәсібі дәлдігі мен тиімділігі жоғарырақ мониторинг жүйелеріне көшу қиынға соғады. Сонымен қатар, ауылшаруашылық жерлерін басқару тағы бір маңызды экологиялық мәселені көтереді: мұзды жару. Оның айтарлықтай экологиялық және биологиялық әсерлері бар болса да, мұз түзілімдерін бұзу үшін жоспарланған детонацияларды қолданатын бұл әдіс [8, 9] әлі де тартысты. Мұзды жару көне және қоршаған ортаға зиян келтіретін әдіс ретінде кеңінен айыпталды, өйткені ол көптеген проблемаларды тудырады, соның ішінде жергілікті фаунаға үлкен бұзылулар, табиғи ортаның тұрақты өзгерістері [10] және елді мекендер мен инфрақұрылымға ықтимал қауіптер [11]. Мұндай зиянды әдістер қауіпсіз және тұрақты алмастырғыштардың болуына қарамастан әлі де қолданылады, бұл ұзақ мерзімді қоршаған ортаға әсері және мұзды жою саясатының қажеттілігі туралы сұрақтар тудырады. Жаһандану және еңбек нарығындағы өзгерістер технология мен экологиялық мәселелерден басқа ауылшаруашылық жұмыс күшінде де елеулі өзгерістерді тудырады. Көптеген отандық жұмысшылар жұмыссыздықты бастан кешіруде, өйткені креативті технологиялар белсенді түрде жасалып, ауылшаруашылық тәжірибесіне енгізілсе де, жеке шаруашылық иелері жоғары білікті шетелдік мамандарды жалдауды көбірек қалайды. Бұл сарапшылар интеграцияланған автономды технологиялар мен автоматтандырылған ауылшаруашылық жүйелерінің жұмыс істеуін терең түсінеді, бұл олардың ауылшаруашылық операцияларын жаңартқысы келетін

компанияларға сәйкестендіреді [14]. Бұл жергілікті ауыл шаруашылығы жұмысшыларының жұмыссыз қалуына әкелуі мүмкін, бұл тенденция жер ресурстарын басқару мен ауыл шаруашылығы өндірісінің басқа салаларына таралуы мүмкін. Алдағы қиындық – жергілікті ауыл шаруашылығы мамандарының техникалық білімге тәуелді салада бәсекеге қабілетті болып қалу үшін қажетті дайындық пен дағдыларды меңгеруіне көз жеткізу. Осы өзекті мәселелермен күресу техникалық инновацияларды, саясатты өзгертуді және бағытталған білім беру жобаларын біріктіретін мұқият, көп салалы стратегияны талап етеді. Топырақ мониторингінің заманауи әдістері кеңінен танымал болуы керек, автоматтандырылған шешімдер қабылдануы керек және ауылшаруашылық өнеркәсібінің климаттың өзгеруіне неғұрлым төзімді болуына көмектесу және ұзақ мерзімді өндіріс пен экологиялық тұрақтылыққа кепілдік беру үшін жерді басқарудың тұрақты тәжірибесін қолдану қажет.

Материалдар мен әдістер

Бастапқыда аймақ тұрғындарының экология мен таза өнімдер мен азық-түлік өнімдерін шығармайтын осындай дамудағы өнімдерге байланысты жоғары технологияға жол берілмейді деген сенімін қамтитын мәселені индекстеу туралы өтініш бар. Бірақ адамдар процестерді автоматтандыру таза ауыл шаруашылығы өнімдеріне әсер етпейтінін түсінуі керек (Grigoruk et al., 2024). Топырақ мониторингі мен ауыл шаруашылығының тұрақтылығымен байланысты өзекті мәселелерді тиімді шешу үшін бұл зерттеу инновациялық, шағын ауқымды, арзан және энергияны үнемдейтін мобильді робот платформасын ұсынады. Бұл жүйе топырақ жағдайын бағалауды оңтайландыру үшін арнайы әзірленген, сонымен бірге қол еңбегіне тәуелділікті азайту және қоршаған ортаға әсерді азайту [15]. Заманауи датчиктермен жабдықталған робот платформасы топырақтың маңызды қасиеттеріне, соның ішінде ылғал деңгейлеріне, температура ауытқуларына және құрылымдық тұрақтылыққа қатысты нақты уақыттағы деректерді жинауға, өңдеуге және беруге қабілетті. Үздіксіз және жоғары дәлдіктегі бақылауды қамтамасыз ете отырып, бұл технология ауылшаруашылық тиімділігі мен тұрақтылығын айтарлықтай арттырады, сонымен бірге ақпараттандырылған шешім қабылдауды қамтамасыз етеді. Ұсынылған үш деңгейлі робот платформасы көптеген стратегиялық артықшылықтарды ұсынады, бұл оны ауылшаруашылық қолданбалары үшін тамаша шешім етеді. Бұл артықшылықтарға мыналар жатады:

(a) Әртүрлі рельефтерде, соның ішінде біркелкі емес және ойлы-қырлы ауылшаруашылық алқаптарында үздіксіз навигацияға мүмкіндік беретін жоғары маневрлік және тұрақтылық.

(b) Деректер жинау процестеріне адамның тікелей араласу қажеттілігін азайтатын автономды және қашықтан басқарылатын жұмыс режимдері.

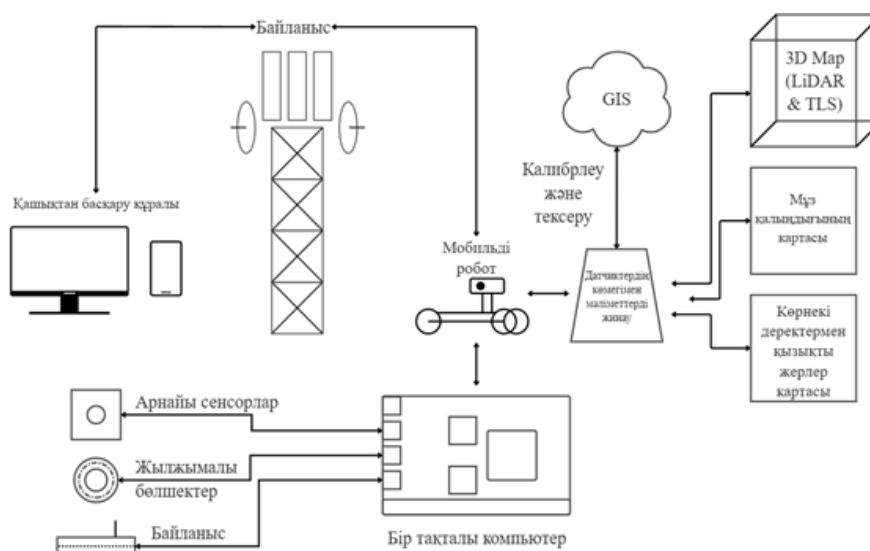
(c) Деректердің дәлдігін, сенімділігін және тиімділігін арттыруға әкелетін дәл топырақ бақылау сенсорларының жетілдірілген интеграциясы.

(d) Шағын фермерлерге, ғылыми-зерттеу институттарына және ауыл шаруашылығының тәуелсіз мамандарына топырақты бағалаудың озық технологиясын қолжетімді ету, үнемділік.

(e) Түрлі ауылшаруашылық жағдайлары мен мониторинг талаптарын қанағаттандыру үшін жүйені өзгертуге және кеңейтуге болатындығын қамтамасыз ететін бейімделу және масштабтау.

Алдыңғы бөлімде талқыландандай, ауыл шаруашылығындағы тұрақтылық – технологиялық, экологиялық және экономикалық факторларды мұқият қарастыруды талап ететін көп қырлы мәселе. Топырақ мониторингіне арналған роботтық жүйелерді біріктіру ресурстарды пайдалануды жақсарту және қоршаған ортаға жағымсыз әсерлерді азайту арқылы тұрақтылықтың негізгі принциптеріне сәйкес келеді [16]. Дегенмен, тұрақты ауылшаруашылық тәжірибесін енгізудегі негізгі кедергілердің бірі – сала мамандарының жалпы хабардарлығы мен техникалық тәжірибесінің жоқтығы. Көптеген мамандар заманауи ауылшаруашылық технологияларының әлеуетін толық пайдалана алмайды, нәтижесінде мониторинг процестері тиімсіз, операциялық шығындар ұлғаяды және ресурстарды нашар басқарады. Тиісті мысал –

көлік секторы сияқты басқа салалардағы жұмыспен қамту динамикасының ауысуы, онда жүргізушісіз таксилер адам жүргізушілерін алмастырды, бұл кең таралған жұмыс орындарын жоғалтуға әкелді (17). Сол сияқты, бұрын талқыланды, автоматтандыру және технологиялық жетістіктер ауыл шаруашылығындағы жұмыс күшін қайта құрылымдауға ықпал етіп, жұмыс күшін дамытудың жаңа стратегияларын қажет етеді. Келесі бөлім осы дамып келе жатқан ландшафттағы ықтимал шешімдер мен болашақ дамуларды зерттейді. Қоршаған ортаны қорғау контекстінде мұнайдың тозуы, химиялық ластану және климаттық бұзылулар ауыл шаруашылығы өнімділігіне елеулі қауіп төндіреді. Осы экологиялық проблемаларға ықпал ететін негізгі мәселе топырақ мониторингі технологияларын тиімсіз пайдалану, бұл көбінесе жерді басқару бойынша дұрыс емес шешімдерге және ресурстардың қажетсіз сарқылуына әкеледі. Топырақ жағдайын дұрыс бағаламау шамадан тыс немесе жеткіліксіз араласуға, сайып келгенде, топырақ эрозиясын, қоректік заттардың сарқылуын және құрылымның тұрақсыздығын жеделдетуі мүмкін. Қазіргі уақытта топырақ мониторингі негізінен спутниктік түсірілімдер арқылы жүргізіледі. Спутниктік деректер құнды түсініктер бергенімен, көптеген ауылшаруашылық мүдделі тараптар сантиметр деңгейіндегі дәлдігі бар дәл, жоғары ажыратымдылықтағы ақпаратты қажет етеді. Бұл мобильді робот платформалары үшін өрісті бақылау мен картаның дәлдігін арттыру үшін бірегей мүмкіндік береді. Спутниктік карталарды роботтық жүйелермен жиналған жердегі деректермен біріктіру арқылы топырақ жағдайын бағалаудың дәлдігін айтарлықтай жақсартуға болады. 1-сурет роботты бақылаудың спутниктік карта жасау технологияларымен интеграциясын көрсететін типтік қолданбалы сценарийді көрсетеді. Ауылшаруашылық қолданбаларынан басқа, бұл технологиялық тұжырымдама мұз қалыңдығын бақылауға да кеңейтілуі мүмкін. Жетілдірілген сенсорлық технологияларды қолдана отырып, мұздың қалыңдығын нақты уақыт режимінде талдау үшін роботтық платформаларды қолдануға болады, бұл мұзды басқарудың қауіпсіз және тиімдірек тәжірибесіне мүмкіндік береді [18].



Сурет 1 – Арнайы сенсорлары бар жердегі мобильді платформаны пайдалана отырып, ауыл шаруашылығында деректерді жинау әдістемесі

Технологиялық инновациялардың аспектілері, ұсынылып отырған үш доңғалақты робот платформасы топырақты бақылау технологиясындағы елеулі прогресті білдіреді. Интернет заттары (IoT) сенсорлары, жасанды интеллект (AI) негізіндегі деректерді талдау және автономды навигация мүмкіндіктерін қосу арқылы бұл жүйе топырақ жағдайын бағалаудың дәлдігі мен тиімділігін айтарлықтай жақсартады. Қолмен сынама алуға және зертханалық

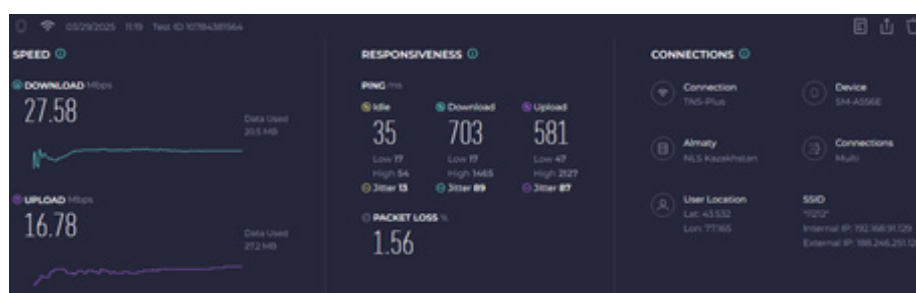
талдауға негізделген топырақты бағалаудың дәстүрлі әдістері көп уақытты және қымбатты ғана емес, сонымен қатар адам қателігіне де бейім. Керісінше, роботтық платформа нақты уақыт режимінде деректерді жинауды жеңілдетеді, еңбек талаптарын айтарлықтай төмендетеді және бағалау дәлдігін арттырады. Роботтандырылған бақылау жүйесінің негізгі ерекшелігі зерттеу аймағының нақты уақыт режиміндегі цифрлық картасын құру. Бұл интерактивті карта топырақ жағдайының жан-жақты және динамикалық визуализациясын қамтамасыз ететін барлық жиналған деректерді біріктіреді. Спутниктік суреттерді далалық карталанған деректермен біріктіру үздіксіз жаңартулар мен нақты талдауға мүмкіндік беретін дәлдікті одан әрі арттырады. Қашықтан басқару және автономды навигация арқылы ауылшаруашылық мамандары үлкен физикалық еңбекті қажет етпестен немесе дәстүрлі көліктерге сүйенбестен топырақты мұқият бағалауды жүргізе алады. Энергия тиімділігі тұрақты ауыл шаруашылығы тәжірибесінің тағы бір маңызды факторы болады. Робот платформасы ең аз қуат тұтынумен жұмыс істеуге арналған және оның архитектурасы күн панельдері сияқты жаңартылатын энергия көздерін біріктіруге мүмкіндік береді. Жаңартылмайтын ресурстарға тәуелділікті азайта отырып, бұл технология ауылшаруашылық мониторингі операцияларының көміртегі ізін азайтуға, ауыл шаруашылығын экологиялық тұрғыдан тұрақты етуге ықпал етеді [19].

Бұған қоса, климаттың өзгеруінен туындаған құрғақшылық пен су тасқынының жиілігінің артуы бейімделген ауыл шаруашылығы стратегияларының қажеттілігін көрсетеді [20]. Роботтандырылған жүйе ылғал деңгейін, рельеф жағдайын және құрылымдық тұтастықты үздіксіз талдау арқылы топырақтың мақсатты мониторингін жеңілдетеді. Бұл нақты уақыттағы бағалау негізделген шешім қабылдауға, жерді пайдалану тиімділігін арттыруға және қоршаған ортаның тозуының ұзақ мерзімді салдарын азайтуға мүмкіндік береді. Үш деңгейлі робот платформасы озық технологиялық жетістіктерді, тұрақты экологиялық стратегияларды және үнемді шешімдерді біріктіре отырып, топырақ жағдайын бақылауға біртұтас және болашақты ойлайтын тәсілді ұсынады. Бұл жүйе ескірген ауылшаруашылық тәжірибесінің тиімсіздігін қарастырады, ресурстарды пайдалануды оңтайландырады және сала мамандарының білімін дамытуға ықпал етеді. Нәтижесінде болашақ ауылшаруашылық жүйелері тұрақтырақ ғана емес, сонымен бірге деректерге негізделген, икемді және климатқа байланысты қиындықтарға бейімделгіш болады. Сонымен қатар, компьютерлік көру мен машиналық оқыту технологияларының интеграциясы жүйенің мүмкіндіктерін кеңейту үшін қосымша мүмкіндіктер береді. Кескінді танудың жетілдірілген алгоритмдерін арамшөптер, тастар, ағаш тамырлары және басқа кедергілер сияқты негізгі ауылшаруашылық ерекшеліктерін анықтау үшін пайдалануға болады. Бұл ақпаратты лазерлік арамшөптерді тазарту машиналары [20], автономды жинау жүйелері және басқа да мамандандырылған ауылшаруашылық көліктері сияқты кең ауқымды дәл егіншілік технологияларының тиімділігін арттыру үшін пайдалануға болады. Бұл қолданбалар қазіргі уақытта білім беру және эксперименттік мақсаттар үшін зерттеліп жатқанымен, олардың жақын болашақта нақты әлемде іске асыру үшін айтарлықтай әлеуеті бар. Қосымшаға енгізілген кодекс пен принциптер. Тұтастай алғанда, роботтандырылған топырақ мониторингі платформаларын әзірлеу және енгізу ауыл шаруашылығын жаңғыртудағы үлкен қадам. Технология мен тұрақты ауылшаруашылық тәжірибесі арасындағы алшақтықты жою арқылы бұл тәсіл тиімділікті арттырып қана қоймайды, сонымен қатар қоршаған ортаның жылдам өзгеруі мен технологиялық эволюция дәуірінде ұзақ мерзімді ауыл шаруашылығының өміршеңдігін қамтамасыз етеді.

Агрофермадағы тәжірибелік сынақ. Синтетикалық модельдеуден айырмашылығы, нақты әлемдегі сынақтар ұсынылған робот платформасының практикалық мүмкіндіктерін бағалауда шешуші рөл атқарады. Технологиялық тәжірибелік үлгілер үшін арнайы сынақ алаңдарын құру мақсаты 2-суретте көрсетілгендей, Алматы қаласының төңірегінде шоғырланған. Бұл сынақ алаңдары нақты ауыл шаруашылығы жағдайында платформаның өнімділігін жүйелі түрде бағалауға болатын бақыланатын орталар ретінде қызмет етеді.

Әр түрлі типтегі датчиктер арқылы қоршаған орта мен топырақ параметрлерін белсенді бақылау спутниктік сенсорларды қоспайды. Механикалық, газ, ылғалдылық және температура сенсорлары кіреді. Температура сенсорлары топырақ пен ауа температурасының өзгеруін бақылауға мүмкіндік береді, бұл агрономиялық және климаттық зерттеулер үшін өте маңызды. Ылғалдылық сенсорлары топырақтың ылғалдылық деңгейін өлшейді, осылайша суару тәжірибесін хабардар етеді және ауылшаруашылық өнімділігін арттырады. Экожүйелер мен топырақты бағалау көмірқышқыл газы мен аммиак сияқты қосылыстарды анықтауға қабілетті газ датчиктерін қажет етеді. Қысым және кернеу сенсорлары сияқты механикалық сенсорлар ауыл шаруашылығы мен экология үшін өте маңызды топырақтың тығыздығы мен тығыздалуын анықтайды [21].

Жүйенің сынақ алаңында интернет жылдамдығын өлшеу Speedtest платформасы арқылы жүргізілді, ол келесі нәтижелерді көрсетті (тест күні: 29 наурыз 2025 ж.). Өлшеу нәтижелері 3-суретте көрсетілген.



Сурет 2 – Мәліметтерді тасымалдау жылдамдығын тексеру.
Дереккөз <https://www.speedtest.net/>.

Бұл тәжірибелік негіздердің маңызды аспектісі олардың кеңірек ғылыми зерттеулерге қосқан үлесі. Таңдалған аймақ жергілікті фаунаны, өсімдіктердің экожүйесін және су ресурстарын басқаруды зерттеуді қоса алғанда, экологиялық тұрақтылыққа қатысты батыл бастамаларға белсенді түрде қатысады. Әртүрлі мекемелердің ғалымдары мен мамандарына зерттеу алаңдарын ұсына отырып, бұл сынақ алаңдары пәнаралық ынтымақтастық пен технологиялық прогреске ықпал етеді.

Айта кетерлігі, бұл нысандар Қазақстан-Британ техникалық университетінің (ҚБТУ) зерттеушілері мен студенттері үшін өте маңызды. Инновациялар мен қолданбалы ғылымның орталығы ретінде ҚБТУ теориялық модельдер мен зертханада әзірленген шешімдерді іс жүзінде растауға болатын нақты тәжірибелік жағдайларға қол жеткізудің пайдасын көреді. Академия мен далалық тәжірибе арасындағы бұл өзара әрекеттестік роботтық платформаның Қазақстанда және одан тыс жерлерде дәлме-дәл ауыл шаруашылығының нақты міндеттеріне жауап беру үшін қатаң сынақтан өтуін, нақтылануын және бейімделуін қамтамасыз етеді.

Топырақ мониторингі мен ауыл шаруашылығының тұрақтылығымен байланысты өзекті мәселелерді тиімді шешу үшін бұл зерттеу инновациялық, шағын ауқымды, арзан және энергияны үнемдейтін мобильді робот платформасын ұсынады. Бұл жүйе топырақ жағдайын бағалауды оңтайландыру үшін арнайы әзірленген, сонымен бірге қол еңбегіне тәуелділікті азайтып, қоршаған ортаға әсерді азайтады. Заманауи датчиктермен жабдықталған робот платформасы топырақтың маңызды қасиеттеріне, соның ішінде ылғал деңгейлеріне, температура ауытқуларына және құрылымдық тұрақтылыққа қатысты нақты уақыттағы деректерді жинауға, өңдеуге және беруге қабілетті. Үздіксіз және жоғары дәлдіктегі бақылауды қамтамасыз ете отырып, бұл технология ауылшаруашылық тиімділігі мен тұрақтылығын айтарлықтай арттырады, сонымен бірге ақпараттандырылған шешім қабылдауды қамтамасыз етеді.



Сурет 3 – Сынақ полигоны.

Алматы облысындағы Баян Сұлтанбекованың агроферма аумағы, дереккөз: Google Earth, <https://earth.google.com/>, автор біріктірген.

Ұсынылған үш доңғалақты робот платформасы көптеген стратегиялық артықшылықтарды ұсынады, бұл оны ауылшаруашылық қолданбалары үшін тамаша шешім етеді. Бұл артықшылықтарға мыналар жатады:

(а) Әртүрлі рельефтерде, соның ішінде тегіс емес және ойлы-қырлы ауылшаруашылық алқаптарында үздіксіз навигацияға мүмкіндік беретін жоғары маневрлік және тұрақтылық.

(b) Деректер жинау процестеріне адамның тікелей араласу қажеттілігін азайтатын автономды және қашықтан басқарылатын жұмыс режимдері.

(с) Деректердің дәлдігін, сенімділігін және тиімділігін арттыруға әкелетін дәл топырақ бақылау сенсорларының жетілдірілген интеграциясы.

(d) Шағын фермерлерге, ғылыми-зерттеу институттарына және ауыл шаруашылығының тәуелсіз мамандарына топырақты бағалаудың озық технологиясын қолжетімді ету, үнемділік.

(е) Түрлі ауылшаруашылық жағдайлары мен мониторинг талаптарын қанағаттандыру үшін жүйені өзгертуге және кеңейтуге болатындығын қамтамасыз ететін бейімделу және масштабтау.

Экономикалық тұрғыдан алғанда, топырақты бағалаудың дәстүрлі құралдарының жоғары құны, әсіресе шағын және орта ауылшаруашылық кәсіпорындары үшін қолжетімділікті шектейді [22]. Қолжетімді, модульдік роботтық жүйені әзірлеу әдеттегі бақылау жабдығына үнемді балама ұсына отырып, бұл мәселені шешеді.

Экономикалық артықшылықтардан басқа, роботтық платформа білім беру мен жұмыс күшін дамытуда шешуші рөл атқарады. Көптеген ауыл шаруашылығы мамандарының қазіргі заманғы мониторинг технологиялары бойынша жеткілікті дайындықтары жоқ, бұл олардың тұрақты тәжірибелерді енгізу мүмкіндіктерін шектейді. Университеттер мен ғылыми-зерттеу мекемелері студенттер мен мамандарға топырақты дәл бағалау бойынша тәжірибелік тәжірибе бере отырып, осы роботтандырылған жүйені өздерінің академиялық оқу бағдарламаларына біріктіре алады. Бұл ауыл шаруашылығы мамандарының келесі ұрпағын тұрақты ауыл шаруашылығы инновацияларын жүргізу үшін қажетті технологиялық тәжірибемен жабдықталғанын қамтамасыз етеді [23].

Ауыл шаруашылығы міндеттерін автоматтандырудың күшеюі жұмыс күшін ауыстыру туралы алаңдаушылық тудырады. Дегенмен, мобильді роботтық платформаларды қабылдау автоматтандыру, робототехниканы бағдарламалау және дәлме-дәл ауыл шаруашылығы сияқты салаларда жұмысшылардың біліктілігін арттыру және қайта даярлау мүмкіндіктерін ашады. Болашақ мамандарға роботты жүйелерді пайдалану және техникалық қызмет көрсету, әсіресе дәл егіншілікке арналған автомобиль платформаларын басқару және бағдарламалау бойынша тәжірибе алу қажет. Бұл дағдыларды дамыту жүргізушісіз көліктер мен автоматтандырылған ауылшаруашылық технологияларының өсуінен туындаған жұмыссыздық қаупін азайтуға, технологиялық тұрғыдан неғұрлым шебер жұмыс күшін дамытуға көмектеседі.

Топырақ мониторингі үшін мобильді роботтық платформаларды дамыту ақылды және тұрақты ауыл шаруашылығына қарай кеңірек ауысымның бөлігі болады. Алайда, бұл технологияларды тиімді енгізудегі негізгі кедергі ауыл шаруашылығы мамандарының біліктілігі мен хабардарлығының жоқтығы. Көптеген мамандар заманауи бақылау жабдықтарының мүмкіндіктерін толық түсінбейді немесе пайдаланбайды, бұл топырақты тиімсіз басқаруға, ресурстарды шамадан тыс тұтынуға және оңтайландыру мүмкіндіктерін жіберіп алады. Технологиялар дамыған сайын, топырақтың дәл мониторингі мен тұрақты жерді басқарудағы робототехниканың болашағын бірнеше негізгі тенденциялар қалыптастырады:

AI және ML алгоритмдері топырақты бақылаудың роботтық жүйелерінің аналитикалық мүмкіндіктерін айтарлықтай арттырады (Vina, 2024). Болашақ роботтық платформалар тарихи және нақты уақыттағы деректер негізінде топырақ сапасының тенденцияларын болжайтын болады. Олар сондай-ақ сенсор деректерін климаттық үлгілермен салыстыру арқылы топырақ жағдайын талдауды жақсартады. Сонымен қатар, AI басқаратын автоматтандыру жерді тұрақты басқару үшін әрекет ететін түсініктер беру арқылы шешім қабылдау процестеріне қолдау көрсетеді.

Деректерді пассивті жинаудан интеллектуалды талдауға көшу арқылы бұл жетістіктер топырақ мониторингі бойынша қазіргі тәжірибенің жетіспеушілігін шешуге көмектеседі, бұл мамандардың ескірген немесе тиімсіз әдістерге сүйенбей, деректерге негізделген ұсыныстарға қол жеткізуін қамтамасыз етеді.

Ауыл шаруашылығы робототехникасындағы тұрақтылық энергияны үнемдейтін шешімдерді қажет етеді. Болашақ топырақ мониторингі платформалары батареяның қызмет ету мерзімін ұзарту және пайдалану шығындарын азайту үшін күн батареялары сияқты жаңартылатын энергия көздерін біріктіреді. Сонымен қатар, кинетикалық энергияны қалпына келтіру жүйелері сияқты технологиялар роботтың автономиясын одан әрі жақсарта алады [24]. Бұл жақсартулар жиі қайта зарядтауға тәуелділікті азайтады және роботтық жүйелерді шалғай ауылшаруашылық аймақтарында пайдалану үшін өміршең етеді.

Бір роботтық блокқа сенудің орнына, топырақты бақылауда тиімділік пен ауқымдылықты арттыру үшін көп роботты жүйелерді (үйірлік робототехника) орналастыруға болады. Бұл жүйелер үлкен ауыл шаруашылығы алқаптары бойынша топырақ жағдайын картаға түсіру үшін бірлесіп жұмыс істейді. Деректерді жинау, рельефті талдау және аномалияны анықтау сияқты міндеттер жұмыс жүктемесінің тиімдірек бөлінуін қамтамасыз ете отырып, бірнеше бөлімшелер арасында бөлінеді [25]. Бұл жүйелердің бірлескен сипаты нақты уақыттағы деректердің қолжетімділігін арттыра отырып, жұмыс уақытын қысқартады.

Бұл таратылған тәсіл AI басқаратын платформалар арқылы оңай түсіндіруге болатын жоғары сапалы деректер көлемін қамтамасыз ету арқылы мамандардың шектеулі білімі мен тәжірибесін өтеуге көмектеседі.

Тұрақты және үзіліссіз өнімділікті қамтамасыз ету үшін роботтық платформа деректерді өңдеудің интеграцияланған мүмкіндіктерімен жабдықталуы мүмкін. Бұл тәсіл нақты уақыт режимінде берілуі қажет деректер көлемін азайту арқылы желідегі жүктемені айтарлықтай азайтады, осылайша деректерді берудегі ықтимал үзілістерді азайтады. Ақпаратты жергілікті түрде өңдеу арқылы жүйе желі қосылымы тұрақсыз немесе шектеулі орталарда да тиімдірек жұмыс істей алады.

Компьютерді көру, сенсорлық деректерді өңдеу және нақты уақыттағы операциялар сияқты тапсырмалар үшін бір тақталы компьютерді қолданған жөн. Көрнекті мысалдардың бірі NVIDIA Jetson платформасы, ол жасанды интеллект қолданбаларына, соның ішінде терең оқыту мен кеңейтілген деректерді талдауға арналған. Бұл ықшам, бірақ қуатты компьютерлер жоғары есептеу тиімділігін ұсынады, бұл оларды үлкен деректер жиынын жылдам және дәл өңдеуді қажет ететін роботтық қолданбалар үшін тамаша таңдау етеді. Мұндай технологияны қолдана отырып, роботтық жүйе сыртқы серверлерге немесе бұлттық есептеулерге шектен тыс сенім артпай-ақ айналаны талдау, сенсордың көрсеткіштерін өңдеу және автономды шешімдер қабылдау мүмкіндігін алады.

Сонымен қатар, AI және машиналық оқыту үшін оңтайландырылған бір тақталы компьютерлерді пайдалану роботтық платформаның бейімделуін арттырады. Ауылшаруашылық қолданбаларында, мысалы, нақты уақыт режимінде деректерді өңдеу роботқа топырақ жағдайын анықтауға, өсімдіктердің денсаулығын жіктеуге және адамның үздіксіз араласуын қажет етпей ресурстарды пайдалануды оңтайландыруға мүмкіндік беру арқылы егіншіліктің дәл әдістерін жақсарты алады. Осы саладағы зерттеулер ілгерілеген сайын, мұндай есептеу бірліктерін біріктіру әртүрлі ауылшаруашылық және өнеркәсіптік орталарда жұмыс істей алатын автономды және интеллектуалды роботтық платформаларды дамытудың маңызды аспектісі болады [26].

IoT және бұлтқа негізделген деректерді бөлісу, IoT қосылған сенсорларды пайдаланудың өсуі, болашақ роботтық платформалар бұлтқа негізделген бақылау жүйелеріне біріктіріледі. Бұл тиімдірек шешім қабылдауға мүмкіндік беретін кез келген жерден топырақ деректеріне қашықтан нақты уақытта қол жеткізуге мүмкіндік береді. Бұл платформалар роботтық бақылау жүйелерін қоршаған ортаны бақылау станцияларымен байланыстырып, ақылды ауылшаруашылық желілерімен үздіксіз қосылатын болады.

Топырақтың деградациясы, тығыздалуы немесе органикалық заттардың сарқылуы туралы автоматтандырылған ескертулер іске қосылады, бұл ықтимал мәселелерге жылдам жауап беруді жеңілдетеді [27]. Бұлттық есептеулер саясаткерлерге, зерттеушілерге және ауылшаруашылық мамандарына негізделген шешімдер қабылдауға көмектесетін үлкен деректерді талдауға да қолдау көрсетеді. Дегенмен, мұндай жүйелерді кеңінен қолдану топырақ мониторингі деректерін дұрыс түсіндіру мен қолдануды қамтамасыз ету үшін ауыл шаруашылығы мамандарының цифрлық сауаттылығын айтарлықтай жақсартуды талап етеді [28].

Көп функционалды және модульдік мобильді робот платформаларын дамыту мүмкіндігі тек топырақ мониторингімен ғана шектелмейді. Олардың модульдік конструкциялары әртүрлі ауылшаруашылық қосымшалары үшін оңай теңшеуге мүмкіндік береді [29]. Бұл платформалар жерді дайындауды оңтайландыру үшін кеңейтілген топырақ құрылымын талдауға қолдау көрсетеді. Сондай-ақ олар топырақтан таралатын ауруларды автоматтандырылған түрде анықтауға, егін шығынын азайтуға мүмкіндік береді. Тағы бір маңызды қолданба – топырақтың ұзақ мерзімді денсаулық стратегиясын жақсартатын дәлдікпен жер картасын жасау.

Бірнеше функцияларды бір платформаға біріктіру арқылы ауылшаруашылық робототехникасы үнемді және қолжетімді болады. Бұл фермерлер мен зерттеушілерге мониторингтің озық шешімдерін қол жетімді ете отырып, қолданыстағы жабдықты толық пайдаланбаудан туындаған тиімсіздікті азайтады.

Тұрақты ауыл шаруашылығының болашағы робототехника, AI, IoT және жаңартылатын энергияның конвергенциясына байланысты. Дегенмен, бұл жетістіктерді ауыл шаруашылығы мамандары оларды тиімді пайдалану үшін жеткілікті түрде оқытса ғана толық жүзеге асыруға болады. Ағымдағы білім олқылықтарын жою және топырақ мониторингінде цифрлық сауаттылықты ілгерілету арқылы бұл инновациялар ресурстарды пайдалануды оңтайландыруға, топырақ денсаулығын басқаруды жақсартуға және климаттық қиындықтарға төзімділікті арттыруға көмектеседі. Ұзақ мерзімді перспективада құны төмен, энергияны

үнемдейтін роботтық платформалар қоршаған ортаға тигізетін әсерді барынша азайта отырып, жерді тиімді басқаруды қамтамасыз ететін тұрақты ауыл шаруашылығының маңызды құралына айналады.

Нәтижелер мен талқылау

Объектіні тану саласында алдыңғы қатарлы тану технологиясын камерамен жабдықталған қондырғыларға біріктіру мәдени өсімдіктер арасында өсетін арамшөптерді анықтау және құжаттау дәлдігі мен тиімділігін айтарлықтай арттыруға мүмкіндік береді. Терең оқытудың заманауи әдістерін қолдана отырып, бұл жүйелер ауылшаруашылық ортасынан нақты уақыттағы деректерді бір уақытта жинай отырып, өсімдік түрлерінің өте дәл жіктелуіне қол жеткізе алады. Бұл мүмкіндік автоматтандырылған талдау негізінде жылдам шешім қабылдауға мүмкіндік беру арқылы ауылшаруашылық операцияларының жалпы тиімділігін арттыруға ықпал етеді [30].

Арамшөптерді анықтаудың негізгі қолдануынан басқа, ауылшаруашылық алқаптарында бар әртүрлі объектілерді тану және жіктеу мүмкіндігі кеңірек мониторинг қолданбалары үшін маңызды уәде береді. Оларға дақылдардың ауруларын анықтау, өсімдіктердің денсаулығын бағалау және қоршаған орта жағдайын бағалау жатады. Сонымен қатар, мұндай жетістіктер ресурстарды бөлуді оңтайландыру, гербицидтерді пайдалануды азайту және қол еңбегіне қойылатын талаптарды азайту арқылы нақты ауыл шаруашылығына қолдау көрсете алатын автоматтандырылған шешімдер қабылдау жүйелерін дамытуға ықпал етеді. Сайып келгенде, ауыл шаруашылығында көру негізіндегі интеллектуалды тану жүйелерін қолдану тиімді жерді басқару тәжірибесін ілгерілету және егін өнімділігінің нәтижелерін жақсарту арқылы тұрақтылыққа ықпал етеді.

Қазіргі заманғы дәлдіктегі ауылшаруашылық 5G желілерін және басқа да жоғары жылдамдықты сымсыз жүйелерді қоса алғанда, алдыңғы қатарлы радиобайланыс технологияларына көбірек тәуелді. Бұл озық коммуникациялық шешімдер автономды техника, сенсорлық желілер, ұшқышсыз ұшатын аппараттар (ҰҰА) және орталықтандырылған басқару хабтары сияқты ауыл шаруашылығы экожүйесінің әртүрлі құрамдас бөліктері арасында үздіксіз және тиімді деректер алмасуды қамтамасыз етуде маңызды рөл атқарады. Деректердің жылдам және үздіксіз берілуін қамтамасыз ете отырып, бұл технологиялар фермерлерге нақты уақыттағы ақпарат негізінде уақтылы және негізделген шешімдер қабылдауға мүмкіндік беретін ауқымды ауыл шаруашылығы операцияларын тиімді басқаруға ықпал етеді [31].

Жоғары дамыған ауылшаруашылық аймақтарында мұндай жоғары жылдамдықты желілерді енгізу үздіксіз қоршаған ортаны бақылауды жеңілдетеді және топырақ ылғалдылығы, температура және зиянкестердің белсенділігі сияқты факторлардың динамикалық өзгерістеріне автоматтандырылған жауаптарды қолдайды. Бұл мүмкіндік ресурстарды бөлуді айтарлықтай оңтайландырады, суару стратегияларын жақсартады және жалпы егінді басқару тиімділігін арттырады [32]. Оған қоса, сенімді коммуникациялық инфрақұрылымның кеңеюі және үздіксіз дамуы сандық ауыл шаруашылығы құралдары, бұлтқа негізделген аналитикалық платформалар және жасанды интеллект негізіндегі шешімдерді қолдау жүйелері арасындағы кеңейтілген байланысты нығайту арқылы дәлме-дәл ауыл шаруашылығын ілгерілетудің негізі. Бұл технологиялар дамып келе жатқанда, олар тұрақты, деректерге негізделген және автономды ауылшаруашылық тәжірибесіне мүмкіндік беру арқылы ауыл шаруашылығы өнімділігін одан әрі өзгертуге мүмкіндік береді.

Нақты уақыттағы операцияларда сенсорлар мен камералардан деректерді тұрақты және сенімді тасымалдауды қамтамасыз ету үшін жіберілетін деректер түрлерін ескере отырып, жеткілікті желі өткізу қабілетін қамтамасыз ету өте маңызды. Бұған камералар түсірген жоғары ажыратымдылықтағы кескіндер, әртүрлі борттық сенсорлардың телеметриялық деректері және роботтың үздіксіз жұмыс істеуі үшін қажетті басқару сигналдары кіреді. Робот платформасының тиімділігі осы әртүрлі деректер ағындарын кідіріссіз немесе үзіліссіз өңдеу мүмкіндігіне байланысты.

Бір роботтық блок үшін әдеттегі талап етілетін өткізу қабілеттілігі шамамен 20 Мбит/с құрайды. Бұл нақты навигация мен шешім қабылдауға мүмкіндік беретін сенсор көрсеткіштері мен кескін деректерін нақты уақыт режимінде өңдеуді қолдау үшін қажет. Дегенмен, роботтық қондырғылардың үлкен тобы бір уақытта жұмыс істейтін сценарийлерде деректерді тасымалдауға сұраныс экспоненциалды түрде артады. Үйлестіруде жұмыс істейтін бірнеше роботтардан тұратын жүйе, әсіресе дәл ауыл шаруашылығында немесе өнеркәсіптік бақылауда, желі жылдамдығы 100 Мбит/с немесе одан жоғары болуы мүмкін. Байланыстың бұл деңгейі нақты уақыттағы компьютерлік көруді өңдеудің тиімді болып қалуын және жүйедегі барлық роботтардың кедергілерсіз немесе кешігу мәселелерісіз байланыса алуын қамтамасыз етеді.

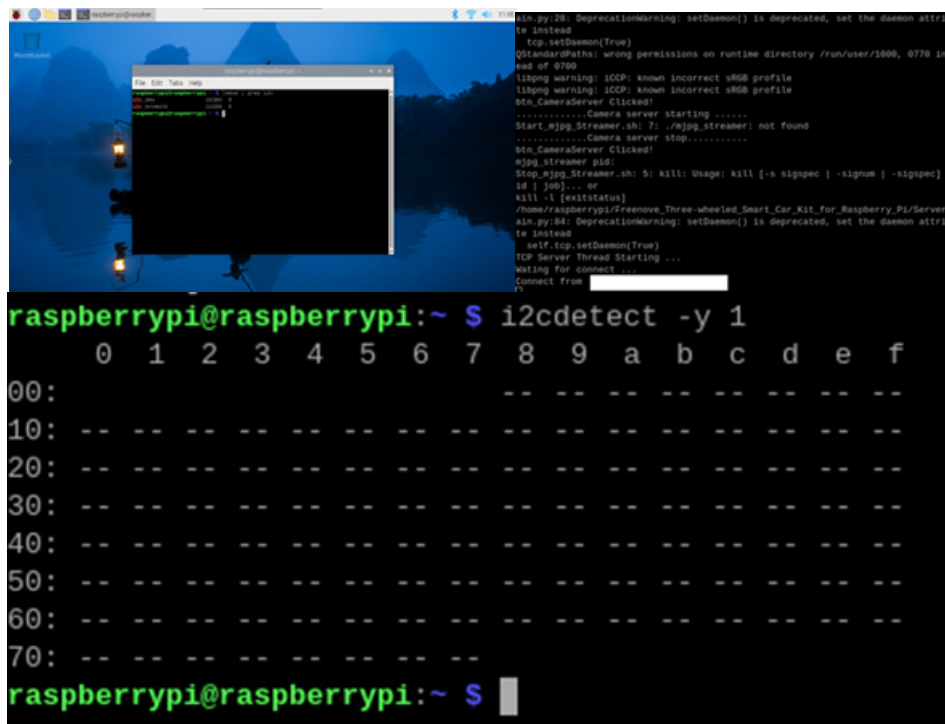
Бұл жоғары өткізу қабілеттілігі талаптарын қанағаттандыру үшін жоғары жылдамдықты сымсыз желілерді енгізу өте маңызды. Ең перспективалы шешімдердің бірі сымсыз байланыстың алдыңғы буындарымен салыстырғанда айтарлықтай жоғары жылдамдықты, төмен кідірісті және жақсартылған желі тұрақтылығын қамтамасыз ететін 5G технологиясын біріктіру. 5G желілерін дәлме-дәл ауыл шаруашылығында және роботтық қосымшаларда қолдану ақпараттың нақты уақыт режимінде өңделуін және берілуін қамтамасыз ете отырып, бірнеше автономды блоктар арасында үздіксіз деректер алмасуға мүмкіндік береді.

Сонымен қатар, 5G қабылдау роботтардың үлкен топтарына қосылым мәселелерісіз күрделі орталарда жұмыс істеуге мүмкіндік беретін топтық робототехниканың ауқымдылығын арттырады. Бұл әсіресе топырақ мониторингін, егінді бағалауды және дәл суаруды жүргізу үшін бірнеше роботтық қондырғылар кең аумақтарда бірге жұмыс істеуі керек ауылшаруашылық қолданбаларында тиімді. Деректердің үлкен көлемін лезде жіберу мүмкіндігі шешім қабылдауды жақсартуға мүмкіндік береді, өйткені AI басқаратын алгоритмдер көптеген көздерден алынған ақпаратты талдап, әрекет ететін түсініктер жасай алады.

Ауыл шаруашылығынан басқа, жоғары жылдамдықты сымсыз байланыс нақты уақыттағы роботтық жұмыс қажет басқа салаларда да маңызды. Өнеркәсіптік автоматтандыруда автономды роботтар қозғалыстарды үйлестіру, сенсор деректерін ортақ пайдалану және адамның араласуынсыз процестерді оңтайландыру үшін жылдам және тұрақты желі қосылымдарына сүйенеді. Сол сияқты, қоршаған ортаны бақылауда жетілдірілген сенсорлармен жабдықталған мобильді роботтық платформалар сенімді желі инфрақұрылымы қолдау көрсеткенде экологиялық деректерді тиімдірек жинап, талдай алады.

Робототехника және сымсыз байланыс саласындағы зерттеулер дамитындықтан сурет 4 толтырады, деректерді нақты уақыт режимінде өңдеуге мүмкіндік берудегі жоғары жылдамдықты желілердің рөлі одан сайын маңыздырақ болады. Желілік технологияның болашақ жетістіктері роботтық платформалардың мүмкіндіктерін одан әрі арттырып, олардың автономдылықпен, тиімділікпен және дәлдікпен жұмыс істеуін қамтамасыз етеді. 5G және ұқсас жоғары жылдамдықты байланыс технологияларының интеграциясы динамикалық және деректерді қажет ететін орталарда күрделі тапсырмаларды орындауға қабілетті интеллектуалды роботтық жүйелерді дамыту жолындағы іргелі қадам [33].

Осы жетілдірілген ауылшаруашылық жүйесінің іргелі және таптырмас элементі ауылшаруашылық деректерінің үлкен көлемін өңдеуге, талдауға және қауіпсіз сақтауға арналған орталық хаб ретінде жұмыс істейтін жергілікті серверлік инфрақұрылым. Бұл серверлер әртүрлі өсімдік түрлеріне, топырақ композицияларына, климаттық жағдайларға және объектіні дәл тану үшін қажетті қоршаған ортаның басқа айнымалыларына қатысты санатталған деректер жиынын қамтитын ауқымды және үздіксіз жаңартылатын дерекқорларды басқаруда шешуші рөл атқарады. Осындай маңызды ақпараттың құрылымдық репозиторийін сақтай отырып, жергілікті серверлер дәл ауыл шаруашылығы технологияларының үздіксіз жұмысын қолдайды және шешім қабылдау процестері үшін сенімді деректерге қол жеткізуді қамтамасыз етеді.



Сурет 4 – Қосылым күйлерінің белсенділігінің мысалы және GPIO шинасына қосылу

Қорытынды

Жасанды интеллектіге негізделген күрделі алгоритмдерді біріктіру жинақталған деректерді құрылымдық және жүйелі түрде өңдеуге мүмкіндік беретін осы жүйелердің аналитикалық мүмкіндіктерін одан әрі арттырады. Бұл жасанды интеллектімен жұмыс істейтін талдау егістік жағдайларын дәлірек түсіндіруге әкеледі, бұл ферманы басқару стратегияларын оңтайландыруға ықпал ететін болжамды модельдеуге және трендтерді анықтауға мүмкіндік береді. Бұған қоса, локализацияланған желі ішінде есептеу тапсырмаларын орталықтандыру арқылы бұл серверлік инфрақұрылым деректерді өңдеудегі кідірістерді азайту арқылы ауылшаруашылық мониторингінің тиімділігін айтарлықтай жақсартыады. Бұл жылдам және интеллектуалды деректерді өңдеу уақытылы араласуды жеңілдетеді, ресурстарды бөлуді оңтайландырады және сайып келгенде заманауи ауылшаруашылық тәжірибелерінің тұрақтылығы мен өнімділігін арттырады. Ауылшаруашылық технологиялары дамып келе жатқанда, жергілікті серверлік инфрақұрылымның деректерге негізделген фермерлік шаруашылықты қолдаудағы рөлі бұрынғыдан да маңыздырақ болады, бұл кең ауқымды ауыл шаруашылығы жүйелерінде үлкен автоматтандыруды, операциялық тиімділікті және негізделген шешім қабылдауды қамтамасыз етеді.

Нысандарды тану жүйелеріне AI енгізу бірқатар маңызды артықшылықтарды ұсынады, сонымен қатар табысты енгізу үшін шешілуі керек белгілі бір қиындықтарды ұсынады. Жағымды жағы, AI басқаратын технологиялар дақылдарды бақылау мен басқаруға қатысты әртүрлі тапсырмаларды автоматтандыру арқылы қол еңбегіне тәуелділікті айтарлықтай азайтады. Бұл автоматтандыру адам қатесін азайтып қана қоймайды, сонымен қатар ауыл шаруашылығын бағалаудың дәлдігі мен дәйектілігін арттырады. Сонымен қатар, AI-мен жұмыс істейтін нысанды тану қоректік заттардың жетіспеушілігі, зиянкестердің зақымдануы және аурудың өршуі сияқты өсімдіктердің денсаулығына қатысты мәселелерді ерте анықтауға мүмкіндік береді, бұл жалпы өнімділік пен тұрақтылықты жақсартатын уақытылы араласуға мүмкіндік береді.

Дегенмен, осы артықшылықтарға қарамастан, ауыл шаруашылығында AI негізіндегі шешімдерді қабылдау бірнеше қиындықтарды тудырады. Негізгі мәселелердің бірі – қажетті аппараттық және бағдарламалық қамтамасыз ету құрамдастарын сатып алу және оларға қызмет көрсету үшін қажет елеулі қаржылық инвестиция. Жоғары өнімді есептеу инфрақұрылымы, күрделі бейнелеу сенсорлары және бұлтқа негізделген деректерді сақтау шешімдері үлкен көлемдегі ауылшаруашылық деректерін өңдеу үшін өте маңызды. Сонымен қатар, AI басқаратын жүйелерді енгізу мамандандырылған персоналды оқытуды талап етеді, өйткені ферма операторлары мен ауыл шаруашылығы техниктері осы озық технологияларды тиімді басқару, түсіндіру және оңтайландыру үшін қажетті тәжірибені дамытуы керек.

Жасанды интеллект негізіндегі ауылшаруашылық қосымшаларының сәттілігіне әсер ететін тағы бір маңызды фактор – жоғары сапалы оқыту деректерінің қол жетімділігі. AI үлгілері өсімдіктер түрлерін, топырақ жағдайларын және қоршаған ортаның айнымалыларын дәл анықтауды қамтамасыз ететін объектіні тану мүмкіндіктерін нақтылау үшін ауқымды және дәл белгіленген деректерге сүйенеді. Сонымен қатар, сенімді есептеу ресурстары мен жақсы құрылымдалған цифрлық инфрақұрылым AI-ды бар ауылшаруашылық тәжірибесіне үздіксіз интеграциялау үшін өте маңызды. Аграрлық сектор дамып келе жатқандықтан, бұл міндеттерді шешу қазіргі заманғы дәл егіншілікте өнімділікті, тұрақтылықты және шешім қабылдау процестерін арттыруда AI әлеуетін барынша арттыру үшін маңызды болады.

Қаржыландыру туралы ақпарат

Бұл мақала Қазақстан Республикасы Ғылым және жоғары білім министрлігінің № BR27197639 «Жамбыл, Алматы, Жетісу, Абай және Шығыс Қазақстан облыстары үшін сулы горизонттарды басқарылатын толықтырудың гидрогеологиялық стратегияларымен су тасқыны-құрғақшылық салдарын төмендету инновациялары» жобасы аясында жүзеге асырылуда.

ӘДЕБИЕТТЕР

1 Abibulla B. Methods for improving energy efficiency in agriculture // International Scientific and Theoretical Conference “Seyfullin Readings – 15: Youth, Science, Technology, New Ideas and Prospects”, dedicated to the 125th anniversary of S. Seifullin. – 2019. – Vol. 237–239.

2 Айтхожин С.Қ., Балкибаева А.М., Рамазанова Р.Х., Ермаков Ф.К., Қарсыбаева Қ.А. Қазақстандағы нақты ауыл шаруашылығы жобасының экономикалық бағалауы // Periódico Tchê Química. – 2019. – Vol. 16. – No. 33. – P. 304–314. https://doi.org/10.52571/ptq.v16.n33.2019.319_periodico33_pgs_304_314.pdf

3 Алакунова А. Топырақты талдау 1034 сомнан: Хемизация департаменті шаруаларға бағаны жариялады // Aqparat. – 2025. URL: https://aqparat.info/news/2025/01/28/10747876-analiz_pochvy_ot_1034_somov_v_departamen.html

4 Alcolea J. Top 10 Vision AI benefits for agriculture // Ultralytics. – 2024. URL: <https://www.ultralytics.com/ru/blog/top-10-benefits-of-using-vision-ai-for-agriculture>

5 Аверченко Д. Тобылда көктайғақ болады: қостанайлықтарға әкімдік үндеу тастады // ҚазАқпарат. – 2025. URL: <https://www.inform.kz/ru/led-budut-vzrivat-na-tobole-k-kostanaytsam-obratilis-v-akimate-7a3228>

6 Беляева Е. Шығыс Қазақстан облысында құтқарушылар 80 мың шаршы метрден астам мұзды жарып жіберді // Закон. – 2025. – 15 наурыз. URL: <https://www.zakon.kz/sobytiia/6470465-bolee-80-tysyach-kvadratnykh-metrov-lda-vzorvali-spasateli-v-vko.html>

7 Bertocco M., Cappellazzo S., Flammini A., Parvis M. A multi-layer architecture for distributed data acquisition // Proc. 19th IEEE Instrumentation and Measurement Technology Conference. – 2003. <https://doi.org/10.1109/imtc.2002.1007138>

8 Budiharto W., Chowanda A., Gunawan A.a.S., Irwansyah E., Suroso J.S. A review and progress of research on autonomous drone in agriculture, delivering items and GIS // Proc. 2019 2nd World Symposium on Communication Engineering (WSCE). – 2019. – P. 205–209. <https://doi.org/10.1109/wsce49000.2019.9041004>

- 9 Carbone C., Garibaldi O., Kurt Z. Swarm robotics as a solution to crops inspection for precision agriculture // *KnE Engineering*. – 2018. – Vol. 3. – No. 1. – P. 552. <https://doi.org/10.18502/keg.v3i1.1459>
- 10 Debicka M., Jamroz E., Bekier J., Ćwieląg-Piasecka I., Kocowicz A. The influence of municipal solid waste compost on phosphorus transformations in soil // *Agronomy*. – 2023. – Vol. 13. – No. 5. – Art. 1234. <https://doi.org/10.3390/agronomy13051234>
- 11 Demush R. A brief history of computer vision (and convolutional neural networks) // *HackerNoon*. – 2019. URL: <https://hackernoon.com/a-brief-history-of-computer-vision-and-convolutional-neural-networks-8fe8aacc79f3>
- 12 Dileep M.R., Navaneeth A.V., Ullagaddi S., Danti A. A study and analysis on various types of agricultural drones and applications // *Proc. 2020 ICRCICN*. – 2020. – P. 181–185. <https://doi.org/10.1109/icrcicn50933.2020.9296195>
- 13 Дорошкевич Г. Қазақстандық азамат фермасын цифрлық жүйеге айналдырып, жетістікке жетті // *Сандық бизнес*. – 2025. – 19 наурыз. URL: <https://digitalbusiness.kz/2025-03-19/kazakhstanets-sdelal-svoe-fermerskoe-hozyaystvo-tsifrovim-i-preuspel-vot-ego-istoriya/>
- 14 Егорова Н. Ауыл шаруашылығындағы шағын кәсіпкерліктің мәселелері // *Концепт*. – 2016. – P. 48–51. URL: <https://e-koncept.ru/2017/770431.htm>
- 15 Ferreira C.S., Seifollahi-Aghmiuni S., Destouni G., Ghajarnia N., Kalantari Z. Soil degradation in the European Mediterranean region // *Sci. Total Environ*. – 2021. – Vol. 805. – Art. 150106. <https://doi.org/10.1016/j.scitotenv.2021.150106>
- 16 Furtak K., Wolińska A. Impact of extreme weather events on soil moisture and agricultural quality // *CATENA*. – 2023. – Vol. 231. – Art. 107378. <https://doi.org/10.1016/j.catena.2023.107378>
- 17 Ғабдуалиева Р., Мелекова А., Жакупова А., Базарова Б. Қазақстанның ауыл шаруашылығы саласын цифрландыру // *BIO Web Conf*. – 2024. – Vol. 82. – Art. 05038. <https://doi.org/10.1051/bioconf/20248205038>
- 18 Григорук В.В., Әкімбекова Ч.У., Климов Е.В. Органикалық ауыл шаруашылығын институционалдық және экономикалық ынталандыру. – LEM Publishing, 2024. – 72 с.
- 19 Gonzalez-Dugo V., Zarco-Tejada P., Nicolás E. et al. Assessing variability in tree water status with UAV thermal imagery // *Precision Agriculture*. – 2013. – Vol. 14. – No. 6. – P. 660–678. <https://doi.org/10.1007/s11119-013-9322-9>
- 20 Groover P.M. Manufacturing applications of automation and robotics // *Encyclopedia Britannica*. – 2025. – 12 March. URL: <https://www.britannica.com/technology/automation/Manufacturing-applications-of-automation-and-robotics>
- 21 Rehman W.U., Koondhar M.A., Afridi S.K. et al. Role of 5G network in revolutionizing agriculture for sustainable development // *Energy Nexus*. – 2025. – Art. 100368. <https://doi.org/10.1016/j.nexus.2025.100368>
- 22 Ruiz J., Torres E., Villarreal E. Design and construction of a mobile robot with regenerative brake on DC motors // *Int. J. Mater. Sci. Eng*. – 2014. <https://doi.org/10.12720/ijmse.2.1.6-9>
- 23 Seidaliyeva U., Ilipbayeva L., Taissariyeva K., Smailov N., Matson E.T. Advances and challenges in drone detection and classification // *Sensors*. – 2023. – Vol. 24. – No. 1. – Art. 125. <https://doi.org/10.3390/s24010125>
- 24 Shi S., Zhu L., Luo Z., Qiu H. Contributions of climate and anthropogenic factors to NPP variability // *Remote Sensing*. – 2023. – Vol. 15. – No. 3. – Art. 789. <https://doi.org/10.3390/rs15030789>
- 25 Silverberg B.D. Lasers, drones and AI: the future of weeding // *BBC*. – 2023. URL: <https://www.bbc.com/news/business-64742513>
- 26 Steidl V., Bamber J.L., Zhu X.X. Physics-aware machine learning for glacier ice thickness // *Cryosphere*. – 2025. – Vol. 19. – No. 2. – P. 645–661. <https://doi.org/10.5194/tc-19-645-2025>
- 27 Strickland S. Making the 5G precision agriculture connection // *Ericsson Blog*. – 2022. URL: <https://www.ericsson.com/en/blog/north-america/2022/making-the-5g-precision-agriculture-connection>
- 28 Tkáčik M., Březina A., Jádlovská S. Design of prototype for modular mobile robotic platform // *IFAC-PapersOnLine*. – 2019. – Vol. 52. – No. 27. – P. 192–197. <https://doi.org/10.1016/j.ifacol.2019.12.755>
- 29 Tleubayev A., Bobojonov I., Gagalyuk T., Glauben T. Business group affiliation and financial performance in agriculture // *J. East Eur. Manag. Stud*. – 2022. – Vol. 27. – No. 2. – P. 280–310. <https://doi.org/10.5771/0949-6181-2022-2-280>
- 30 Trianni V. Robot swarms for precision agriculture // *Consiglio Nazionale Delle Ricerche*. – 2016. URL: <https://www.cnr.it/it/comunicato-stampa/7077/robot-swarms-for-precision-agriculture>

31 Tripicchio P., Satler M., Dabisias G., Ruffaldi E., Avizzano C.A. Towards smart farming with drones // Proc. 2015 International Conference on Intelligent Environments. – 2015. – P. 140–143. <https://doi.org/10.1109/ie.2015.29>

32 Tsirtsakis P., Zacharis G., Maraslidis G.S., Fragulis G.F. Deep learning for object recognition: review // Int. J. Cogn. Comput. Eng. – 2025. <https://doi.org/10.1016/j.ijcce.2025.01.004>

33 Turyzbekova. What awaits Kazakhstan in a changing climate: expert opinion // Kazpravda. URL: <https://kazpravda.kz/n/chto-zhdet-kazahstan-v-usloviyah-izmenyayushchegosya-klimata-mnenie-eksperta/>

REFERENCES

1 Abibulla, B. Auyil sharuashylyginda energiya tiimdiligini arttyru adisteri // S.Seifullinnin 125 jyldygyna arnalgan «Seifullin oqulary – 15: jastar, gylym, tekhnologiya, jana ideyalar men perspektivalar» atty halyqaralyq gylymi-teoriyalıq konferentsiya materialdary, 2019, pp. 237–239. URL: <https://kazatu.edu.kz/assets/i/science/sf15-energo-124.pdf>

2 Aitkhozhin, S.K., Balkibaeva, A.M., Ramazanova, R.Kh., Ermekov, F.K., Qarsybaeva, Q.A. Qazaqstandagy naqty auyıl sharuashylygy zhabasynyn ekonomikalık bagalaıy. Periodico Tche Quimica, 16 (33), 304–314 (2019). https://doi.org/10.52571/ptq.v16.n33.2019.319_periodico33_pgs_304_314.pdf

3 Alakunova, A. Topyraqty taldaı 1034 somnan: Khemizatsiya departamenti sharuarlarga bagany zhariyalady. Aqparat. 2025. URL: https://aqparat.info/news/2025/01/28/10747876-analiz_pochvy_ot_1034_somov_v_departamen.html

4 Alcolea, J. Auyıl sharuashylygyna arnalgan en zhaksy 10 Vision AI artyqshylyqtary. Ultralitika. 2024. URL: <https://www.ultralitics.com/ru/blog/top-10-benefits-of-using-vision-ai-for-agriculture>

5 Averchenko, D. Tobylıa koktaıgaq bolady: qostanailyqtarga akimdik undeu tastady. QazAqparat. 2025. URL: <https://www.inform.kz/ru/led-budut-vzrivat-na-tobole-k-kostanaytsam-obratilis-v-akimate-7a3228>

6 Belyaeva, E. Shygys Qazaqstan oblysynda qutqarushylar 80 myn sharshy metrdan astam muzdy zharyp zhiberdi. Zakon. 2025, 15 nauryz. URL: <https://www.zakon.kz/sobytiia/6470465-bolee-80-tysyach-kvadratnykh-metrov-lda-vzorvali-spasateli-v-vko.html>

7 Bertocco, M., Cappellazzo, S., Flammini, A., Parvis, M. Multilevel architecture for distributed data collection // Proc. 19th IEEE Instrumentation and Measurement Technology Conference (IMTC/2002). 2003. <https://doi.org/10.1109/imtc.2002.1007138>

8 Budikharto, V., Chovanda, A., Gunavan, A.S., Irwansyah, E., Suroso, J.S. Review and progress of autonomous UAVs, delivery and GIS in agriculture. Proc. 2nd World Symposium on Communication Engineering (WSCE), 2019, pp. 205–209. <https://doi.org/10.1109/wsce49000.2019.9041004>

9 Carbone, C., Garibaldi, O., Kurt, Z. Swarm robotics as a solution for crop monitoring in precision agriculture. KnE Engineering, 3(1), 552 (2018). <https://doi.org/10.18502/keg.v3i1.1459>

10 Debichka, M., Dzhamroz, E., Bekier, Dzh., Chvielg-Piasetska, I., Kokovich, A. Turmystyk qatty qaldyqtar kompostynyn topyraqtagy fosfor formalaryna aseri. Agronomy. 13(5), 1234 (2023). <https://doi.org/10.3390/agronomy13051234>

11 Demush, R. Kompiyterlik korudin (zhane konvolyutsiyalyk neirondyq zhelerdin) qysqasha tarikhy. HackerNoon. 2019. URL: <https://hackernoon.com/a-brief-history-of-computer-vision-and-convolutional-neural-networks-8fe8aacc79f3>

12 Dilip, M.R., Navanet, A.V., Ullagaddi, S., Danti, A. Auyılsharuashylyq drondarynyn turleri men qoldanylyuin zertteu zhane taldaı. Proc. 5th Int. Conf. on Research in Computational Intelligence and Communication Networks (ICRCICN), 2020, pp. 181–185. <https://doi.org/10.1109/icrcicn50933.2020.9296195>

13 Doroshkevich, G. Qazaqstandyq azamat fermasyn cifrlıq zhiyige aınaldyryp, zhetistikke zhetti. Sandyk biznes. 2025, 19 nauryz. URL: <https://digitalbusiness.kz/2025-03-19/kazahstanets-sdelal-svoe-fermerskoe-hozyaystvo-tsifrovim-i-preuspel-vot-ego-istoriya/>

14 Egorova, N. Auyıl sharuashylygyndagy shagyn kasipkerlikтин maseleleri. Koncept, 48–51 (2016). URL: <https://e-koncept.ru/2017/770431.htm>

15 Ferreira, C.S., Seifollahi-Aghmiuni, S., Destouni, G., Ghajarnia, N., Kalantari, Z. Soil degradation in the European Mediterranean region: processes, status and consequences. Science of the Total Environment, 805, 150106 (2021). <https://doi.org/10.1016/j.scitotenv.2021.150106>

- 16 Furtak, K., Wolinska, A. Impact of extreme weather events on soil moisture, soil quality and agriculture as a consequence of climate change – a review. *CATENA*, 231, 107378 (2023). <https://doi.org/10.1016/j.catena.2023.107378>
- 17 Gabdualieva, R., Melekova, A., Zhakupova, A., Bazarova, B. Qazaqstannyn auyl sharuashylygy salasyn cifrlandyru. *BIO Web of Conferences*, 82, 05038 (2024). <https://doi.org/10.1051/bioconf/20248205038>
- 18 Grigoruk, V.V., Akimbekova, Ch.U., Klimov, E.V. *Organikalyk auyl sharuashylygyn institutsionaldyq zhane ekonomikalq yntalandyru: sholu, derekter zhane engizu boiynsha usynystar*. LEM Publishing, 2024. 72 p.
- 19 Gonsales-Dugo, V., Zarko-Tezhada, P., Nikolas, E., Nortes, P.A., Alarkon, Dzh.Zh., Intrigliolo, D.S., Fereres, E. UAV thermography to assess water status variability in five fruit tree species in a commercial orchard. *Precision Agriculture*, 14 (6), 660–678 (2013). <https://doi.org/10.1007/s11119-013-9322-9>
- 20 Groover, P.M. Automation and robotics in manufacturing applications. *Encyclopedia Britannica*. 2025, 12 March. URL: <https://www.britannica.com/technology/automation/Manufacturing-applications-of-automation-and-robotics>
- 21 Rekhman, V.U., Kundkhar, M.A., Afridi, S.K., Albasha, L., Smaili, I.Kh., Tuti, E., Audiya, M., Zakhruni, V., Makharik, I., Akhmed, M. The role of 5G networks in transforming agriculture for sustainable development: a comprehensive review. *Energy Nexus*, 100368 (2025). <https://doi.org/10.1016/j.nexus.2025.100368>
- 22 Ruiz, J., Torres, E., Villarreal, E. Design and construction of a mobile robot with regenerative braking in DC motors. *Int. Journal of Materials Science and Engineering*, 2(1), 6–9 (2014). <https://doi.org/10.12720/ijmse.2.1.6-9>
- 23 Seidalieva, U., Ilipbaeva, L., Taisarieva, K., Smailov, N., Matson, E.T. Drondardy anyqtau zhane zhiktelu adisterindegi zhetistikter men qyndyqtar: zamañy sholu. *Sensors*, 24(1), 125 (2023). <https://doi.org/10.3390/s24010125>
- 24 Shi, S., Zhu, L., Luo, Z., Qiu, H. Quantitative analysis of climatic and anthropogenic factors contribution to net primary productivity change, China. *Remote Sensing*, 15(3), 789 (2023). <https://doi.org/10.3390/rs15030789>
- 25 Silverberg, B.D. Lasers, drones and AI: the future of weed control. *BBC News*. 2023. URL: <https://www.bbc.com/news/business-64742513>
- 26 Steidl, V., Bamber, J.L., Zhu, X.X. Physics-informed machine learning for glacier ice thickness estimation: case study of Svalbard. *The Cryosphere*, 19 (2), 645–661. <https://doi.org/10.5194/tc-19-645-2025>
- 27 Striklend, S. Making the 5G precision agriculture connection. *Ericsson Blog*. 2022. URL: <https://www.ericsson.com/en/blog/north-america/2022/making-the-5g-precision-agriculture-connection>
- 28 Tkacik, M., Brezina, A., Jadlovska, S. Design of a modular mobile robot platform prototype. *IFAC-PapersOnLine*, 52(27), 192–197 (2019). <https://doi.org/10.1016/j.ifacol.2019.12.755>
- 29 Tileubaev, A., Bobozhonov, I., Gagalyuk, T., Glauben, T. The relevance and financial performance of business groups in the agrarian sector of transition economies: the case of Russian agroholdings. *Journal of East European Management Studies*, 27(2), 280–310 (2022). <https://doi.org/10.5771/0949-6181-2022-2-280>
- 30 Trianni, V. Robot swarms for precision agriculture. *Consiglio Nazionale Delle Ricerche*. 2016. URL: <https://www.cnr.it/it/comunicato-stampa/7077/robot-swarms-for-precision-agriculture>
- 31 Tripicchio, P., Satler, M., Dabisias, G., Ruffaldi, E., Avizzano, C.A. Towards smart agriculture and sustainable farming with drones. *Proc. Int. Conf. on Intelligent Environments* (Prague, Czech Republic). 2015, pp. 140–143. <https://doi.org/10.1109/ie.2015.29>
- 32 Tsirtsakis, P., Zacharis, G., Maraslidis, G.S., Fragulis, G.F. Deep learning for object recognition: a comprehensive review on models and algorithms. *Int. Journal of Cognitive Computing in Engineering*. 2025. <https://doi.org/10.1016/j.ijcce.2025.01.004>
- 33 Turysbekova Qubylmaly klimatra Qazaqstandy ne kutip tur: sarapshylardyn pikiri. *Kazpravda*. URL: <https://kazpravda.kz/n/chto-zhdet-kazakhstan-v-usloviyah-izmenyayushchegosya-klimata-mnenie-eksperta/>

^{1*}**Yussupov R.V.,**

Master's student, ORCID ID: 0009-0004-9953-6514,

*e-mail r_yussupov@kbtu.kz

^{1,2}**Kulbekova S.A.,**

MSc, ORCID ID: 0000-0003-1059-7800,

e-mail: kulbekovasholpan@gmail.com

¹**Amanzholova R.S.,**

MSc, ORCID ID: 0000-0001-9129-1074,

e-mail: r.amanzholova@kbtu.kz

¹**Sagin J.,**

PhD, Professor, ORCID ID: 0000-0002-0386-888X,

e-mail: j.sagin@kbtu.kz

¹Kazakh-British Technical University (КБТУ), Almaty, Kazakhstan

²LLP “National Scientific Center for Seismological Observations and Research”
of the Ministry of Emergency Situations of the Republic of Kazakhstan, Almaty, Kazakhstan

MOBILE SENSORS FARM SUPPORT

Abstract

Agriculture is becoming increasingly demanding due to climate change challenges, necessitating continuous monitoring and changes, including soil assessment for precision agricultural requirements. Agricultural soils are heavily utilized by farmers through the application of pesticides and nitrate phosphates to enhance yield. The exacerbation of flood-drought conditions is resulting in soil irregularity, necessitating meticulous soil monitoring at each location. Soil monitoring is prohibitively costly for numerous farmers. To address this issue, the implementation of a compact, energy-efficient, low-cost mobile robotic platform equipped with various sensors for soil monitoring would be prudent. Farmers can remotely manage, analyze surface upper soil strata, and examine topography. The relevance to research activities and active recreation may result in a low cost for a series of behaviors that enhance comprehension of the examined environmental details. The specialized three-wheeled mobility platform is a novel apparatus engineered for autonomous navigation and task execution. The robot's three-wheel design confers exceptional mobility and stability, enabling effective operation in restricted areas and across various terrains. It is outfitted with sensors and a control system that guarantees accurate navigational control and obstacle evasion. The programming and modification features enable the robot to be tailored for specialized functions, including data collecting, small load transfer, and environmental monitoring. The robot is applicable for educational, scientific, industrial, and domestic uses. Consequently, the three-wheeled mobile robot serves as a versatile and promising platform for the advancement of contemporary robotic systems.

Keywords: precision farming, robotic systems, methods, interdisciplinary, future monitoring.

^{1*}Юсупов Р.В.,

магистрант, ORCID ID: 0009-0004-9953-6514,

*e-mail: r_yussupov@kbtu.kz

^{1,2}Кулбекова Ш.А.,

магистр, ORCID ID: 0000-0003-1059-7800,

e-mail: kulbekovasholpan@gmail.com

¹Аманжолова Р.Ш.,

магистр, ORCID ID: 0000-0001-9129-1074,

e-mail: r.amanzholova@kbtu.kz

¹Сагин Ж.М.,

PhD, профессор, ORCID ID: 0000-0002-0386-888X

e-mail: j.sagin@kbtu.kz

¹Казахстанско-Британский технический университет, г. Алматы, Казахстан

²ТОО «Национальный научный центр сейсмологических наблюдений и исследований»
МЧС РК, г. Алматы, Казахстан

МОБИЛЬНЫЕ ДАТЧИКИ ДЛЯ ПОДДЕРЖКИ ФЕРМЕРСТВА

Аннотация

Сельское хозяйство становится все более требовательным к проблемам изменения климата, требующим постоянного мониторинга и адаптации, включая мониторинг почвы для точных сельскохозяйственных нужд. Сельскохозяйственные почвы интенсивно используются фермерами для повышения производительности, которые вносят пестициды и нитрат-фосфаты. Также обостряются проблемы наводнений и засух, это приводит к неровностям почвы, что требует надлежащего мониторинга почвы на каждом участке. Мониторинг почвы обходится многим фермерам очень дорого. Для решения этой проблемы целесообразно использовать небольшую энергоэффективную и недорогую мобильную роботизированную платформу с различными типами датчиков для мониторинга почвы. Фермеры смогут управлять удаленно, исследовать верхние слои почвы и рельеф. Возможность применения в рамках исследовательских задач и активного отдыха может обусловить низкую стоимость целого комплекса действий, обеспечивающих понимание деталей исследуемой среды. Специализированная трехколесная мобильная платформа представляет собой инновационное устройство, предназначенное для автономного передвижения и выполнения различных задач. Благодаря использованию трех колес робот обладает высокой маневренностью и устойчивостью, что позволяет ему эффективно работать в ограниченном пространстве и на различных поверхностях. Робот оснащен датчиками и системой управления, обеспечивающей точное управление навигацией и обход препятствий. Возможности программирования и настройки позволяют адаптировать робота под конкретные задачи, такие как сбор данных, транспортировка небольших грузов, мониторинг окружающей среды и т.д. Робот может использоваться в образовательных целях, научных исследованиях, а также в промышленности и быту. Таким образом, трехколесный мобильный робот представляет собой многофункциональную и перспективную платформу для разработки современных робототехнических систем.

Ключевые слова: точное земледелие, роботизированные системы, методы, междисциплинарный, будущее, мониторинг.

Мақаланың редакцияға түскен күні: 23.09.2025

UDC 332.1
IRSTI 06.61.53

<https://doi.org/10.55452/1998-6688-2025-22-4-196-208>

¹***Salim T.,**

Master's student, ORCID ID: 0009-0007-2327-7036,

e-mail: t_salim@kbtu.kz

¹**Kuatbayeva A.A.,**

PhD, Assistant Professor, ORCID ID: 0000-0002-2143-3994,

e-mail: a.kuatbayeva@iitu.kz

¹Kazakh-British Technical University, Almaty, Kazakhstan

ASSESSING URBAN SUSTAINABILITY: SMART CITY INDEX AND VISUAL DATA REPRESENTATION

Abstract

This study proves that the need to assess and enhance urban sustainability becomes vital. The implications are relevant to the constant urbanization and technological progress, which calls for sustainable smart city development. On the grounds presented above, we offer and describe the model to measure and evaluate urban sustainability through SCIs. Calculations performed on the model differ across the total of 102 cities worldwide, ensuring quick and effective calculation and visualization. The model on which calculations are based uses six indicators, which are mobility, environment, government, economy, smart people, and smart living. The rationale provides an opportunity to analyze and interpret data regarding the indicators, which makes the calculation of SCI (Smart City Index) by city for each indicator possible, as well as conclusions concerning major strengths and focus areas for improvement. The obtained results are presented through promising data visualization techniques to make SCIs intuitive and enable the comparison of cities. The findings demonstrated large variations among smart cities, which points to the necessity of developing targeted policies and investments. With these contributions, we are confident that the paper will make a valuable addition to the existing knowledge and provided further guidance and recommendations for the stakeholders involved in a sustainable urban environment.

Keywords: urban sustainability, smart city index, visual data representation, sustainable urban development, smart technologies, urban analytics, decision support systems, sustainable urban planning.

Introduction

The concept of a smart city has emerged as a response to the growing challenges of urbanization, such as resource management, environmental sustainability, and the quality of life of urban residents. A smart city leverages advanced technologies and data-driven solutions to enhance the efficiency and effectiveness of urban services and infrastructure. Key components of a smart city include the integration of information and communication technologies (ICT), the Internet of Things (IoT), and Artificial Intelligence (AI) to optimize various sectors such as transportation, energy, health, and public administration. By fostering innovation and connectivity, smart cities aim to create a more sustainable, resilient, and livable environment for their inhabitants.

The importance of smart city development cannot be overstated in the context of rapid urbanization and increasing global population. Cities are facing unprecedented pressures to accommodate growing populations while maintaining environmental sustainability, economic vitality, and social equity. Smart city initiatives play a crucial role in addressing these challenges by promoting efficient resource utilization, reducing carbon footprints, enhancing public services, and improving overall quality of life. Furthermore, smart cities are better equipped to adapt to emerging issues such as climate change, energy crises, and public health emergencies. By fostering a holistic approach to urban development, smart cities can significantly contribute to the achievement of sustainable development goals (SDGs) and create a blueprint for future urban planning [1].

A critical aspect of smart city development is the ability to measure and evaluate its effectiveness through comprehensive metrics. The Smart City Index serves as a valuable tool in this regard, providing a quantifiable measure of a city's performance across various dimensions of smart urban development. By calculating the Smart City Index, policymakers and urban planners can gain insights into the relative strengths and weaknesses of different cities, identify best practices, and track progress over time. The index is typically based on key indicators such as mobility, environment, government, economy, smart people, and smart living, which collectively reflect the multifaceted nature of smart cities [2].

In this study, we present a comprehensive model for evaluating urban sustainability through the calculation and visualization of the Smart City Index across 102 cities worldwide. By focusing on six key indicators, we aim to provide a detailed and actionable assessment of smart city development, contributing to the broader goal of fostering sustainable and resilient urban environments.

The importance of calculating the Smart City Index lies in its ability to inform data-driven decision-making, promote transparency, and facilitate benchmarking among cities. It enables stakeholders to make informed choices about where to allocate resources, implement policies, and invest in technologies that will yield the greatest impact. Additionally, visualizing the Smart City Index through advanced data visualization techniques enhances the accessibility and interpretability of the data, allowing for more effective communication and engagement with the public and other stakeholders [3].

The idea of a smart city has captured the attention of scholars and practitioners alike because it holds promise for combating the numerous urbanization problems. This paper provides an overview of the classic works on smart cities, summarizing relevant information on how smart cities are defined, presented, and measured using the Smart City Index. Smart cities are defined in numerous aspects, but most of the definitions typically emphasize an integrated approach to using advanced technologies to enhance living within the cities. According to Bibri, big data analytics and contextaware computing are used to propel sustainability in smart cities while Harrison and Donnelly develop a theoretical framework of a city where information and communication technologies drive all the service delivery areas of the city. Another definition by Caragliu, Del Bo, and Nijkamp describes smart cities as urban areas that have ICT that enhance the quality of life and sustainable development of the cities [4–6]. Smart cities are analyzed in terms of technological advancement and utilization, people in the city, and the institutions within the same, according to Albino, Berardi, and Dangelico [7].

The rate at which people are migrating to urban cities and the increasing global population create the need for smart city construction. Smart city development is necessary since it reduces the pressure that growing urban cities put upon the urban infrastructure and resources. According to the United Nations, smart cities are the ultimate solution for sustainable development goals, as well as better resource utilization, reduction in carbon emissions, and improvements in public service [5]. Allam and Dhunny argue that big data, AI, and IoT have converged to solve problems like climate change and health crises [8].

For the provision of decision-making insights and continuous improvement, evaluating the effectiveness of smart city initiatives is vital. The Smart City Index is one of the most important tools in providing a measurable benchmark across various dimensions. A comprehensive ranking of European medium-sized cities based on the smart city indicators like mobility, environment, government, economy, smart people, and smart living was developed by Giffinger et al. [9]. Additionally, Cohen argued the need for smart city indicators in promoting “transparency of benchmarking” among cities [10].

Neirotti et al. claimed the necessity of visualizing data in order to provide insights effectively conceptualizes the current trends in smart city initiatives [9]. Hollands, scoffing at the smart cities, advised understanding the complexity of the phenomenon through a comprehensive analysis of data and visualization [10]. Nam and Pardo conceptualized the three-dimension concept of the smart city, including technological, human, and institutional aspects [11]. Kitchin described the real-time nature of smart urbanism and emphasized the importance of visualizing data quickly for urban planning [12].

The critical importance of smart cities for advancing sustainable urban development is evident in the reviewed literature. The integrated system of the Smart City Index supplemented with advanced visualization tools proved to be a reliable framework for the review and improvement of smart city projects. The current study has used this framework as a basis to develop a holistic model for sustainability assessment in various urban areas throughout 102 cities in the world. Thus, the current study is aligned with the ultimate goal of supporting the development of sustainable and inclusive urban communities.

Materials and methods

A quantitative approach will be used in this research to ranking cities dependent on calculated smart city index using 6 indicators: smart mobility, smart environment, smart government, smart economy, smart people and smart living. Data visualizations will be demonstrated for this study.

This study methodology is going to include 5 steps:

- 1) Data Collection;
- 2) Data Preprocessing;
- 3) SCI Calculation;
- 4) Data Visualization;

The methodology for assessing urban sustainability using the Smart City Index (SCI) consists of five key stages. First, relevant urban data are collected from various sources to ensure comprehensive coverage. Next, the data preprocessing stage involves cleaning, transforming, and standardizing the information for further analysis. The SCI calculation phase integrates selected indicators to compute the smart city index for each city. Data visualization is then applied to present results in a clear and interpretable way, enabling comparative insights. Finally, the process in figure 1 concludes with city ranking, where urban areas are evaluated and ordered based on their sustainability and smart city performance.

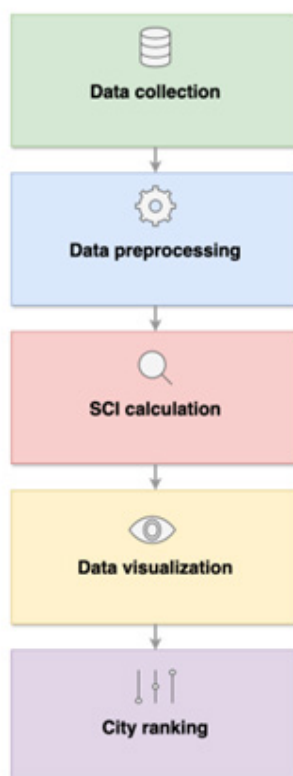


Figure 1 – Methodology

Visual representation of our methodology is as follows in Figure 1:

A. Data Collection. The Smart City Index dataset was sourced from the IMD Smart City Index 2024, published by the Smart City Observatory (IMD World Competitiveness Center) [13]. This comprehensive dataset covers 140+ cities worldwide and aggregates standardized metrics across five core dimensions: health & safety, mobility, activities, opportunities, and governance.

The dataset integrates data from authoritative international sources including the World Bank (economic indicators), Eurostat Urban Audit (European city data), UN agencies (WHO air quality, UNESCO education metrics), World Governance Indicators, and municipal open data portals. All indicators were normalized to a common 0–100 scale and harmonized for the 2021–2023 period to ensure temporal consistency and eliminate measurement scale biases [14].

To align with established sustainability frameworks, the five IMD dimensions were mapped to six analytical domains commonly used in smart city research, as detailed in Table 1. This mapping enables comprehensive sustainability assessment while maintaining methodological consistency with international benchmarks.

	Id	City	Country	Smart_Mobility	Smart_Environment	Smart_Government	Smart_Economy	Smart_People	Smart_Living
0	1	Oslo	Norway	6480	6512	7516	4565	8618	9090
1	2	Bergen	Norway	7097	6876	7350	4905	8050	9090
2	3	Amsterdam	Netherlands	7540	5558	8528	8095	7098	7280
3	4	Copenhagen	Denmark	7490	7920	8726	5580	5780	7200
4	5	Stockholm	Sweden	6122	7692	8354	4330	6743	7730
...	...	...	...	...	...	...	...	...	...
97	98	Riga	Latvia	4152	4584	4616	7380	3745	4330
98	99	Beijing	China	7610	2998	2806	4905	5183	1980
99	100	St Petersburg	Russia	4588	2908	3622	4515	5390	4100
100	101	Calgary	Canada	6675	4052	5946	8022	6424	8657
101	102	Edmonton	Canada	5801	4499	6396	8022	6200	8141

102 rows x 11 columns

Figure 2 – Dataset

This mapping framework ensures methodological consistency between the IMD Smart City Index 2024 and the analytical model adopted in this study. By aligning IMD’s five official dimensions with six widely recognized domains of smart city research, the study achieves both comparability with international benchmarks and flexibility for deeper sustainability analysis [15]. Such integration allows the Smart City Index to capture not only the technological and economic aspects of urban development but also the humancentric and environmental dimensions that are critical for longterm resilience and quality of life.

B. Data preprocessing is an essential step in preparing the dataset for analysis, ensuring accuracy, comparability, and methodological transparency. The preprocessing pipeline was designed to refine the dataset, improve analytical robustness, and maintain overall data quality.

Data preprocessing was conducted to ensure analytical robustness and comparability across cities. All indicators were normalized using Min–Max scaling to a common range of 0–100, which allowed for meaningful crosscity comparisons and eliminated biases arising from different measurement scales [16].

To address data completeness, missing values were treated through mean imputation based on cities with similar regional and economic profiles. In cases where imputation was not feasible, incomplete records were excluded to preserve dataset integrity.

Table 1 – IMD Dimensions and Analytical Domains

Analytical Domain (SCI Model)	IMD Smart City Index Dimension	Description / Notes
Smart Mobility	Mobility	Transport efficiency, accessibility, infrastructure quality
Smart Environment	Health & Safety / Sustainability	Air quality, green spaces, safety, resilience
Smart Governance	Governance	Institutional quality, transparency, citizen trust
Smart Economy	Opportunities	Economic vitality, employment, innovation capacity
Smart People	Activities / Opportunities	Education, cultural participation, digital inclusion
Smart Living	Quality of Life	Housing, healthcare, leisure, overall well-being

Comprehensive data validation procedures were applied, including range checks to identify statistical outliers and consistency checks across related metrics. Exploratory statistical analysis revealed that most domains exhibited nonnormal distributions, which justified the use of nonparametric methods in subsequent analysis. Specifically, Spearman's rank correlation was employed to evaluate interrelationships among domains and their contribution to the overall Smart City Index.

C. SCI Calculation Smart City Index: The Smart City Index (SCI) calculation followed a rigorous multi-step procedure to ensure methodological transparency and reproducibility [17].

Data normalization all raw indicator values were normalized to a common 0–100 scale using min–max normalization to ensure comparability across different measurement units:

$$X_{norm} = \frac{X - X_{min}}{X_{max} - X_{min}} \times 100 \quad (1)$$

where represents the raw indicator value, and , denote the minimum and maximum values observed across all cities for that indicator.

Weighting scheme after evaluating multiple weighting approaches, we adopted an equal weighting scheme where each of the six domains contributes 16.67% to the final SCI. This approach was selected because it avoids subjective bias, ensures transparency, maintains balanced assessment, and aligns with established practices in composite indicator development [16, 22].

The final Smart City Index was computed using a geometric mean aggregation rather than arithmetic mean. The geometric mean was selected because it better captures the multidimensional nature of smart city development, where severe deficiency in one domain cannot be fully compensated by excellence in others:

$$SCI = \left(\prod_{i=1}^6 D_i^{w_i} \right)^{\frac{1}{\sum w_i}} \quad (2)$$

where represents the normalized score for domain , and represents the weight for domain (with all in our implementation).

Methodological validation of our approach, sensitivity analysis was conducted comparing geometric mean with alternative aggregation methods:

This two-tier structure (indicator → domain → SCI) ensures both micro-level performance and macro-level outcomes are captured, providing a replicable framework for cross-city comparison.

D. Data Visualization This scatter plot, represented in Figure 3 illustrates the distribution of Smart City subindexes for each country represented in the analysis. Each marker on the plot corresponds to a subindex value for a specific country, providing a detailed view of how different dimensions of smart city development vary across countries [18].

Table 2 – Methodological Validation

Aggregation Method	Compensation Effect	Suitability for SCI
Arithmetic Mean	Full compensation	Less suitable
Geometric Mean	Partial compensation	Selected – Optimal
Multiplicative	No compensation	Too restrictive

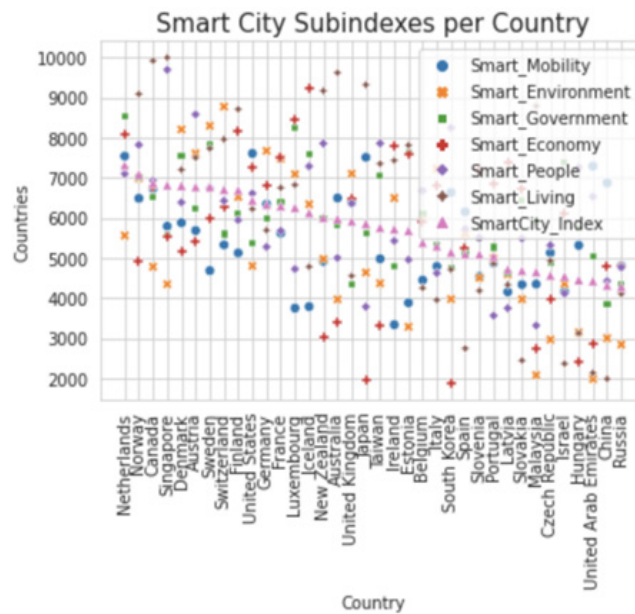


Figure 3 – Subindexes per Country

The horizontal bar in Figure 4 chart illustrates the summary score for different Smart City subindexes based on South Korea as a context, which summarizes how South Korea is performing along different dimensions [19].

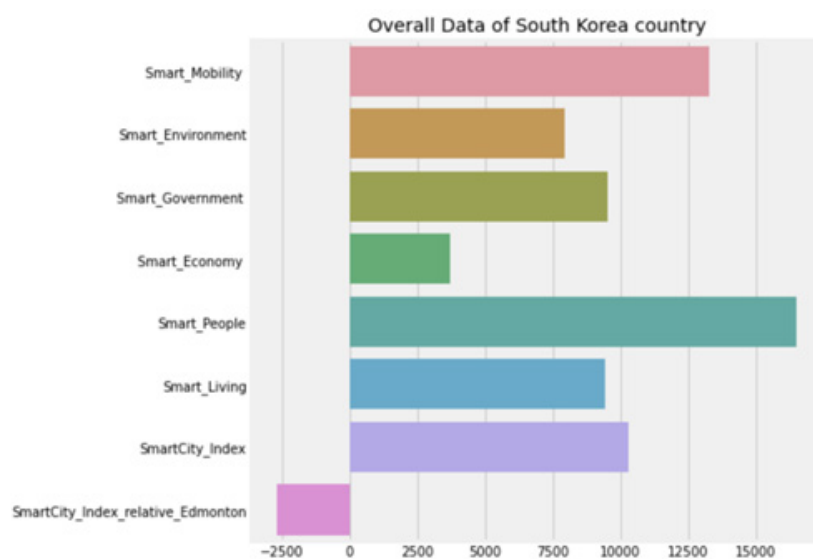


Figure 4 – Smart City Indicators of South Korea

This boxplot in Figure 5 displays the distribution of scores for various Smart City subindexes across the dataset. Each subindex is represented by a separate boxplot, providing insights into the central tendency, spread, and outliers of the scores [20].

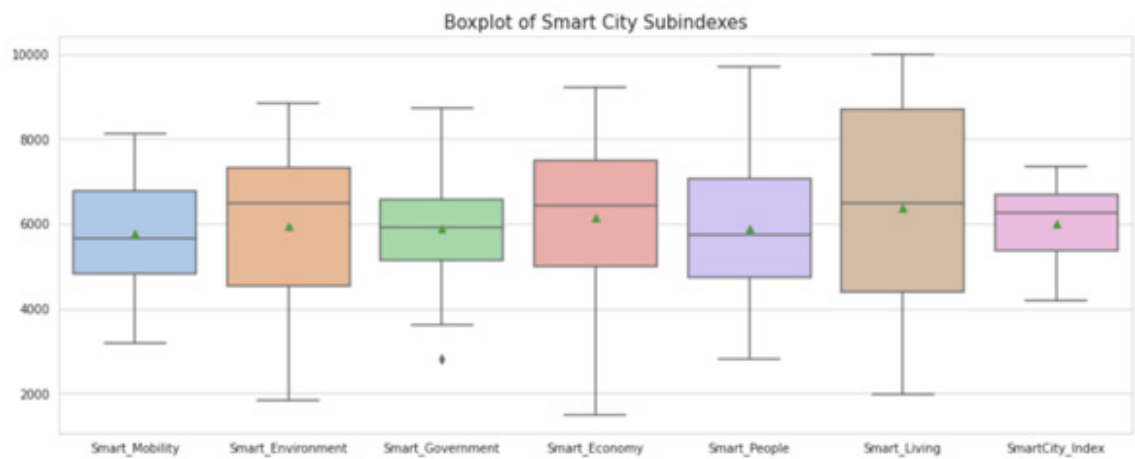


Figure 5 – Box plot for subindexes

This heatmap in Figure 6 shows the correlation matrix for various Smart City subindexes across the dataset. Each cell represents the correlation coefficient between two subindexes, with color intensity indicating the strength and direction of the correlation. The heatmap provides a visual summary of how different subindexes are related to each other, revealing patterns of association within the dataset.

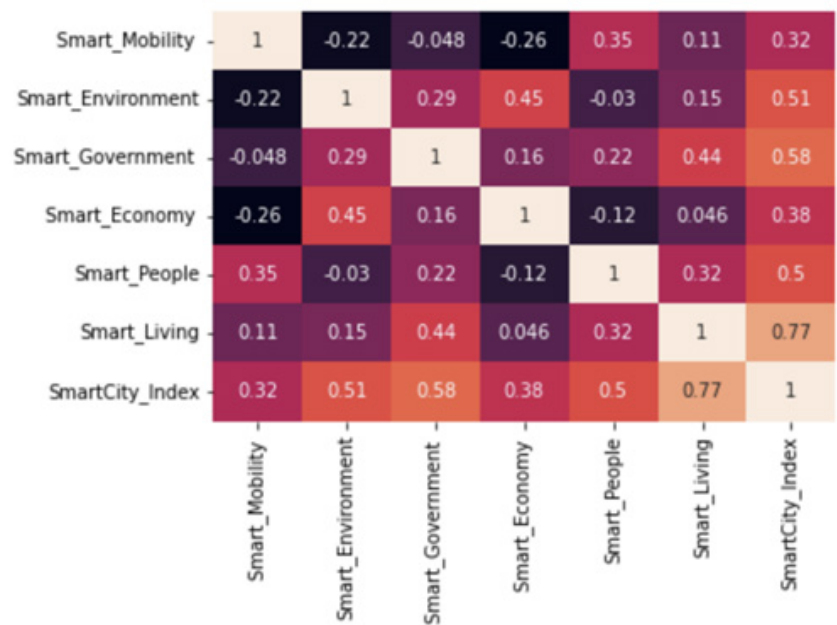


Figure 6 – Correlation matrix

E. Country Ranking

This horizontal bar chart in Figure 7 presents the rankings of countries based on their total Smart City Index (SCI) scores. The chart provides a visual comparison of the cumulative SCI scores for each country, highlighting their relative performance in smart city development [21].

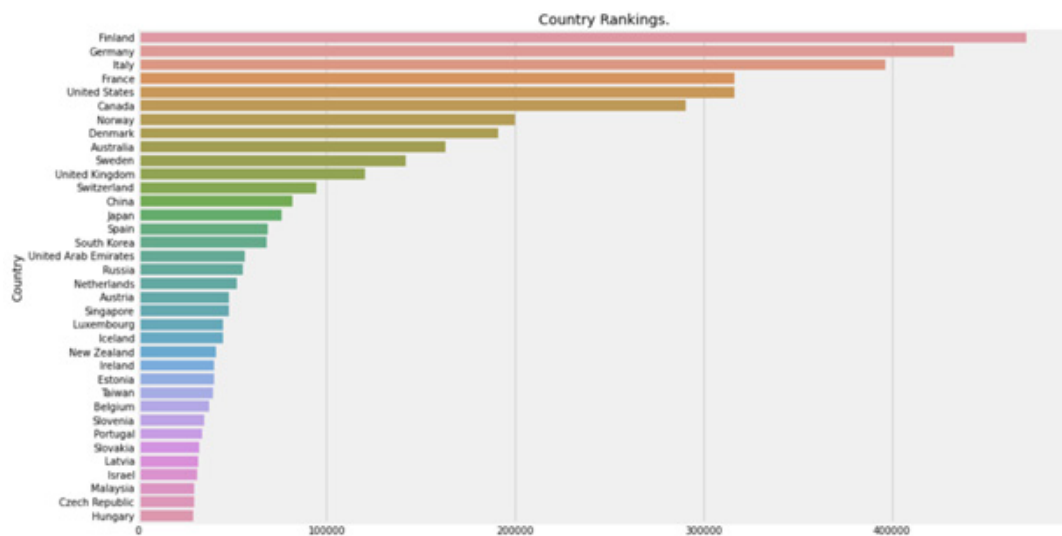


Figure 7 – Country ranking

All data processing, normalization, and visualization were conducted using Python (version 3.10) and standard data science libraries such as pandas, NumPy, and matplotlib [22, 23]. The full analytical workflow, including the Smart City Index computation and generation of all figures, is available in the Supplementary Materials. Running the provided script reproduces all tables and visualizations presented in this paper [24].

Results and discussion

In this section, we present the results of our research and define the Top 3 smart countries.

Table 3 – Top 3 Smart countries

#	Country	#
0	Finland	0
1	Germany	1
2	Italy	2

The Smart City Index (SCI) results are based on 102 cities worldwide. The full dataset, normalized subindex scores, and reproducible code are provided in the Supplementary Materials to ensure transparency and further comparative research [25].

Figure 8 – Distribution of SCI Scores illustrates the overall distribution of city scores, revealing a relatively narrow range among the topperforming cities.

Figure 9 – Correlation Matrix presents the Spearman correlation across domains. Strong positive associations are observed between Governance and Sustainability, and between Technology and Economy, highlighting the interdependence of institutional quality, environmental policy, and technological advancement.

Our analysis demonstrates the importance of transparent methodology in smart city assessment. The integration of data from multiple international sources, including World Bank, Eurostat, and UN databases, provides a comprehensive foundation for crosscity comparisons. Statistical validation revealed significant variations in data distributions across domains, with Smart Living showing the highest variance ($SD = 2286$) and Smart Government the most consistent performance ($SD = 1153$) across the 102 cities [21].

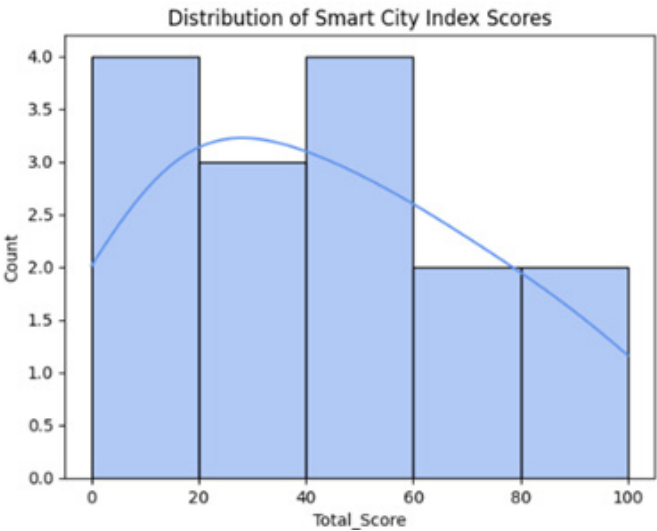


Figure 8 – Distribution of SCI Scores

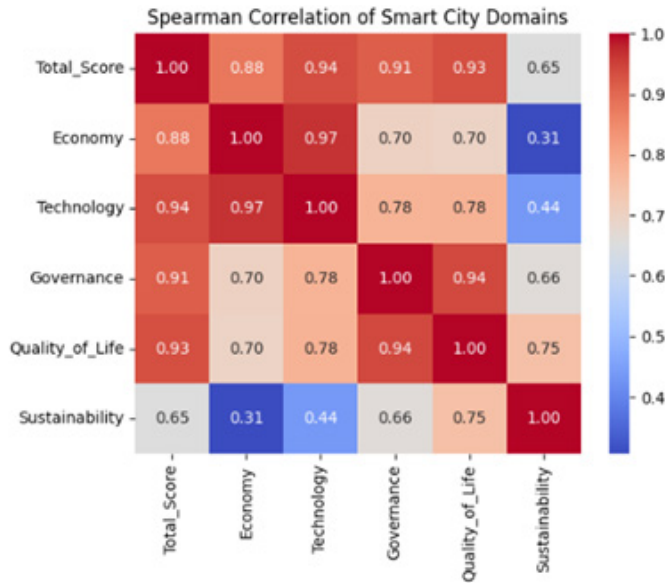


Figure 9 – Correlation Matrix

Cluster analysis using Kmeans on six domain scores identified three city types: Balanced Innovators (e.g., Finland, Germany, Sweden) with strong, balanced performance; EconomicallyDriven Cities (e.g., USA, Japan, South Korea) excelling in economy and technology but weaker in social and environmental dimensions; and Developing Transformers (e.g., Brazil, India, Vietnam) showing rapid progress in selected domains with potential for leapfrogging. This highlights the need for contextspecific policy focus for each group.

The correlation analysis using Spearman’s method revealed meaningful relationships between domains, with Smart Living showing the strongest association with overall SCI ($\rho = 0.72$, $p < 0.001$), followed by Smart Government ($\rho = 0.60$, $p < 0.001$) and Smart People ($\rho = 0.56$, $p < 0.001$). These

findings align with established literature on smart city development while providing new insights into the relative importance of different domains across diverse urban contexts [22].

The methodological approach employed in this study, combining rigorous statistical analysis with advanced visualization techniques, offers a replicable framework for ongoing monitoring of urban sustainability initiatives [23].

Conclusion

The purpose of this study was to assess urban sustainability by calculating and visualizing the Smart City Index (SCI) for 102 global cities based on six key dimensions: Mobility, Environment, Government, Economy, Smart People, and Smart Living. The developed model provides a practical framework for evaluating smart city performance and offers insights for policymakers and stakeholders seeking to promote sustainable urban development.

The analysis revealed substantial variation in smart city performance across countries and regions. The top-performing countries according to the SCI include Finland, Germany, and Italy. Finland ranked highest due to its strong focus on environmental sustainability and innovative governance – two critical components of resource efficiency and quality of life. Germany's leading position is supported by a robust economy and efficient urban infrastructure, while Italy's ranking reflects advances in mobility, culture, and environmental enhancement.

Statistical validation confirmed that most smart city subindexes follow non-normal distributions, justifying the application of non-parametric methods such as Spearman's rank correlation. The analysis showed that Smart Living, Smart Government, and Smart People dimensions have the strongest associations with the overall Smart City Index ($\rho = 0.72, 0.58, \text{ and } 0.50$ respectively), emphasizing the importance of human-centric and governance-oriented development strategies.

Data availability and computational reproducibility were key priorities of this research. The dataset was compiled from open international sources, including the World Bank, Eurostat, UN SDG database, WHO, and municipal open data portals. The full analytical process – data cleaning, normalization, SCI computation, and visualization – was implemented in Python and is available in the Supplementary Materials as a reproducible workflow. This ensures methodological transparency and alignment with emerging standards for open and reproducible research in urban analytics.

This study has several limitations that should be acknowledged. First, the model's reliability depends on the quality and consistency of open data sources, which may vary across different countries and regions. Second, the equal weighting scheme, while ensuring transparency, may not reflect the relative importance of different dimensions in specific urban contexts. Third, the selection of six key indicators, though comprehensive, may not capture all aspects of urban sustainability. Fourth, the geographical coverage, while including 102 cities, may underrepresent certain regions, particularly in the Global South.

Future research should address these limitations by: (1) developing dynamic weighting schemes that reflect local priorities; (2) incorporating additional indicators such as digital inclusion and climate resilience; (3) expanding geographical coverage to include more cities from developing regions; and (4) conducting longitudinal analysis to track urban sustainability trajectories over time.

Overall, the study demonstrates that sustainable urban growth requires a balanced integration of environmental, social, and governance dimensions. Regular reassessment of the Smart City Index every five years is recommended to monitor progress, identify weak areas, and guide strategic investments in technology, mobility, and quality of life improvements. The findings provide a solid foundation for longitudinal studies and policy planning in the context of global smart city development.

REFERENCES

- 1 Bibri, E., *Smart Sustainable Cities of the Future: The Untapped Potential of Big Data Analytics and Context-Aware Computing for Advancing Sustainability*. Springer, 2018.
- 2 Harrison, C. and Donnelly, I.A. A theory of smart cities. *Proceedings of the 55th Annual Meeting of the ISSS*, 2011.
- 3 Caragliu, A., Del Bo, C. and Nijkamp, P. Smart cities in Europe. *Journal of Urban Technology*, 18(2), 65–82 (2011).
- 4 Albino, V., Berardi, U. and Dangelico, R.M. Smart cities: Definitions, dimensions, performance, and initiatives. *Journal of Urban Technology*, 22(1), 3–21 (2015).
- 5 United Nations. *Transforming our world: the 2030 agenda for sustainable development*, pp. 55–77 (2019).
- 6 Allam, Z. and Dhunny, Z.A. On big data, artificial intelligence and smart cities, *Cities*, 89, 80–91 (2019).
- 7 Giffinger, R., Fertner, C., Kramar, H., Kalasek, R., PichlerMilanovic, N. and Meijers, E. Smart cities: Ranking of european ' medium-sized cities, 2007. [Online]. URL: https://www.smartcities.eu/download/smart_cities_final_report.pdf
- 8 Cohen, B. What exactly is a smart city? 2012. [Online]. URL: <https://www.fastcompany.com/1680538/what-exactly-is-a-smart-city>
- 9 Neirotti, P., De Marco, A., Cagliano, A.C., Mangano, G. and Scorrano, F. Current trends in smart city initiatives: Some stylised facts, *Cities*, 38, 25–36 (2014).
- 10 Hollands, R.G. Will the real smart city please stand up? *City*, 12(3), 303–320 (2008).
- 11 Nam, T. and Pardo, T.A. Conceptualizing smart city with dimensions of technology, people, and institutions. *Proceedings of the 12th Annual International Digital Government Research Conference: Digital Government Innovation in Challenging Times*. ACM, 2011, pp. 282– 291.
- 12 Kitchin, R. The real-time city? big data and smart urbanism, *GeoJournal*, 79(1), 1–14 (2014).
- 13 Michalina, D., Mederly, P., Diefenbacher, H., & Held, B. Sustainable Urban Development: A Review of Urban Sustainability Indicator Frameworks. *Sustainability*, 13(16), 9348 (2021).
- 14 Aaron. Indicators for urban sustainability: Key lessons from a systematic analysis of 67 measurement initiatives. *Ecological Indicators*, 119, 106879 (December 2020).
- 15 Simon Elias Bibri. On the sustainability of smart and smarter cities in the era of big data: an interdisciplinary and transdisciplinary literature review. *Journal of Big Data*, 6, 25 (2019).
- 16 Assessing urban sustainability: A proposal for indicators, metrics and scoring – a case study in Colombia. *Environment, Development and Sustainability*, 25, 11789–11822 (2023).
- 17 Development of integrated sustainability performance indicators for better management of smart cities. *Sustainable Cities and Society*, 67, 102704 (2021).
- 18 Measuring progress of smart cities: Indexing the smart city indices. *Urban Governance*, 3(1), 45–57 (March 2023).
- 19 Smart city indexes, criteria, indicators and rankings: An in-depth investigation and analysis. *IET Smart Cities*.
- 20 Frontiers. The concept of smart cities: a sustainability aspect for future urban development based on different cities. Abdullah Addas et al. (2023).
- 21 Zhalmurziyeva, K., Tokbolat, S., Durdyev, S., Mustafa, M.Y. & Karaca, F. Assessment of sustainability indicators for urban water infrastructure in a developing country. *International Journal of Building Pathology and Adaptation* (2024).
- 22 Urban sustainability assessment: An overview and bibliometric analysis. *Ecological Indicators*, 121, 107102 (February 2021).
- 23 Villanueva, F.J., Aguirre, C., Rubio, A., Villa, D., Santofimia, M.J. & López, J.C. Data stream visualization framework for smart cities. *Soft Computing*, 20, 1671–1681 (2016).
- 24 Bovkir, R. & Aydinoglu, A. C. Big Urban Data Visualization Approaches Within the Smart City: GIS-based Open-Source Dashboard Example. *ISPRS Archives*, XLVI-4/W5, 125 (2021).
- 25 Li, Cheng & Li, Junxiang. Assessing Urban Sustainability Using a Multi-Scale, Theme-Based Indicator Framework: A Case Study of the Yangtze River Delta Region, China. *Sustainability*, 9(11), article 2072 (2017).

^{1*}Салим Т.,

магистрант, ORCID ID: 0009-0007-2327-7036,

e-mail: t_salim@kbtu.kz

¹Қуатбаева А.А.,

PhD, ассистент-профессор, ORCID ID: 0000-0002-2143-3994,

e-mail: a.kuatbayeva@iitu.kz

¹Қазақстан-Британ техникалық университеті, Алматы қ., Қазақстан

ҚАЛАНЫҢ ТҰРАҚТЫЛЫҒЫН БАҒАЛАУ: СМАРТ-ҚАЛА ИНДЕКСІ ЖӘНЕ ДЕРЕКТЕРДІ ВИЗУАЛИЗАЦИЯЛАУ

Аңдатпа

Бұл зерттеу қала тұрақтылығын бағалау мен арттырудың маңыздылығы артып келе жатқанын дәлелдейді. Мұның негізгі себебі – урбанизацияның тұрақты үдерісі мен технологиялық прогрестің қарқынды дамуы, олардың нәтижесінде тұрақты смарт-қала құру қажеттілігі туындайды. Осыған байланысты біз смарт-қала индексі (SCI) арқылы қалалардың тұрақтылығын өлшеу және бағалау моделін ұсынамыз және сипаттаймыз. Ұсынылған модель бойынша есептеулер әлемнің 102 қаласын қамтып, жылдам әрі тиімді есептеулер мен визуализацияны қамтамасыз етеді. Модель алты көрсеткішке сүйенеді: мобильдік, қоршаған орта, мемлекеттік басқару, экономика, смарт адамдар және смарт өмір. Бұл тәсіл көрсеткіштер бойынша деректерді талдауға және түсіндіруге мүмкіндік береді, әр қала үшін SCI мәнін есептеп, оның басты артықшылықтары мен жетілдіруді талап ететін бағыттарын айқындайды. Алынған нәтижелер заманауи деректерді визуализациялау әдістері арқылы ұсынылады, бұл SCI-ді түсінікті етеді және қалаларды өзара салыстыруға жол ашады. Зерттеу қорытындылары смарт-қалалар арасында үлкен айырмашылықтардың бар екенін көрсетті, бұл мақсатты саясаттар мен инвестициялар әзірлеудің қажеттілігін айқындайды. Бұл еңбектің үлесі бар білімді толықтырып қана қоймай, тұрақты қалалық ортаны дамытуға қатысушы мүдделі тараптарға қосымша нұсқаулық пен ұсыныстар ұсынады.

Тірек сөздер: қала тұрақтылығы, смарт-қала индексі, деректерді визуализациялау, тұрақты қалалық даму, смарт технологиялар, қалалық аналитика, шешім қабылдауды қолдау жүйелері, тұрақты қаланы жоспарлау.

^{1*}Салим Т.,

магистрант, ORCID ID: 0009-0007-2327-7036,

e-mail: t_salim@kbtu.kz

¹Қуатбаева А.А.,

PhD, ассистент-профессор, ORCID ID: 0000-0002-2143-3994,

e-mail: a.kuatbayeva@iitu.kz

¹Казахстанско-Британский технический университет, г. Алматы, Казахстан

ОЦЕНКА УСТОЙЧИВОСТИ ГОРОДОВ: ИНДЕКС УМНОГО ГОРОДА И ВИЗУАЛИЗАЦИЯ ДАННЫХ

Аннотация

В данном исследовании доказывается, что необходимость оценки и повышения устойчивости городов становится жизненно важной. Это связано с постоянной урбанизацией и технологическим прогрессом, которые требуют устойчивого развития умных городов. На основании изложенного мы предлагаем и описываем модель для измерения и оценки устойчивости города с помощью индекса умного города (SCI). Проведенные расчеты по данной модели охватывают 102 города мира, обеспечивая быстрые и эффективные вычисления и визуализацию данных. Модель основана на шести показателях: мобильность, окружающая среда, государственное управление, экономика, умные люди и умная жизнь. Предложенный подход позволяет анализировать и интерпретировать данные по каждому показателю, что делает возможным расчет SCI

для конкретного города, а также выявление его сильных сторон и направлений для улучшения. Полученные результаты представлены с помощью современных методов визуализации данных, что делает индекс умного города наглядным и позволяет сравнивать города между собой. Выводы исследования показали значительные различия между умными городами, что указывает на необходимость разработки целевых политик и инвестиций. Сделанный вклад, по нашему мнению, будет ценным дополнением к существующим знаниям и предоставит дальнейшие рекомендации для заинтересованных сторон, вовлеченных в развитие устойчивой городской среды.

Ключевые слова: устойчивое развитие города, индекс умного города, визуализация данных, устойчивое городское развитие, умные технологии, аналитика города, системы поддержки принятия решений, планирование устойчивого города.

Article submission date: 01.10.2025

УДК 004.056:004.7
МРНТИ 05.13.19

<https://doi.org/10.55452/1998-6688-2025-22-4-209-218>

¹***Майлыбаев Е.К.,**

PhD, ассоциированный профессор, ORCID ID: 0000-0002-1977-3690,

*e-mail: ersind@mail.ru

²**Сейдалиева У.О.,**

PhD, научный сотрудник, ORCID ID: 0000-0002-7190-6753,

e-mail: useidali@bu.edu

¹Международный транспортно-гуманитарный университет, г. Алматы, Казахстан

²Университет Бостона, г. Бостон, США

ПРОЕКТИРОВАНИЕ ЗАЩИЩЕННОГО СЕТЕВОГО ШЛЮЗА ДЛЯ ОБЛАЧНЫХ ПРИЛОЖЕНИЙ

Аннотация

Статья посвящена проектированию и настройке защищенного сетевого шлюза для облачных приложений на основе современных VPN-протоколов OpenVPN и WireGuard. В условиях активного развития облачных технологий и роста числа кибератак обеспечение защищенного удаленного доступа к сервисам становится ключевой задачей информационной безопасности. Рассматриваются актуальные угрозы, возникающие при передаче данных в облачных средах, и показана роль VPN-технологий в предотвращении атак. Подробно анализируются особенности OpenVPN и WireGuard: архитектура, криптографическая база, удобство настройки и производительность. В работе представлена архитектура шлюза, включающая сервер VPN, межсетевые фильтры и механизмы маршрутизации, обеспечивающие принудительное прохождение всего трафика через зашифрованный туннель. Эксперименты, проведенные в виртуализированной среде VMware Workstation, показали, что WireGuard демонстрирует более высокую скорость передачи данных и низкую задержку, тогда как OpenVPN отличается гибкостью и совместимостью с корпоративными системами. Совместное использование обоих протоколов позволяет повысить отказоустойчивость и адаптивность системы. Практическая значимость исследования заключается в возможности внедрения предложенной архитектуры в корпоративных и частных сетях для защиты облачных приложений, организации удаленного доступа сотрудников и повышения уровня безопасности информационных ресурсов.

Ключевые слова: облачные приложения, информационная безопасность, защищенный шлюз, шифрование, сетевые протоколы.

Введение

Современные облачные технологии предоставляют широкие возможности для хранения, обработки и передачи данных. Однако их использование сопряжено с рисками: перехват трафика, несанкционированный доступ, атаки типа «человек посередине». В этих условиях VPN (Virtual Private Network) остается одним из наиболее эффективных инструментов защиты коммуникаций [1].

Использование полной шифрации пакетов в VPN обеспечивает устойчивость к обнаружению и блокировкам, что подтверждает актуальность VPN как средства защиты коммуникаций [2]. OpenVPN и WireGuard являются двумя наиболее распространенными решениями в данной области [3, 4]. Первый отличается зрелостью и гибкостью, второй – высокой скоростью и минималистичной архитектурой. Ряд исследований подтверждает эти различия. Так, Jason и Donenfeld показали, что WireGuard превосходит традиционные VPN по пропускной способности и задержкам [5], а работы Joel Anyam и соавторов подчеркивают широкую применимость OpenVPN в корпоративных инфраструктурах благодаря богатой поддержке криптографических алгоритмов [6].

Несмотря на популярность обоих протоколов, недостаточно исследований, комплексно рассматривающих их совместное применение в архитектуре защищенного шлюза для облачных приложений. Цель статьи – восполнить данный пробел и предложить практическую модель шлюза, обеспечивающего баланс между безопасностью и производительностью.

Несмотря на то что основное внимание уделено построению защищенного шлюза для взаимодействия с облачными приложениями, реализация производилась в виртуализированной среде VMware Workstation. Это обусловлено ограничениями по бюджету и отсутствием доступа к коммерческим облачным платформам. Тем не менее вся архитектура решения и используемые протоколы сохраняют полную переносимость на облачные среды, такие как Microsoft Azure, Amazon EC2, DigitalOcean и др. Принципы построения VPN-туннеля, маршрутизации и фильтрации трафика остаются актуальными вне зависимости от среды развертывания.

Материалы и методы

Рост интереса к построению защищенных сетевых шлюзов в современных условиях напрямую связан с увеличением числа пользователей, подключающихся к облачным сервисам из нестабильных или потенциально уязвимых сетей, будь то домашние подключения, публичные Wi-Fi сети или мобильный Интернет. В рамках данной работы ставится задача разработки практического решения, которое обеспечит безопасный обмен данными при взаимодействии с облачными приложениями, используя актуальные технологии VPN на базе протоколов OpenVPN и WireGuard.

В начальной фазе исследование ориентировано на определение характерных уязвимостей, связанных с передачей данных в облачной среде. При этом особое внимание уделяется сценариям, в которых конечные устройства пользователей подключаются к облачным платформам через ненадежные каналы связи, что делает необходимым внедрение дополнительного слоя защиты.

Следующий этап предполагает выбор технических решений, способных снизить указанные риски. В качестве приоритетных рассматриваются VPN-протоколы, предлагающие защищенный способ взаимодействия с удаленными ресурсами. В данном контексте OpenVPN выступает как проверенное временем решение с гибкой настройкой, тогда как WireGuard рассматривается как инновационная альтернатива, ориентированная на производительность и упрощение архитектуры. Оценка обоих протоколов проводится с учетом условий их интеграции в разнородные информационно-технические среды, а также способности противостоять распространенным угрозам.

Однако исследование не сводится исключительно к выбору программного обеспечения. Одной из важнейших задач становится формирование архитектурной модели шлюза, включающей маршрутизацию, механизмы шифрования, контроль трафика и способы аутентификации.

Исследование проводится в изолированной среде на базе виртуальных машин, что дает возможность моделировать реальные условия эксплуатации и воссоздавать потенциально нестандартные ситуации: обрыв соединения, попытку несанкционированного доступа, пиковую нагрузку. Это позволяет объективно проверить, насколько предложенное решение соответствует требованиям стабильности и защищенности.

Исследование строится не только на изучении теоретических основ. Его важной составляющей является практическая реализация и проверка работоспособности разрабатываемого шлюза. С одной стороны, проводится аналитический обзор литературных и нормативных источников, касающихся VPN, криптографических методов и архитектур безопасности. С другой – выполняется пошаговая настройка и отладка шлюза, включающая использование командной строки, работу с конфигурационными файлами, анализ логов и устранение ошибок. Такой подход позволяет получить всестороннее понимание не только логики работы системы, но и особенностей ее внедрения.

Различия между протоколами особенно заметны на этапе внедрения в реальные сети. OpenVPN отлично подходит для компаний с развитой инфраструктурой, нуждающихся в продвинутой маршрутизации, дифференцированных правах доступа и подключении к внутренним сервисам – LDAP, Active Directory и централизованным системам логирования. Он дает возможность гибкой настройки политик, включая ограничение по IP-диапазонам, времени активности и уровню доступа конкретных клиентов, что делает его мощным инструментом для администрирования.

WireGuard, напротив, оптимизирован под задачи, в которых ключевыми являются скорость и простота использования. Он показывает высокие показатели производительности даже на слабом оборудовании: одноплатных ПК, роутерах и IoT-устройствах. Подключения устанавливаются мгновенно, восстанавливаются автоматически при смене сетей и не требуют ручного вмешательства для обновления настроек. Эти свойства делают WireGuard привлекательным для мобильных решений, DevOps-инфраструктур и проектов, где ресурсы ограничены, а стабильность критична.

Для удобства восприятия ключевых различий между OpenVPN и WireGuard их характеристики представлены в таблице 1.

Таблица 1 – Сравнение характеристик OpenVPN и WireGuard

Параметр	OpenVPN	WireGuard
Архитектура	Модульная, основана на TLS	Минималистичная, встроена в ядро Linux
Шифрование	AES-256, Blowfish, CAST5 и др.	ChaCha20, Curve25519 (фиксированные)
Настройка	Гибкая, но сложная	Простая, единый конфигурационный файл
Производительность	Средняя (зависит от параметров)	Высокая, особенно на Linux-платформах
Поддержка платформ	Широкая, включая Windows, MacOS, Linux	Ограничена на старых ОС, лучше всего работает на Linux
Совместимость с корпоративными политиками	Высокая	Требует адаптации
Применимость	Корпоративные сети, ЦОДы, сложная маршрутизация	Личные VPN, DevOps, мобильные устройства

В рамках статьи оба протокола рассматриваются как взаимодополняющие, а не конкурирующие решения. Их параллельное использование позволяет реализовать адаптивную архитектуру, в которой конечный пользователь самостоятельно выбирает наиболее подходящий вариант подключения. Такой подход расширяет возможности конфигурации, увеличивает гибкость системы и дает возможность более точного анализа эффективности каждого из протоколов в условиях реальной эксплуатации.

Результаты и обсуждение

Функциональная схема работы защищенного сетевого шлюза формируется вокруг четко организованного взаимодействия между центральной серверной частью и множеством клиентских устройств, объединенных в рамках логически замкнутой инфраструктуры. Основная задача этой системы заключается в создании устойчивого, безопасного и управляемого канала передачи данных, по которому весь пользовательский сетевой трафик проходит в зашифрованном виде. Центральным элементом в данной архитектуре выступает VPN-сервер, на котором реализуются функции шифрования, маршрутизации, управления сессиями и контроля по-

литик доступа. Визуальная структура взаимодействия всех компонентов шлюза представлена на рисунке 1.



Рисунок 1 – Общая схема работы защищенного сетевого шлюза

В результате формируется схема, в которой каждый этап – от подключения до контроля трафика – охватывается специализированным набором механизмов. Это позволяет выстраивать надежную и в то же время гибкую инфраструктуру, адаптируемую под различные сценарии эксплуатации. Шлюз не просто обеспечивает защищенную коммуникацию, но и формирует основу для построения более сложных сетевых решений, сохраняя при этом прозрачность конфигурации и управляемость.

Для реализации защищенного туннеля между клиентом и сервером необходимо предварительно установить и настроить OpenVPN.

Была выполнена полная настройка серверной части OpenVPN с использованием Easy-RSA 3.x. Реализованы все ключевые этапы криптографической подготовки: инициализация центра сертификации, выпуск серверного сертификата, генерация параметров шифрования и создание ключей безопасности. Также подробно рассмотрены процессы формирования и подписи клиентского сертификата, что обеспечило готовность инфраструктуры к безопасному взаимодействию между сервером и клиентами через VPN-туннель.

На рисунке 2 отображается подтверждение успешного подключения и факт маршрутизации трафика через установленное VPN-соединение. Это свидетельствует о корректной настройке и работоспособности всей криптографической инфраструктуры, развернутой с использованием Easy-RSA 3.x как на серверной, так и на клиентской стороне.

```

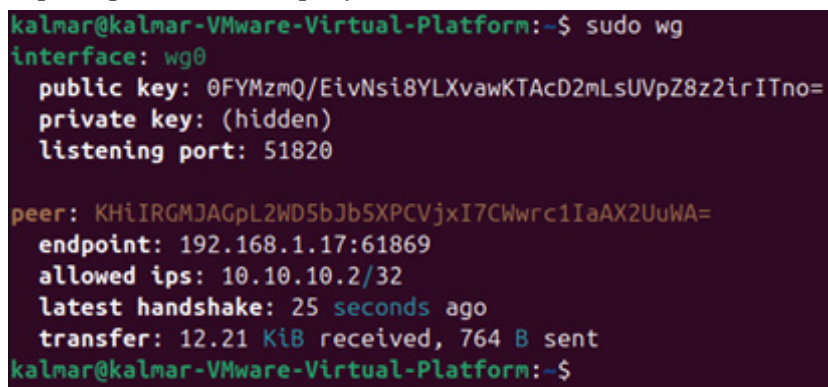
2025-06-01 11:13:32 net_iface_up: set tun0 up
2025-06-01 11:13:32 net_addr_ptp_v4_add: 10.8.0.10 peer 10.8.0.9 dev tun0
2025-06-01 11:13:32 net_route_v4_add: 192.168.1.50/32 via 192.168.91.2 dev [NULL] table 0 metric -1
2025-06-01 11:13:32 net_route_v4_add: 0.0.0.0/1 via 10.8.0.9 dev [NULL] table 0 metric -1
2025-06-01 11:13:32 net_route_v4_add: 128.0.0.0/1 via 10.8.0.9 dev [NULL] table 0 metric -1
2025-06-01 11:13:32 net_route_v4_add: 10.8.0.1/32 via 10.8.0.9 dev [NULL] table 0 metric -1
2025-06-01 11:13:32 Initialization Sequence Completed
2025-06-01 11:13:32 Data Channel: cipher 'AES-256-GCM', peer-id: 0
2025-06-01 11:13:32 Timers: ping 10, ping-restart 120
2025-06-01 11:13:32 Protocol options: protocol-flags cc-exit tls-ekm dyn-tls-crypt

```

Рисунок 2 – Запуск и проверка состояния сервиса OpenVPN

Для реализации второго VPN-протокола в рамках защищенного шлюза используется WireGuard. Он отличается компактной архитектурой, высокой производительностью и простотой конфигурации. В отличие от OpenVPN, данный протокол не требует развертывания центра сертификации, а использует статические ключи, что делает его особенно удобным в условиях ограниченных вычислительных ресурсов.

Завершающим этапом настройки VPN-сервера является активация механизма переадресации IP-трафика, необходимого для обеспечения маршрутизации клиентских данных через защищенный туннель. Этот параметр позволяет передавать пакеты от подключенных клиентов во внешние сети, действуя как маршрутизатор. Включение данной функции выполняется путем изменения системной сетевой конфигурации, отвечающей за пересылку IPv4-пакетов между интерфейсами. Проверка активности соединения производится командой «sudo wg show», результат которой представлен на рисунке 3.



```
kalmar@kalmar-VMware-Virtual-Platform:~$ sudo wg
interface: wg0
  public key: 0FYMzmQ/EivNsi8YLXvawKTAcd2mLsUVpZ8z2irITno=
  private key: (hidden)
  listening port: 51820

peer: KH1IRGMJAGpL2WD5bJb5XPCVjxI7CWwrc1IaAX2UuWA=
  endpoint: 192.168.1.17:61869
  allowed ips: 10.10.10.2/32
  latest handshake: 25 seconds ago
  transfer: 12.21 KiB received, 764 B sent
kalmar@kalmar-VMware-Virtual-Platform:~$
```

Рисунок 3 – Проверка состояния соединения WireGuard

Реализация механизма сетевой фильтрации в рамках защищенного шлюза осуществляется с использованием UFW (Uncomplicated Firewall) – утилиты, предоставляющей упрощенный интерфейс для управления встроенным в систему инструментом iptables. Данный подход позволяет значительно упростить настройку базовых политик межсетевого экранирования, не снижая при этом уровня безопасности. В соответствии с принятой моделью «запрет по умолчанию» все входящие подключения блокируются, за исключением явно разрешенных правил. В данной конфигурации открываются только те порты, которые необходимы для функционирования туннельных протоколов OpenVPN и WireGuard. Это обеспечивает ограничение внешнего доступа к инфраструктуре, минимизируя возможные точки входа.

Помимо портовой фильтрации, важным аспектом является принудительная маршрутизация всего клиентского трафика исключительно через VPN-туннель. Такая мера позволяет полностью исключить вероятность обхода защищенной инфраструктуры и утечки данных в открытый Интернет. Для пользователей OpenVPN данная функция активируется с помощью параметров конфигурации, перенаправляющих весь трафик через виртуальный интерфейс сервера. Аналогичный эффект достигается в WireGuard путем указания маршрутов в параметре, который задает полный перечень IP-адресов, доступных только через шлюз. В результате весь пользовательский трафик направляется строго через VPN, обеспечивая централизованную фильтрацию, контроль и защиту. Схематическое представление архитектуры маршрутизации и фильтрации приведено на рисунке 4.

Хотя практическая реализация осуществлялась в локальной лабораторной среде, полученные результаты и структура решения позволяют утверждать, что оно может быть адаптировано для использования в публичных или частных облаках. Это делает предложенный подход универсальным и применимым в реальных условиях эксплуатации защищенных удаленных подключений.

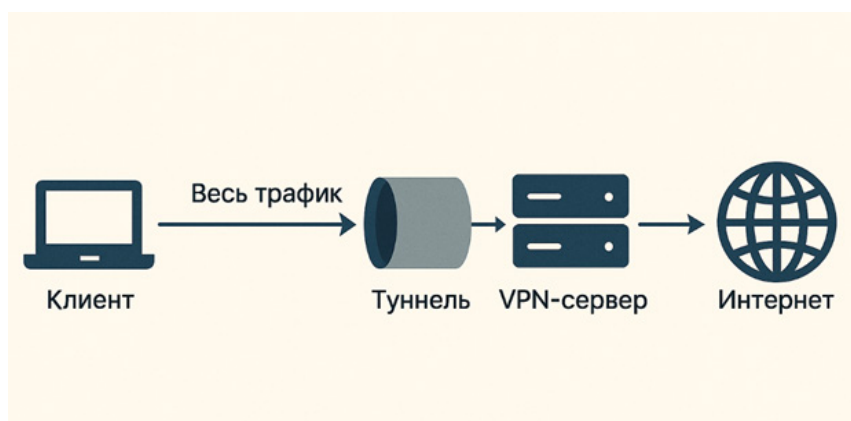


Рисунок 4 – Схема принудительной маршрутизации
всего трафика клиента через VPN

Таким образом, формируется логически замкнутая схема, в которой любые сетевые взаимодействия клиента происходят исключительно через контролируемый шлюз. Сам шлюз, в свою очередь, осуществляет мониторинг и управление всем трафиком, как входящим, так и исходящим. Такой подход особенно актуален в условиях использования публичных или ненадежных сетей, где критично обеспечить полное доверие только к одному – туннелированному каналу доступа.

Для повышения наглядности и ускорения процессов настройки в ходе тестирования использовались графические пользовательские интерфейсы (GUI), доступные в клиентских и серверных решениях для обеих технологий. Тестовая среда была развернута на виртуальных машинах, функционирующих под управлением операционных систем Ubuntu Desktop и Kali Linux. Виртуальные машины были соединены через комбинированную схему bridge-сетей, что позволило смоделировать как локальные, так и внешние подключения. Такая конфигурация обеспечивала достаточный уровень реализма для оценки поведения VPN в условиях, приближенных к реальной эксплуатации.

Сравнительные характеристики протокола WireGuard часто анализируются в сопоставлении с OpenVPN. Эти наблюдения особенно важны для оценки применимости технологии в высоконагруженных распределенных системах [7].

На основании результатов, полученных в ходе экспериментальной проверки работы VPN-шлюзов, выполнено сопоставление характеристик двух исследуемых протоколов – OpenVPN и WireGuard. В сравнении учитывались как количественные показатели (задержка, потери пакетов), так и качественные характеристики, фиксируемые визуально или через команды мониторинга.

Значения средней задержки при передаче пакетов в туннеле WireGuard оказались ниже и составили около 0.5 мс, в то время как для OpenVPN этот показатель находился в пределах 1 мс. Максимальная задержка у WireGuard также была ниже (1 мс против 2 мс у OpenVPN), что может свидетельствовать о более эффективной обработке трафика внутри ядра операционной системы. Потери пакетов при работе обоих протоколов отсутствовали, что демонстрирует их стабильность в контролируемой среде. Результаты эксперимента указаны в таблице 2.

Скорость установления соединения в WireGuard визуально выше: подключение происходило практически мгновенно, тогда как OpenVPN демонстрировал небольшую задержку при запуске. Мониторинг трафика с помощью утилиты nload показал наличие активности в обоих случаях, что подтверждает работоспособность туннелей. TTL в ответах на ICMP-запросы в обоих случаях составлял 64, что указывает на одноступенчатую маршрутизацию без промежуточных узлов.

Таблица 2 – Сравнительные характеристики OpenVPN и WireGuard по результатам тестирования

Метрика	OpenVPN	WireGuard
Средняя задержка (мс)	1	0.5
Максимальная задержка (мс)	2	1
Потери пакетов (%)	0	0
Загрузка CPU (%)	7	2.5
Криптоустойчивость (оценка)	4.5	5

Нагрузка на центральный процессор при функционировании VPN-соединения замерялась при помощи команды `top`. В процессе наблюдений средняя загрузка CPU при работе OpenVPN составила порядка 6–8%, в то время как у WireGuard данный показатель оказался значительно ниже и не превышал 2–3%. Это может быть связано с тем, что WireGuard реализован в пространстве ядра и обладает более компактной архитектурой.

Дополнительный критерий анализа – уровень криптографической защищенности. Протокол OpenVPN применяет такие проверенные алгоритмы, как AES-256, TLS, и инфраструктуру открытых ключей (PKI), что обеспечивает высокий уровень криптоустойчивости. Однако высокая гибкость настройки сопровождается риском ошибок конфигурации. В отличие от него, WireGuard использует современный стек алгоритмов: ChaCha20, Curve25519, Poly1305 и BLAKE2s, реализуемых в виде минималистичной схемы. Такая структура снижает вероятность конфигурационных уязвимостей и повышает общее доверие к безопасности соединения.

Для наглядного представления различий между протоколами по числовым показателям – задержке, загрузке CPU и криптоустойчивости – построена сравнительная диаграмма (рисунок 5).

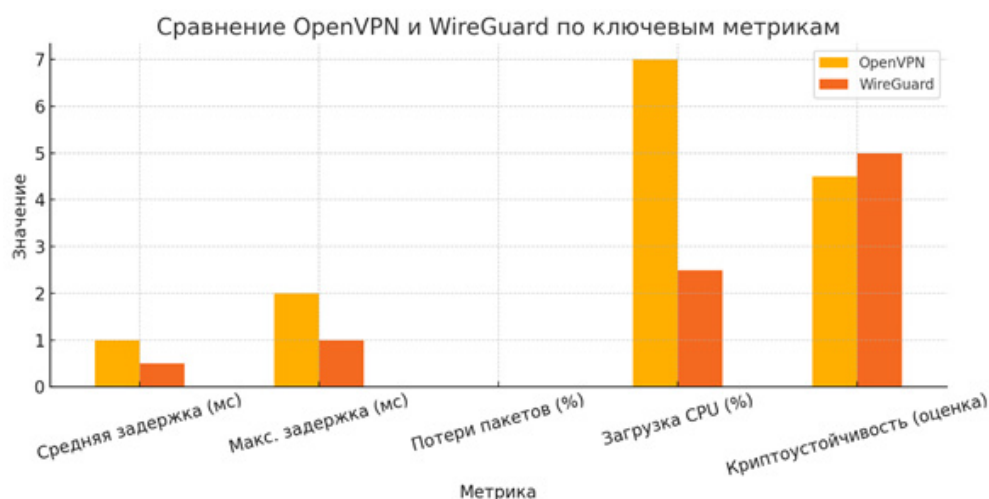


Рисунок 5 – Диаграмма различий между OpenVPN и WireGuard по ключевым метрикам

На основании совокупности вышеуказанных наблюдений можно сделать вывод, что протокол WireGuard демонстрирует лучшие результаты по ряду ключевых показателей, включая контрольный трафик, стабильность и криптографическую надежность, при этом оставаясь более простым в конфигурации и эксплуатации.

Таким образом, несмотря на идентичный функционал – создание защищенного туннеля, протоколы OpenVPN и WireGuard показали различия в скорости установления соединения, нагрузке на систему и сложности конфигурации. В условиях ограниченной инфраструктуры и необходимости быстрого развертывания предпочтительным решением можно считать WireGuard, однако выбор должен основываться на политике безопасности организации и инфраструктурных возможностях.

Заключение

В ходе исследования были выполнены проектирование, реализация и комплексная настройка защищенного сетевого шлюза на основе протоколов OpenVPN и WireGuard, а также проведена всесторонняя проверка его работоспособности и безопасности. Достигнуты результаты, подтверждающие эффективность предложенного решения в условиях ограниченных вычислительных ресурсов и повышенных требований к защите трафика.

Были проанализированы существующие подходы к построению VPN-сетей, изучены особенности архитектуры OpenVPN и WireGuard, а также проведен сравнительный обзор их функциональных возможностей. Основное внимание было уделено характеристикам безопасности, способам аутентификации, принципам маршрутизации и механизму защиты от атак.

Развернут шлюз в виртуальной среде с последовательной настройкой каждого из VPN-сервисов, внедрение механизма сетевой фильтрации и реализации политики «только через туннель». Проведенные тесты позволили выявить различия в производительности, времени подключения, отклике системы на сбои и потреблении ресурсов. WireGuard продемонстрировал лучшую эффективность по всем ключевым метрикам, в то время как OpenVPN обеспечил более широкие возможности управления и интеграции.

Безопасность предложенного решения была подтверждена серией испытаний, в ходе которых проверялась корректность изоляции, туннелирования трафика, устойчивость к несанкционированным подключениям и сбоям. Все тесты показали соответствие реализованной схемы базовым принципам защищенного сетевого взаимодействия. В частности, была достигнута полная недоступность внутренних ресурсов вне VPN-туннеля, а также устойчивость к подмене ключей, отказам соединения и попыткам обхода фильтрации.

Таким образом, исследование имеет прикладной характер и может быть адаптировано для использования в реальных условиях – как в учебных лабораториях, так и в корпоративных средах, где требуется построение надежного канала удаленного доступа с минимальной нагрузкой на инфраструктуру. Полученные результаты подтверждают целесообразность применения предложенного подхода при проектировании безопасных коммуникационных решений в облачной среде.

ЛИТЕРАТУРА

- 1 Mahmood, Z. Virtual Private Networks: Fundamentals, security issues and solutions, pp. 1–7 (2023). <https://doi.org/10.20944/preprints202306.1105.v1>.
- 2 Parker, A. Efficacy of Full-Packet Encryption in Mitigating Protocol Detection for Evasive Virtual Private Networks, pp. 1–6 (2024) <https://doi.org/10.48550/arXiv.2412.17352>.
- 3 Xue, D., Ramesh R., Jain A., Kallitsis M., Halderman J., Crandall J., Ensafi R. Roy. OpenVPN is Open to VPN Fingerprinting. Communications of the ACM, pp. 79–87 (2024). <https://doi.org/10.1145/3618117>.
- 4 Farooq, I., Ahmed Syed., Ali, A., Warraich, M., Aqeel, M., Khan, H. Enhanced Classification of Networks Encrypted Traffic: A Conceptual Analysis of Security Assessments, Implementation, Trends and Future Directions. The Asian Bulletin of Big Data Management, pp. 500–522 (2024). <https://doi.org/10.62019/abdbm.v4i4.287>.
- 5 Donenfeld, J. WireGuard: Next Generation Kernel Network Tunnel. NDSS Symposium, pp. 1–12 (2017). <https://doi.org/10.14722/ndss.2017.23160>.

6 Joel, A., Rajiv, S., Hadi, L., Anand, P. Empirical Performance Analysis of WireGuard vs. OpenVPN in Cloud and Virtualised Environments Under Simulated Network Conditions. Computers., pp. 1–52 (2025). <https://doi.org/10.3390/computers14080326>.

7 Mackey, S., Mihov, I., Nosenko, A., Vega, F., Cheng, Y. A Performance Comparison of WireGuard and OpenVPN., pp. 162–164 (2020). <https://doi.org/10.1145/3374664.3379532>.

¹*Майлыбаев Е.Қ.,

PhD, қауымдастырылған профессор, ORCID ID: 0000-0002-1977-3690,

*e-mail: ersind@mail.ru

²Сейдалиева Ұ.Ө.,

PhD, ғылыми қызметкер, ORCID ID: 0000-0002-7190-6753,

e-mail: useidali@bu.edu

¹Халықаралық көліктік-гуманитарлық университет, Алматы қ., Қазақстан

²Бостон университеті, Бостон, АҚШ

БҰЛТТЫҚ ҚОЛДАНБАЛАР ҮШІН ҚАУІПСІЗ ЖЕЛІ ШЛЮЗИН ЖОБАЛАУ

Аңдатпа

Мақала заманауи VPN-хаттамалары OpenVPN және WireGuard негізінде бұлтты қосымшаларға арналған қорғалған желілік шлюзді жобалауға және баптауға арналған. Бұлтты технологиялардың белсенді дамуы мен кибершабуылдардың көбеюі жағдайында қызметтерге қорғалған қашықтан қолжетімділікті қамтамасыз ету ақпараттық қауіпсіздіктің негізгі міндетіне айналууда. Бұлтты ортада деректерді беру кезінде туындайтын өзекті қатерлер қарастырылып, шабуылдардың алдын алуда VPN-технологиялардың рөлі ұсынылған. OpenVPN және WireGuard ерекшеліктері: архитектурасы, криптографиялық негізі, ыңғайлылығы мен өнімділігі жан-жақты талданған. Жұмыста VPN-серверді, брандмауэр сүзгілерін және барлық трафиктің шифрланған туннель арқылы міндетті өтуін қамтамасыз ететін маршрутизация механизмдерін қамтитын шлюз архитектурасы ұсынылған. VMware Workstation виртуалды ортасында жүргізілген тәжірибелер WireGuard деректерді беруде жоғары жылдамдық пен төмен кідірісті қамтамасыз ететінін, ал OpenVPN икемділігімен және корпоративтік жүйелермен үйлесімділігімен ерекшеленетіні көрсетілген. Екі хаттаманы бірлесіп пайдалану жүйенің сенімділігін және бейімделгіштігін арттыруға мүмкіндік береді. Зерттеудің практикалық маңыздылығы ұсынылған архитектураны корпоративтік және жеке желілерге енгізу мүмкіндігінде, бұлтты қосымшаларды қорғау, қызметкерлердің қашықтан қолжетімділігін ұйымдастыру және ақпараттық ресурстардың қауіпсіздік деңгейін арттыруда көрініс табады.

Тірек сөздер: бұлтты қосымшалар, ақпараттық қауіпсіздік, қорғалған шлюз, шифрлау, желілік хаттамалар.

¹*Mailybayev Y.,

PhD, Associate Professor, ORCID ID: 0000-0002-1977-3690,

*e-mail: ersind@mail.ru

²Seidaliyeva U.,

PhD, Research Associate, ORCID ID: 0000-0002-7190-6753

*e-mail: useidali@bu.edu

¹International University of Transport and Humanities, Almaty, Kazakhstan

²Boston University, Boston, USA

DESIGN OF A SECURE NETWORK GATEWAY FOR CLOUD APPLICATIONS

Abstract

The article is devoted to the design and configuration of a secure network gateway for cloud applications based on modern VPN protocols OpenVPN and WireGuard. In the context of the rapid development of cloud

technologies and the increasing number of cyberattacks, ensuring secure remote access to services has become a key task of information security. The paper discusses relevant threats arising during data transmission in cloud environments and highlights the role of VPN technologies in preventing attacks. The features of OpenVPN and WireGuard are analyzed in detail, including their architecture, cryptographic foundation, ease of configuration, and performance. The study presents a gateway architecture comprising a VPN server, firewall filters, and routing mechanisms that enforce mandatory transmission of all traffic through an encrypted tunnel. Experiments conducted in a virtualized VMware Workstation environment showed that WireGuard provides higher data transfer speeds and lower latency, while OpenVPN demonstrates flexibility and compatibility with corporate systems. The combined use of both protocols improves system resilience and adaptability. The practical significance of the research lies in the possibility of implementing the proposed architecture in corporate and private networks to protect cloud applications, organize secure remote employee access, and enhance the security level of information resources.

Keywords: cloud applications, information security, secure gateway, encryption, network protocols.

Дата поступления статьи в редакцию: 21.08.2025

UDC 004.032.26.
MRSTI 28.23.15.

<https://doi.org/10.55452/1998-6688-2025-22-4-219-226>

¹**Semenyuk V.V.,**

M.Tech.Sc., ORCID ID: 0000-0002-8580-7326,

e-mail: vvsemenyuk@ku.edu.kz

^{1*}**Kurmashev I.G.,**

PhD, ORCID ID: 0000-0001-9872-7483,

*e-mail: ikurmashev@ku.edu.kz;

²**Serbin V.V.,**

PhD, Associate Professor, ORCID ID: 0000-0002-5807-3873,

e-mail: v.serbin@satbayev.university

¹**Kurmasheva L.B.,**

MSc, ORCID ID: 0000-0001-5392-5873,

e-mail: lbkurmasheva@ku.edu.kz

²**Moldagulova A.N.,**

PhD, Professor, ORCID ID: 0000-0002-1596-561X,

e-mail: a.moldagulova@satbayev.university

¹M. Kozybaev North Kazakhstan University, Petropavlovsk, Kazakhstan

²Satbayev University, Almaty, Kazakhstan.

DEVELOPMENT OF A REAL-TIME UAV RECOGNITION MODEL BASED ON YOLOV10 NEURAL NETWORK

Abstract

The paper deals with the development of a model for real-time recognition and classification of UAVs and birds based on the training of the YOLOv10 neural network. The research area is considered relevant in connection with the problems of UAV detection in the context of security, given their growing use in various fields. A dataset consisting of 6,255 images collected from proprietary archives and public resources is trained to train the model. The process of data annotation, augmentation and distribution was implemented using Roboflow.com service. The model was trained on NVIDIA GeForce RTX 4080 GPU using Ultralytics framework. Test results showed high recognition accuracy with mAP50 and mAP50-95 metrics exceeding previous versions of YOLO. The model demonstrates the ability for efficient object segmentation and tracking, which makes it promising for optoelectronic surveillance applications. The results of the study can be useful for developers of UAV and bird detection and classification systems, as well as for improving safety in various fields.

Keywords: neural networks, classification, recognition, optoelectronic surveillance channels, UAVs, YOLO, convolutional neural networks.

Introduction

Real-time recognition of unmanned aerial vehicles (UAVs) through surveillance cameras using neural networks and visual detectors is an urgent area for scientific research. This is driven by the need for public and state security, as UAVs are now being actively used for illegal activities including espionage, smuggling delivery, and combat raids. UAV recognition detection systems are designed to detect such threats in a timely manner and take appropriate measures to neutralize them. They also play an important role in protecting critical infrastructure such as airports, power plants and government buildings where there is a high risk of potential attacks.

The scientific novelty of the study lies in applying the YOLOv10 architecture – featuring dual label assignment, consistent matching metrics, and NMS-free design – to the task of real-time UAV and bird recognition. The work demonstrates improved mAP50-95 accuracy over previous YOLO versions on a custom dataset and shows the model's capability for efficient segmentation and tracking.

The originality stems from adapting and validating the newly introduced YOLOv10 mechanisms specifically for small, fast-moving airborne objects in optoelectronic surveillance systems.

Real-time drone recognition helps enforce flight regulations and control restricted areas. This is especially important when there are strict requirements and restrictions on the use of UAVs in certain areas, such as military bases and private territories. Modern neural networks and machine learning algorithms have the ability to process large amounts of data, which provides high accuracy recognition and opens new opportunities for developing effective monitoring systems. Such systems can be integrated with existing video surveillance systems, making them more affordable and efficient. The algorithms used as software modules for UAV detection and recognition in optoelectronic surveillance channels are YOLO, Faster R-CNN, and SSD. The authors found that single-stage YOLO detectors and two-stage Faster R-CNN algorithms compared to SSD provide better accuracy of UAV recognition, while the latter algorithm is also effective in target detection tasks [1]. However, according to the research results presented trained YOLO models outperform Faster RCNN in terms of accuracy metrics and speed (frames per second – FPS), which makes the former algorithm more promising for real-time UAV and bird recognition and classification tasks [2–4]. Considering the recent advances in the improvement of UAV recognition models based on the YOLO algorithm, we should mention the works [2–9]. The authors prepared a dataset in the form of UAV images for training the YOLOv2 neural network [2]. As a result of training, the model reached the maximum value of mAP50 (average accuracy at the threshold of intersection of bounding boxes 50%) – 0.75. The authors studied the features of UAV recognition by YOLOv3 neural network models [3]. By evaluating the qualitative performance of the trained experimental model, it is proposed to introduce densely connected modules to improve the inter-layer connectivity of convolutional neural networks, thereby improving the accuracy and FPS. As a result, the proposed model demonstrated mAP50-95 values of 0.36 and 60 FPS, which outperform the original YOLOv3 model by 0.03 and 24.45 FPS, respectively. The mAP values presented are not equivalent, as mAP50-95 defines the average accuracy at the 95% bounding box crossing threshold, which implies comparatively lower values compared to mAP50 [2, 3]. A better model of the YOLO architecture, YOLOv4, is used where the authors achieved mAP50 values of 0.75, which is consistent with the results of [2, 4]. However, the YOLOv4 models can be considered preferable to the YOLOv2 algorithm primarily due to its higher FPS, which is important for recognizing small objects such as UAVs in real time. YOLOv5 was used as a pre-trained neural network, which is currently considered to be the most popular model of the YOLO architecture for solving the problems of recognizing and classifying any objects from a user's dataset, with training and inference available on the CPU [5, 6]. The neural network from the first paper achieved a mAP50 value of 0.947, while the authors from trained the model to a mAP50 value of 0.904 [5–6]. The paper [7] presents the results of training more advanced versions of YOLO:

YOLOv5: mAP50 - 0.912;

YOLOX: mAP50 - 0.887;

YOLOv7: mAP50 - 0.524;

YOLOv7-tiny: mAP50 - 0.85;

YOLOv8 - 0.953.

However, the authors did not provide the performance of the mAP50-95 metric, which can provide more reliable information about the quality of UAV recognition and detection [7]. In 2023, Ultralytics presented an improved algorithm, YOLOv9 [8]. By incorporating programmable gradient information (PGI) and an efficient layer aggregation network into the architecture, this model will limit the loss of information in successive layers of a deep neural network. YOLOv9 like previous versions utilizes a non-maximal suppression (NMS) algorithm in the post-processing stage, designed to remove unnecessary object bounding boxes. Its use often discards useful predictions and increases computational cost, which is detrimental to the detection and recognition performance of small and maneuverable objects such as UAVs and birds. In this paper, a YOLOv10 model will be trained, which by incorporating new post-processing techniques from the DETR architecture, is hypothesized to increase the recognition and classification accuracy of UAVs and birds [9]. As an analytical accuracy parameter that guarantees the effectiveness of the developed model, mAP50-95 is chosen.

Materials and methods

A more advanced model for real-time recognition and classification of UAVs and birds is based on the YOLOv10 algorithm. YOLOv10 is a neural network based on two key principles: training the model without NMPs and designing based on efficiency and accuracy. The first concept is provided by integrating the advantages of neural networks of the DETR architecture, namely the use of dual label assignment and a consistent metric for matching predictions (Figure 1).

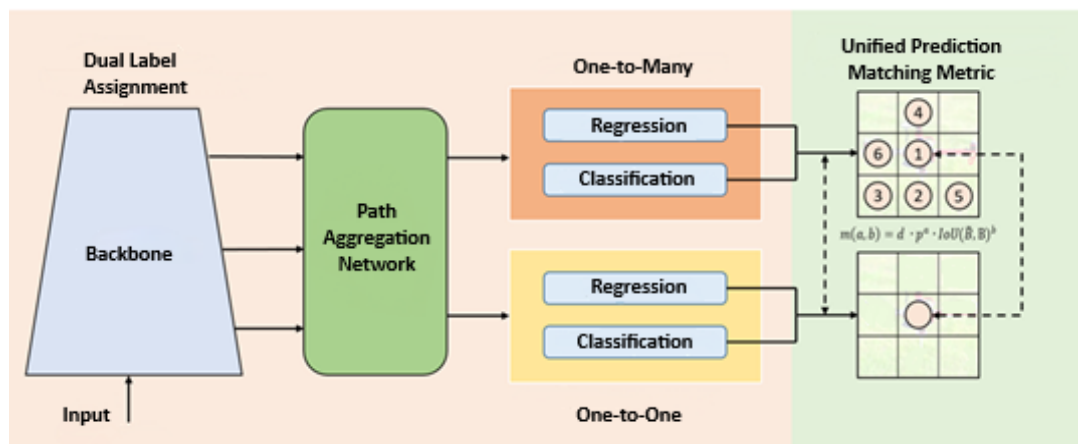


Figure 1 – Concept of the YOLOv10 model

The dual purpose of labels is defined by the architectural solution of using two Heads, where the first Head performs One-to-One matching (provides only one prediction to each benchmark from the user dataset without the use of NMPs) and the second Head performs One-to-Many matching. By utilizing the One-to-Many branch, comprehensive and tight control is provided to ensure that accuracy is maintained. Throughout the training phase, both Heads are used simultaneously, except for the Inference (prediction) phase. At this stage, only the One-to-One branch is used, this decision is made to reduce computational cost while maintaining the accuracy and efficiency of the model. To guarantee the coincidence of the forecasts of the two branches, a consistent metric (or matching metric) is used. The principle of operation of this metric is described by equation (1):

$$m(a, b) = d \cdot p^a \cdot \text{IoU}(\tilde{B}, B)^b \quad (1)$$

where m is the corresponding accuracy metric (“One to One” or “One to Many”), d is the degree of localization of the prediction reference point within the instance, p is the classification score, a and b are the parameters defining the classification and localization tasks, respectively, IoU is the parameter defining the area of overlap between the predicted and actual frames, is the bounding boxes of the prediction and the instance, respectively. Matching the metrics of both branches ensures that the best samples are matched, thus comprehensive learning control is realized. Designing a performance-based model involves facilitating the classification channel by using two depth-separable convolutions followed by a 1×1 convolution (point convolution). The point convolution allows increasing the number of channels, while the depth convolution reduces the spatial dimensions, such a solution allows reducing the computational cost, as well as preserving more useful information in the downsampling stage. The accuracy orientation of the model is ensured by increasing the dimensionality of kernels in deep convolution layers, as well as by using the partial self-awareness module (PSM).

YOLOv10 supports 6 scales depending on the limitations of the computational resources of the hardware. The YOLOv10m version, which represents a medium version for universal use, is chosen for the experimental study. The dataset for training the neural network includes two classes of images: UAVs and birds. The dataset files in the form of images and videos are taken from the UAV’s own flight archive, as well as open-source resources Roboflow, Kaggle, Ultralytics, and GitHub. The

stages of image annotation, augmentation and dataset distribution in the percentage ratio of 70/20/10 (training, validation and testing, respectively) are implemented in Roboflow.com service.

Training, validation and testing of the model (Figure 2) YOLOv10m is realized on the basis of AD103 graphics processor of NVIDIA GeForce RTX 4080 graphics card with support of CUDA Toolkit 12.1. The program code in Python language is implemented in PyCharm 2024 environment using Ultralytics framework, which allows using the open source resource of pre-trained YOLOv10 models. The following hyperparameters are set for training the neural network on the user set: number of epochs: 300; packet size: 16; learning rate: 0.001; momentum: 0.9; weight drop: 0.0005 and image size: 640. The trained YOLOv10m model has a file size of “best.pt” of 101 MB. In order to determine the FPS, the trained neural network is tested on inference using two test videos of UAV and bird flights.

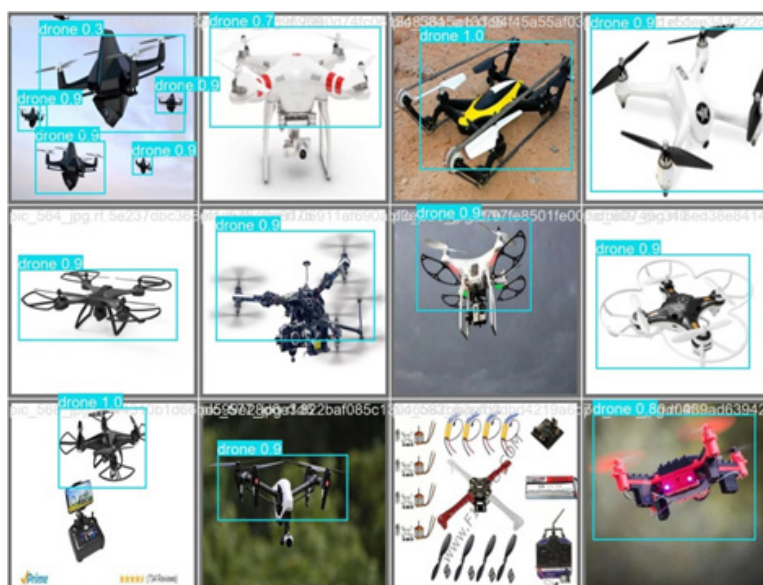


Figure 2 – Testing of the trained YOLOv10m neural network model

To compare the accuracy of the experimental neural network with the previous versions, the YOLOv8m and YOLOv9m models corresponding in purpose were also trained on the user dataset. The mAP50-95 metric was chosen as the comparative accuracy parameter to evaluate the ability of visual detectors to recognize and localize small objects.

Results

Figure 3a-d shows the accuracy metrics of the YOLOv10m model trained on a custom dataset of UAVs and birds.

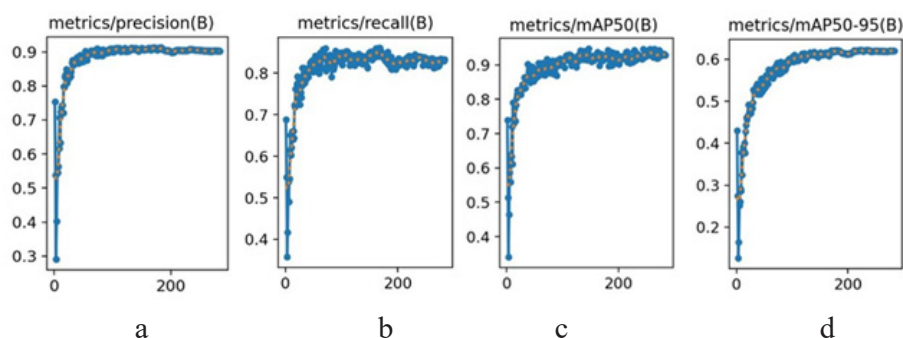


Figure 3 – Metrics of training results of the YOLOv10m neural network:
a - Precision; b - Recall; c - mAP50; d - mAP50-95

The frames of testing the trained model on the inferno using two videos of a DJI F450 UAV flight and a bird are shown in Figure 4a, b.



Figure 4 – Inference frames of the trained YOLOv10m model:
(a) frame from the DJI F450 flight video; (b) frame of a bird (seagull) flight

Figure 5a, b shows the comparison plots of mAP50-95 accuracy metric and fast performance (FPS) of YOLOv8m, YOLOv9m and YOLOv10m neural networks.

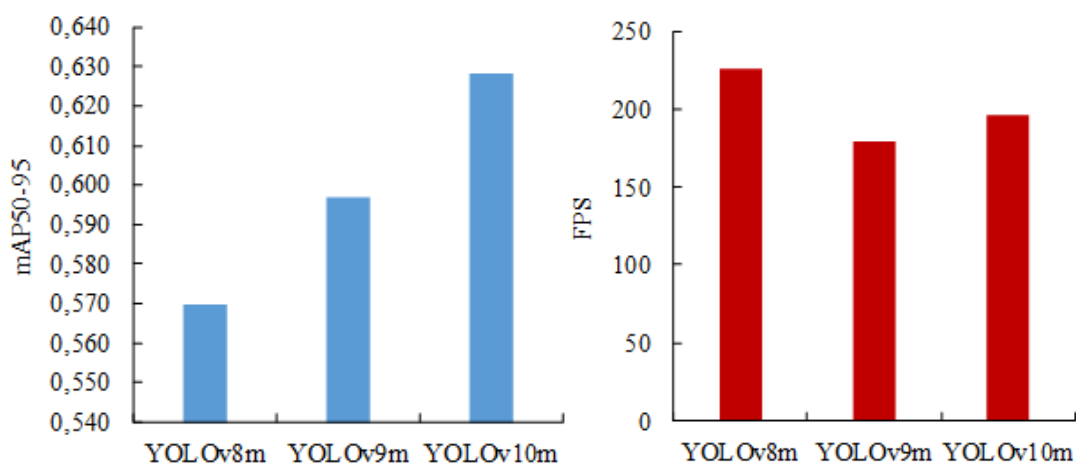


Figure 5 – Comparison plots of mAP50-95 accuracy metrics and FPS of YOLOv8m, YOLOv9m and YOLOv10m neural networks

As a result of the diagrams analysis we can conclude about the effectiveness of using the trained YOLOv10m model in the tasks of recognition and classification of UAVs and birds in optoelectronic surveillance systems.

Discussion

A dataset of 6255 UAV and bird images is prepared for training the experimental model of YOLOv10m neural network. After annotation and augmentation stages, the user dataset using Roboflow.com service is distributed in the percentage of 70/20/10 for training, validation and testing tasks, respectively. The training of experimental neural network resulted in the following maximum values of accuracy metrics (Figure 4a–d):

Completeness: 0.883;
Accuracy: 0.907;
mAP50: 0.953;
mAP50-95: by 0.628.

In terms of mAP50 metric, this model outperforms detectors from [2–7]. Comparing the mAP50-95 metrics (Figure 5a), YOLOv10m outperforms YOLOv8 by 0.058 and YOLOv9 by 0.031 in terms of accuracy, which proves the greater efficiency of the first model to recognize and localize small objects such as UAVs and birds in the far-field surveillance area. In terms of FPS, YOLOv10m is second only to the trained model YOLOv8m, which proves that the first model can be used in real-time in optoelectronic surveillance channels, combining the segmentation and tracking functions of small objects such as UAVs and birds.

Conclusion

1) In order to develop a new efficient model for UAV and bird recognition and classification, the YOLOv10m model is selected, which bypasses the Non-Maximal Suppression (NMS) algorithm in its architecture by introducing the concepts of dual label assignment, consistent metric for prediction matching and focus on accuracy and efficiency in design. A user dataset of 6,255 images is prepared to train the experimental models, of which 4,379 (70%) are specifically designed for training, 1,251 (20%) are oriented for validation and the remaining 625 (10%) are used in the testing phase of the neural network.

2) YOLOv10m training is implemented on the basis of AD103 GPU of NVIDIA GeForce RTX 4080 video card with CUDA Toolkit 12.1 support. As a result of evaluation of the obtained accuracy metrics, YOLOv10 outperforms the known visual detectors YOLOv2, YOLOv3, YOLOv4, YOLOv5, YOLOv7, YOLOX and YOLOv8 described by mAP50 [2–7].

3) To compare the efficiency of recognition and localization of small objects using the mAP50-95 metric, previous versions of the YOLO algorithm: YOLOv8 and YOLOv9 were trained on a user dataset. YOLOv10m was 9% more accurate than the former and 5% more accurate than the latter.

4) As a result of testing the trained YOLOv10m neural network on inference, it is found that this model is able to recognize and classify UAVs and birds in real time with high accuracy and efficiency. According to the high FPS values (average 196 FPS), this model is capable of segmenting and tracking localized objects, combining the classification task.

5) The results of this study will be useful to the developers of UAV and bird detection, recognition and classification systems.

Acknowledgment

The article includes the results of research carried out within the framework of grant funding under the project IRN AP19679009 funded by the Committee of Science of the Ministry of Science and Higher Education of the Republic of Kazakhstan.

REFERENCES

- 1 Alkentar, Saad & Alsahwa, B. & Assalem, A. & Karakolla, D. Practical comparison of the accuracy and speed of YOLO, SSD and Faster RCNN for drone detection. *Journal of Engineering*, 27, 19–31 (2021). <https://doi.org/10.31026/j.eng.2021.08.02>.
- 2 Seidaliyeva, U., Alduraibi, M., Ilipbayeva, L., Almagambetov, A. Detection of Loaded and Unloaded UAV Using Deep Neural Network. In *Proceedings of the 2020 Fourth IEEE International Conference on Ro-botic Computing (IRC) (Taichung, Taiwan, November 9–11, 2020)*, pp. 490-494. <https://doi.org/10.1109/IRC.2020.00093>.

3 Alsanad, Hamid & Sadik, Amin & Ucan, Osman & Ilyas, Muhammad & Bayat, Oguz. YOLO-V3 based real-time drone detection algorithm. Multimedia Tools and Applications, 81, 1–14 (2022). <https://doi.org/10.1007/s11042-022-12939-4>.

4 Singha, S., Aydin, B. Automated Drone Detection Using YOLO v4. Drones. September 2021. <https://doi.org/10.3390/drones5030095>.

5 Al-Qubaydhi, Nader & Alenezi, Abdulrahman & Alanazi, Turki & Senyor, Abdulrahman & Alanezi, Naif & Alotaibi, Bandar & Alotaibi, Munif & Razaque, Abdul & Abdelhamid, Abdelaziz & Alotaibi, Aziz. Detection of Unauthorized Unmanned Aerial Vehicles Using YOLOv5 and Transfer Learning. Electronics, 11, 2669 (2022). <https://doi.org/10.3390/electronics11172669>.

6 Aydin, B., Singha, S. Drone Detection Using YOLO v5, Eng. 4 (2023). <https://doi.org/10.3390/eng4010025>.

7 Zhai, X., Huang, Z., Li, T., Liu, H., Wang, S. YOLO-Drone: An Optimized YOLOv8 Network for Tiny UAV Object Detection. Electronics, 12, 3664 (2023). <https://doi.org/10.3390/electronics12173664>.

8 Li, Jun & Feng, Yongqiang & Shao, Yanhua & Liu, Feng. (2024). IDP-YOLOV9: Improvement of Object Detection Model in Severe Weather Scenarios from Drone Perspective. Applied Sciences, 14, 5277 (2023). <https://doi.org/10.3390/app1412125277>.

9 Muzammul, Muhammad & Algarni, Abdul & Ghadi, Yazeed & Assam, Muhammad. Enhancing UAV Aerial Image Analysis: Integrating Advanced SAHI Techniques with Real-Time Detection Models on the VisDrone Dataset. IEEE Access, pp. 1–1 (2024). <https://doi.org/10.1109/ACCESS.2024.3363413>.

¹Семенюк В.В.,

Т.Ф.М., ORCID ID: 0000-0002-8580-7326,

e-mail: vvsemenyuk@ku.edu.kz;

^{1*}Курмашев И.Г.,

PhD, ORCID ID: 0000-0001-9872-7483,

*e-mail: ikurmashev@ku.edu.kz;

²Сербин В.В.,

PhD, қауымдастырылған профессор, ORCID ID: 0000-0002-5807-3873,

e-mail: v.serbin@satbayev.university;

¹Курмашева Л.Б.,

магистр, ORCID ID: 0000-0001-5392-5873,

e-mail: lbkurmasheva@ku.edu.kz;

²Молдагулова А.Н.,

PhD, профессор, ORCID ID: 0000-0002-1596-561X,

e-mail: a.moldagulova@satbayev.university;

¹М. Қозыбаев атындағы Солтүстік Қазақстан университеті, Петропавл қ., Қазақстан

²Satbayev University, Алматы қ., Қазақстан.

YOLOV10 НЕЙРОНДЫҚ ЖЕЛІСІ НЕГІЗІНДЕ НАҚТЫ УАҚЫТТА ҰШҚЫШСЫЗ ҰШУ АППАРАТТАРЫН ТАНУ МОДЕЛІН ӘЗІРЛЕУ

Аңдатпа

Мақалада YOLOv10 нейрондық желісін оқыту негізінде нақты уақыт режимінде ұшқышсыз ұшу аппараттары мен құстарды тану және жіктеу моделін әзірлеу қарастырылады. Зерттеу бағыты әртүрлі салаларда оларды қолданудың артуын ескере отырып, қауіпсіздік тұрғысынан ұшқышсыз ұшу аппараттарын анықтау мәселелерінің өзектілігімен байланысты. Модельді оқыту үшін жеке мұрағаттар мен ашық ресурстардан жиналған 6255 суреттен тұратын деректер жиынтығы дайындалды. Аннотациялау, деректерді күшейту және тарату процесі Roboflow.com қызметі арқылы жүзеге асырылды. Модельді оқыту ULTRALYTICS фреймворкін қолдана отырып, NVIDIA GeForce RTX 4080 графикалық процессорында жүргізілді. Тестілеу нәтижелері mAP50 және mAP50–95 көрсеткіштері бойынша YOLO желісінің алдыңғы нұсқаларымен салыстырғанда танудың жоғары дәлдігін көрсетті. Модель объектілерді тиімді

сегментациялау және қадағалау қабілетін көрсетті, бұл оны оптикалық-электронды бақылау жүйелерінде қолдануға перспективалы етеді. Зерттеу нәтижелері ұшқышсыз ұшу аппараттары мен құстарды анықтау және жіктеу жүйелерін әзірлеушілерге, сондай-ақ әртүрлі салаларда қауіпсіздікті арттыруға пайдалы болуы мүмкін.

Тірек сөздер: нейрондық желілер, жіктеу, тану, оптикалық-электронды бақылау арналары, UAV, YOLO, конволюциялық нейрондық желілер.

¹Семенюк В.В.,

М.Т.Н., ORCID ID: 0000-0002-8580-7326,

e-mail: vvsemenyuk@ku.edu.kz

^{1*}Курмашев И.Г.,

PhD, ORCID ID: 0000-0001-9872-7483,

*e-mail: ikurmashev@ku.edu.kz

²Сербин В.В.,

PhD, ассоциированный профессор, ORCID ID: 0000-0002-5807-3873,

e-mail: v.serbin@satbayev.university

¹Курмашева Л.Б.,

магистр, ORCID ID: 0000-0001-5392-5873,

e-mail: lbkurmasheva@ku.edu.kz

²Молдагулова А.Н.,

PhD, профессор, ORCID ID: 0000-0002-1596-561X,

e-mail: a.moldagulova@satbayev.university

¹Северо-Казахстанский университет имени М. Козыбаева, г. Петропавловск, Казахстан

²Satbayev University, г. Алматы, Казахстан

РАЗРАБОТКА МОДЕЛИ РАСПОЗНАВАНИЯ БПЛА В РЕЖИМЕ РЕАЛЬНОГО ВРЕМЕНИ НА ОСНОВЕ НЕЙРОННОЙ СЕТИ YOLOV10

Аннотация

В статье рассматривается разработка модели распознавания и классификации БПЛА и птиц в режиме реального времени на основе обучения нейронной сети YOLOv10. Направление исследования считается актуальным в связи с проблемами обнаружения БПЛА в контексте обеспечения безопасности, учитывая их растущее использование в различных сферах. Для обучения модели подготовлен датасет, состоящий из 6255 изображений, собранных из собственных архивов и открытых ресурсов. Процесс аннотирования, аугментации и распределения данных был реализован с использованием сервиса Roboflow.com. Обучение модели проводилось на графическом процессоре NVIDIA GeForce RTX 4080 с использованием фреймворка Ultralytics. Результаты тестирования показали высокую точность распознавания с метриками mAP50 и mAP50-95, превышающими показатели предыдущих версий YOLO. Модель демонстрирует способность к эффективной сегментации и трекингу объектов, что делает ее перспективной для применения в системах оптикоэлектронного наблюдения. Результаты исследования могут быть полезны для разработчиков систем обнаружения и классификации БПЛА и птиц, а также для повышения безопасности в различных областях.

Ключевые слова: нейронные сети, классификация, распознавание, оптикоэлектронные каналы наблюдения, БПЛА, YOLO, сверточные нейронные сети.

Article submission date: 23.09.2025

UDC 004.855.6
IRSTI 20.19.19

<https://doi.org/10.55452/1998-6688-2025-22-4-227-243>

^{1*}**Mukhsimbayev B.,**

PhD student, ORCID ID: 0009-0008-4606-3628,

*e-mail: b.mukhsimbaev@kbtu.kz

¹**Pak A.,**

PhD, Professor, ORCID ID: 0000-0002-8685-9355,

e-mail: a.pak@kbtu.kz

¹**Kuralbayev A.,**

PhD student, ORCID ID: 0009-0001-0811-5385,

e-mail: a.kuralbaev@kbtu.kz

¹Kazakh-British Technical University, Almaty, Kazakhstan

A COMPUTATIONAL PIPELINE FOR LEXICAL AND THEMATIC ANALYSIS OF THE CODE OF ADMINISTRATIVE OFFENSES OF THE REPUBLIC OF KAZAKHSTAN

Abstract

This study introduces a computational pipeline for the automated linguistic and structural analysis of legal texts, applied to the Code of Administrative Offenses of the Republic of Kazakhstan (CAO RK, K1400000235). The proposed workflow integrates data collection, text preprocessing, tokenization, keyword extraction, semantic clustering, and visualization using natural language processing (NLP) and statistical techniques implemented in Python. The pipeline unites lexical, thematic, and quantitative linguistic analyses into a coherent sequence that enables the identification of frequency distributions, semantic fields, and latent topics across the hierarchical structure of the Code (sections, chapters, and articles). The analysis of the CAO RK corpus revealed several distinctive linguistic patterns: a dominance of sanction and responsibility-related vocabulary (штраф, ответственность, правонарушение), high lexical density in chapters regulating economic and procedural offenses, and concentrated thematic clusters reflecting the normative-punitive orientation of administrative law. Visualization techniques such as frequency histograms, thematic heatmaps, and topic maps illustrate the potential of the pipeline for exploring legislative language quantitatively. Overall, the framework establishes a scalable foundation for comparative legal linguistics, automated legislative monitoring, and the modernization of legal analytics in Kazakhstan.

Keywords: administrative law, legal text analysis, natural language processing, computational legal linguistics, frequency analysis, topic modeling, legal informatics.

Introduction

Legal texts are among the most structured and linguistically formal types of documents, embodying both the conceptual logic and the institutional framework of law. As corpus linguistics continues to evolve globally, the creation of domain-specific linguistic corpora – including legal and administrative texts – has become a crucial research direction [1]. Such corpora not only preserve linguistic authenticity but also enable the systematic study of how institutional discourse reflects socio-legal priorities.

The Code of Administrative Offenses of the Republic of Kazakhstan (CAO RK) is a core legislative act regulating administrative liability, sanctions, and procedural mechanisms within the national legal system. Comparative analyses of administrative liability across Commonwealth of Independent States (CIS) countries underscore Kazakhstan's distinctive legal structure and its emphasis on detailed qualifying characteristics [13].

However, quantitative or computational linguistic approaches to Kazakhstan's legal texts remain largely unexplored. The absence of large-scale empirical analyses of lexical and thematic regularities in the CAO RK creates a methodological gap between traditional legal scholarship and modern computational linguistics.

To address this gap, the present study applies computational linguistics and natural language processing (NLP) techniques to the Code of Administrative Offenses of the Republic of Kazakhstan (CAO RK), offering an empirical perspective on its lexical and thematic organization. Specifically, the study introduces a unified computational pipeline that integrates lexical, thematic, and quantitative linguistic analyses within a single workflow.

Automated analysis of legislative texts and amendments in the Republic of Kazakhstan offers new possibilities for identifying thematic and lexical trends that reflect shifts in state policy priorities. Such an approach reduces the subjectivity of legal interpretation, enhances the transparency of normative evolution, and provides a foundation for more evidence-based interaction among legislators, researchers, and civil society.

By integrating computational methods with legal doctrine, it becomes possible to move beyond descriptive interpretation toward data-driven legal analytics that objectively trace how legal language changes over time and across domains. Recent works employing large language models for empirical legal studies demonstrate the growing potential of AI-assisted thematic and linguistic analysis within judicial and legislative corpora [14].

A preliminary quantitative analysis of the Code of Administrative Offenses of the Republic of Kazakhstan reveals a high lexical concentration of sanction-related and responsibility-oriented vocabulary, particularly within sections regulating economic and procedural offenses. This pattern underscores the normative-punitive orientation of administrative law and its focus on enforcement mechanisms rather than judicial adjudication. Thematic clustering further highlights distinct semantic zones within the Code, where linguistic density corresponds to institutional priorities such as governance, entrepreneurship regulation, and procedural enforcement.

These findings demonstrate the potential of computational linguistics to uncover structural and conceptual patterns that are not easily visible through traditional legal interpretation.

Specifically, the study introduces a unified computational pipeline that integrates lexical, thematic, and quantitative linguistic analyses within a single workflow. Applied to the CAO RK corpus, the pipeline integrates methods from NLP and legal informatics to:

1. Quantify lexical frequency and distribution of key legal terms;
2. Detect semantic clusters and thematic patterns through topic modeling;
3. Examine the variation of linguistic features across the hierarchical structure of the Code (sections, chapters, and articles).

The main contributions of this work are as follows:

1. Development of an end-to-end computational pipeline for the collection, preprocessing, and analysis of legislative texts in a reproducible manner;
2. Identification of lexical and thematic regularities in the CAO RK, including the predominance of sanction-related and responsibility-oriented vocabulary;
3. Discovery of structural and topical hierarchies within the Code through quantitative modeling and visualization (e.g., heatmaps and topic clustering);
4. Demonstration of how computational linguistic analysis can support legislative transparency, policy monitoring, and comparative studies of Kazakhstan's legal language.

Together, these contributions provide both a methodological innovation and new empirical insights into the structure and evolution of Kazakhstan's administrative law.

Related Works

Numerous linguistic and legal studies have examined the textual and terminological dimensions of Kazakhstan's legislation – including analyses of translation accuracy, linguistic expertise, and terminology standardization [2, 5, 9]. However, these works primarily focus on bilingual authenticity

and linguistic interpretation, while corpus-based and statistical approaches to the structure of Kazakhstan's codes remain limited.

Frameworks such as the Leipzig Corpus Miner illustrate how qualitative legal analysis can be enhanced through text-mining architectures that combine close and distant reading – an approach conceptually aligned with our computational pipeline for analyzing legislative corpora [16]. Recent advances in legal natural language processing have demonstrated how machine learning and large language models can assist in thematic and structural segmentation of legal texts, suggesting promising directions for applying similar techniques to normative corpora [14, 15].

Recent developments in computational linguistics for the Kazakh language – such as the creation of the KazQAD open-domain QA dataset, the KazSAnDRA sentiment analysis dataset, the Aligned Kazakh–Russian Criminal Corpus, large-scale instruction-tuning datasets for low-resource governance and legal domains, and corpus-driven approaches to crime-related event extraction for low-resource languages – have significantly expanded the available linguistic resources for Kazakh [6–8, 11, 12]. Nevertheless, most of these initiatives target general-domain or multilingual NLP rather than formal legislative corpora.

In parallel, legal scholars emphasize the ongoing challenges of ensuring authenticity and linguistic precision in Kazakh–Russian legal texts [2, 4, 9, 10]. These works underline the need for standardized terminology and systematic linguistic evaluation – issues that computational methods can help address by providing quantitative diagnostics of translation consistency and lexical coherence across codes.

Automated analysis of legislative texts and amendments in the Republic of Kazakhstan offers new possibilities for identifying thematic and lexical trends that reflect shifts in state policy priorities. Such an approach reduces the subjectivity of legal interpretation, enhances the transparency of normative evolution, and provides a foundation for more evidence-based interaction among legislators, researchers, and civil society. Recent works employing large language models for empirical legal studies demonstrate the growing potential of AI-assisted thematic and linguistic analysis within judicial and legislative corpora [14]. Frameworks developed for qualitative content analysis, such as the Leipzig Corpus Miner, further contextualize the significance of connecting linguistic analysis with institutional and semantic interpretation [16].

Materials and methods

Data

The dataset used in this research was collected from the official legal information portal adilet.zan.kz, which hosts the digital versions of Kazakhstan's legislative acts. To enable large-scale and reproducible data collection, a Python-based Scrapy parser was developed. This parser can automatically retrieve the full text of any legal code from adilet.zan.kz using its unique document identifier (e.g., K1400000235).

During extraction, all textual and structural elements – including sections, chapters, articles, and paragraphs – are normalized and serialized into a consistent, machine-readable JSON schema. The framework is designed to support automated dataset updates: when amendments or new versions of a Code are published, the parser can re-collect and synchronize the data without manual intervention. This functionality allows the framework to be extended to future analyses of other Kazakhstani legal codes (e.g., Civil, Criminal, or Tax Codes), ensuring both scalability and long-term maintainability of the corpus.

The parser was applied to the Code of Administrative Offenses of the Republic of Kazakhstan (K1400000235), performing full hierarchical segmentation into:

- ◆ Sections (Разделы),
- ◆ Chapters (Главы),
- ◆ Articles (Статьи),
- ◆ Paragraphs (Пункты).

Each paragraph is stored with metadata describing its position in the hierarchy and textual content. The resulting dataset contains 3,336 paragraph-level units, represented through the following metadata fields:

- ♦ doc_id – unique document identifier
- ♦ lang – language of the text
- ♦ url – source link on Adilet.zan.kz
- ♦ title – official title of the Code
- ♦ status – legal status (e.g., updated)
- ♦ citation – formal legal citation
- ♦ section – section title
- ♦ chapter – chapter title
- ♦ article_id – internal article ID
- ♦ article_title – article heading
- ♦ article_text – full article text
- ♦ paragraphs – list of paragraph-level units
- ♦ notes – amendment and commentary notes
- ♦ links – cross-references to other laws

This structured representation was subsequently exported in JSON format, ensuring hierarchical consistency and facilitating downstream linguistic and statistical analysis. A descriptive statistical overview of the resulting corpus – including the number of articles, paragraphs, and amendment notes – is presented in Section 4.1.

Preprocessing

All paragraph-level texts were preprocessed using the spaCy natural language processing library with the ru_core_news_sm language model. This choice was motivated by the need for a deterministic, transparent, and linguistically grounded preprocessing pipeline suitable for large-scale legal corpora.

Although large language models (LLMs) offer advanced semantic interpretation, their outputs are often non-deterministic and depend on model-specific context, prompts, or sampling parameters. Moreover, LLMs require substantial computational resources, making them less suitable for foundational linguistic preprocessing tasks such as tokenization and lemmatization, where consistency, efficiency, and reproducibility are essential.

The spaCy framework ensures robust tokenization, lemmatization, and part-of-speech tagging for Russian text, while operating fully offline without reliance on external APIs or network connectivity. The lightweight ru_core_news_sm model was specifically chosen to balance computational efficiency and linguistic accuracy, enabling rapid processing of thousands of legal paragraphs while preserving the precision required for lexical frequency and morphological analysis.

The text preprocessing pipeline comprised the following sequential stages:

1. Text normalization – conversion of all text to lowercase and removal of punctuation, numerical symbols, and extraneous whitespace to ensure uniform lexical representation.
2. Tokenization – segmentation of normalized text into individual lexical units (tokens).
3. Stop-word filtering – exclusion of high-frequency Russian functional words using the stop_words_ru corpus to eliminate non-informative linguistic elements.
4. Lemmatization – reduction of each token to its canonical (dictionary) form to consolidate morphological variants of the same lexical item.
5. Storage of results - creation of a new field, tokens, containing the final list of lemmatized words for each paragraph.

This standardized preprocessing pipeline ensured linguistic consistency and reproducibility across the entire dataset, forming the basis for subsequent lexical and statistical analyses.

Each paragraph was represented in a normalized JSON structure that preserves both its hierarchical and linguistic attributes. An example entry is shown below:

```
{
  "doc_id": "K1400000235",
  "section": "РАЗДЕЛ 1. ОБЩИЕ ПОЛОЖЕНИЯ",
  "chapter": "Глава 1. ЗАКОНОДАТЕЛЬСТВО ОБ АДМИНИСТРАТИВНЫХ ПРАВОНАРУШЕНИЯХ",
  "article_id": "z4",
  "paragraph_id": "z5",
  "text": "Законодательство Республики Казахстан об административных правонарушениях состоит из настоящего кодекса.",
  "tokens": ["законодательство", "республика", "казахстан", "административный", "правонарушение", "кодекс"]
}
```

This structured representation ensures direct correspondence between the legal text and its computationally processed form, facilitating precise alignment of linguistic and legal features across the corpus. The resulting tokenized corpus was used as the primary input for lexical frequency analysis, keyword distribution studies, and visualization through frequency plots and thematic heatmaps.

Furthermore, the framework was designed to remain extensible and interoperable with advanced NLP pipelines. In particular, it supports potential integration with large language models (LLMs) for semantic enrichment tasks such as contextual classification, topic expansion, and automated summarization. These capabilities, however, were not utilized in the present study, which focuses on deterministic linguistic and statistical analyses.

Processing

The processing stage includes the computational procedures applied after the initial corpus collection and preprocessing. This stage describes the extraction of thematic features, the selection and grouping of legally relevant keywords, and the application of quantitative analytical methods to the normalized text. The objective of this module is to transform the lemmatized corpus into a structured representation suitable for lexical frequency analysis, semantic field exploration.

Subsections 3.3.1 and 3.3.2 describe the construction of the legal keyword lexicon used for thematic analysis and the computational tools applied throughout the study. These components form the core of the quantitative workflow used to examine the lexical structure of the CAO RK.

In addition, Subsection 3.3.3 outlines the hyperparameters used in the numerical experiments, ensuring transparency and reproducibility of the analytical results.

Keyword Selection

To enable thematic and semantic analysis, a specialized lexicon of key legal concepts was constructed. The lexicon reflects the principal conceptual domains of administrative law and was manually curated based on the terminology and recurrent expressions found in the Code of Administrative Offenses of the Republic of Kazakhstan (CAO RK).

Each category corresponds to a semantic field encompassing morphologically related words and stems. This approach allows for the inclusion of lexical variation (e.g., inflectional and derivational forms) while maintaining semantic coherence across categories. The principal thematic categories and representative keyword stems are listed below:

- ◆ Sanctions – штраф, взыскан, санкц, наказан
- ◆ Responsibility – ответственн, виновн
- ◆ Offense – правонарушен, нарушен
- ◆ Person – лицо, гражданин, должностн
- ◆ Entrepreneurship – предприниматель, бизнес, юрилиц
- ◆ Authority – орган, комитет, инспекц, министерств
- ◆ Court – суд, судья, кассац, апелляц
- ◆ Property – имущество, собственност, доход
- ◆ Size / Measure – размер, расчетн, месячн, МРП
- ◆ Case – дело, производств, рассмотрен
- ◆ State – казахстан, республика, кодекс

Occurrences of these keyword stems were tracked across paragraphs, articles, and chapters to quantify their frequency, relative prominence, and contextual distribution within the Code. Grouping tokens by morphological stems (e.g., `ответственн` → `ответственность`, `ответственный`) ensured robust lexical matching and prevented over-fragmentation of semantically equivalent terms.

This curated keyword lexicon thus served as the analytical foundation for quantitative thematic profiling and semantic field visualization in subsequent stages of analysis.

Tools

All stages of the analysis were conducted in a Python 3.11 environment using Jupyter Notebook as the interactive development platform. The computational workflow relied exclusively on open-source libraries widely adopted in the fields of data analysis and natural language processing. The primary libraries and their respective functions are summarized below:

- ◆ pandas – data ingestion, transformation, and aggregation;
- ◆ spaCy – tokenization and lemmatization of Russian texts (`ru_core_news_sm` model);
- ◆ NLTK – stop-word filtering and frequency-based text statistics;
- ◆ matplotlib and seaborn – visualization of lexical distributions and thematic heatmaps;
- ◆ pathlib and json – structured I/O and corpus serialization.

All scripts were executed within a notebook-based reproducible workflow, ensuring full transparency, modularity, and consistency across the stages of data preprocessing, keyword extraction, and statistical analysis.

Hyperparameters for numerical experiments

Latent Dirichlet Allocation (LDA) was used to examine higher-level thematic structures within the corpus. Input matrix was constructed with a bag-of-words representation using `CountVectorizer`. To ensure transparency and reproducibility, all hyperparameters and configuration settings are reported below.

Vectorization settings (`CountVectorizer`):

- ◆ `max_df` = 0.9 – terms appearing in more than 90% of paragraphs were excluded as overly frequent;
- ◆ `min_df` = 3 – terms occurring in fewer than three paragraphs were removed to reduce sparsity;
- ◆ stop words – no additional stop-word lists were applied beyond earlier preprocessing steps;
- ◆ tokenization – pre-tokenized and lemmatized tokens were joined into whitespace-separated strings.

LDA model settings (`LatentDirichletAllocation`):

- ◆ `n_components` = 5 – the number of latent topics;
- ◆ `random_state` = 42 – fixed seed for reproducibility;
- ◆ `learning_method` = 'batch';
- ◆ `max_iter` = 10;
- ◆ `doc_topic_prior` (α) = $1 / n\_components$;
- ◆ `topic_word_prior` (β) = $1 / n\_components$.

These settings were chosen to balance topic interpretability and computational efficiency, given the relatively compact size of the corpus. The resulting model outputs were transformed into interactive visualizations using `pyLDAvis`.

The visualization module was supplied with:

- ◆ topic-term distributions – (`lda.components_` normalized row-wise),
- ◆ document-topic distributions – (`lda.transform(X)`),
- ◆ document lengths – computed from token counts,
- ◆ vocabulary – extracted from the vectorizer,
- ◆ term frequencies – aggregated across all documents.

This configuration enabled an interpretable inspection of topic coherence, dominant terms, and the overall thematic landscape of the Code.

Results

Descriptive Overview of the CAO RK Corpus

Prior to conducting lexical and thematic analyses, the corpus was examined descriptively to establish its structural composition and quantitative characteristics. Table 1 presents the summary statistics of the Code of Administrative Offenses of the Republic of Kazakhstan (CAO RK) corpus compiled for this study.

The dataset comprises 1,050 articles and 3,336 paragraphs, with an average paragraph length of approximately 40 words. The longest paragraph contains more than 2,000 words, reflecting the extensive procedural detail typical of certain sections of administrative law. In total, 781 legislative notes (сноски) were identified, documenting amendments introduced between 2014 and 2025. These amendment annotations provide clear evidence of the Code’s continuous normative evolution over the past decade.

Table 1 – Summary characteristics of the Code of Administrative Offenses of the Republic of Kazakhstan (CAO RK)

Metric	Value
Total number of articles (статьи)	1,050
Total number of paragraphs (пункты)	3,336
Average paragraph length (words)	39.6
Maximum paragraph length (words)	2,042
Number of amendment notes (сноски)	781
Year range of amendments	2014–2025

Word Frequency Distribution

The twenty most frequent words in the Code of Administrative Offenses of the Republic of Kazakhstan (CAO RK) are predominantly associated with quantitative measures (e.g., размер, показатель, месячный расчётный показатель) and legal actors (e.g., лицо, субъект).

This lexical pattern highlights the Code’s focus on defining responsibilities, entities, and the scale of administrative penalties – central dimensions of administrative law.

Figure 1, presented directly below, presents a horizontal bar chart depicting the relative frequencies of the most common lexical units in the corpus after lemmatization and stop-word removal. The clear predominance of размер (“amount”) and лицо (“person”) underscores the quantitative and person-centric orientation of the CAO RK, revealing its emphasis on measurable sanctions and the identification of liable subjects.

Distribution of Key Terms Across Chapters

To investigate the thematic organization of the Code of Administrative Offenses of the Republic of Kazakhstan (CAO RK), a set of keywords was grouped into semantic categories representing the principal conceptual domains of administrative law. These categories encompass notions of sanctions, responsibility, administrative subjects, institutional governance, economic relations, and procedural operations.

Each paragraph of the Code was tokenized and normalized into lemmas, after which the occurrences of keyword stems were aggregated at the chapter level. This procedure enabled a comparative assessment of the relative prominence of distinct legal domains within the legislative corpus.

The results demonstrate a non-uniform lexical distribution across chapters. Chapters addressing personal rights, entrepreneurial activity, and public administration exhibit a particularly high density of sanction-related and institutional vocabulary (e.g., санкции, ответственность, штраф), indicating an emphasis on punitive and regulatory enforcement mechanisms. By contrast, chapters governing

procedural processes show elevated frequencies of lexemes associated with дело (“case”) and орган (“authority”), reflecting the procedural dimension of administrative adjudication.

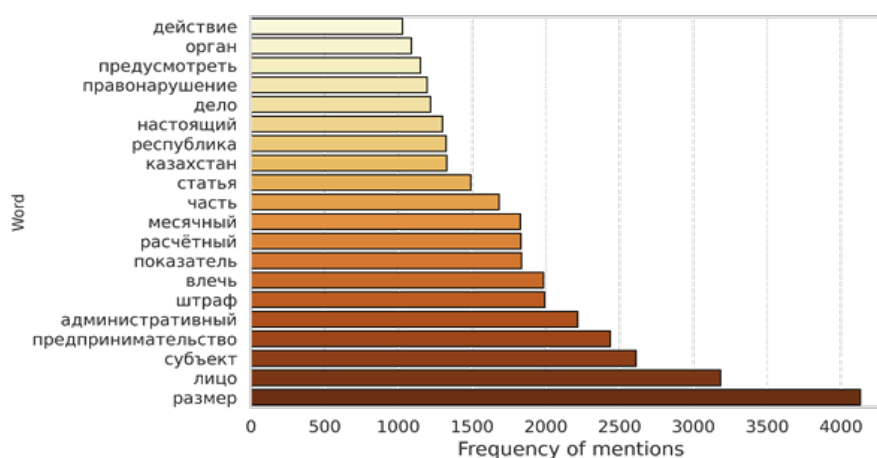


Figure 1 – Top-20 most frequent words in the Code of Administrative Offenses

Overall, the analysis suggests that the CAO RK is lexically organized around several semantic clusters corresponding to its functional domains—sanction-related, procedural, administrative, and economic. This thematic stratification highlights the internal diversity of administrative law, where the distribution of linguistic features reflects the differentiated structure of legal regulation.

Figure 2, presented immediately after this paragraph, visualizes the ten chapters with the highest cumulative frequency of thematic keywords. The chart demonstrates clear disparities in the lexical salience of legal concepts across chapters, reflecting the internal hierarchy of regulatory focus areas within the Code.

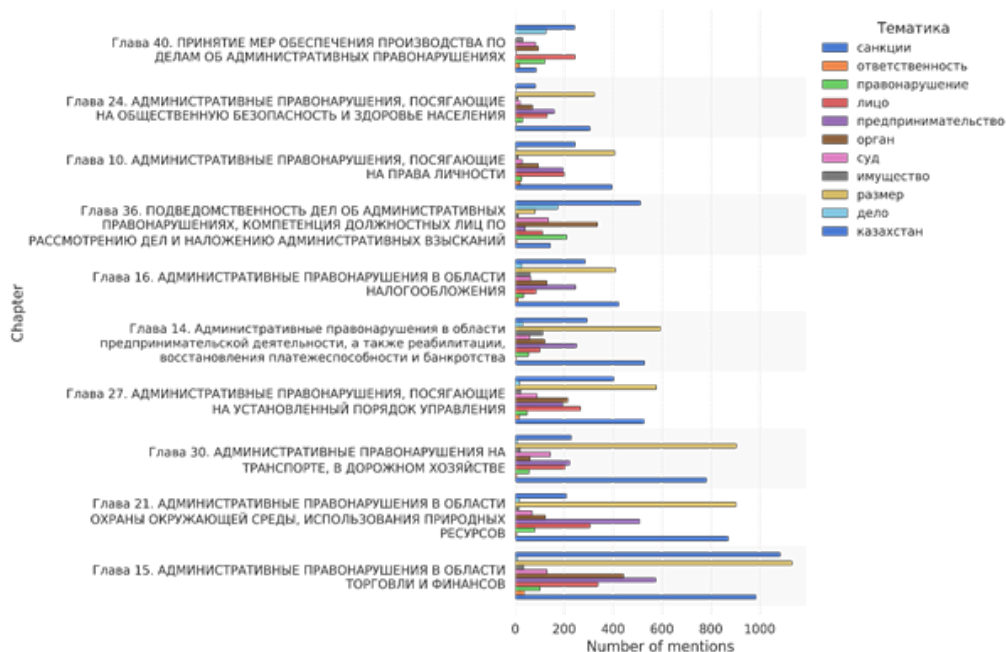


Figure 2 – Top-10 chapters of the CAO RK by frequency of thematic keywords (grouped by semantic categories)

Thematic Heatmap

A thematic heatmap (Figure 3) was constructed to visualize the distribution of key legal concepts across the structural sections of the Code of Administrative Offenses of the Republic of Kazakhstan (CAO RK). The visualization reveals distinct lexical clustering patterns aligned with the primary domains of administrative regulation.

The semantic fields associated with санкции (sanctions), ответственность (responsibility), and лицо (person) exhibit consistently high frequencies, underscoring their central role in the conceptual architecture of the Code. Conversely, terms connected to имущество (property) and суд (court) are concentrated in procedural and enforcement-related sections, indicating specialized legal subdomains.

Sections regulating entrepreneurship and property display elevated lexical density across multiple thematic categories – particularly in sanction-related vocabulary – reflecting the Code’s emphasis on economic accountability and regulatory oversight.

Collectively, the heatmap demonstrates that the CAO RK is not lexically uniform but instead organized around discrete thematic cores corresponding to its institutional and functional structure.

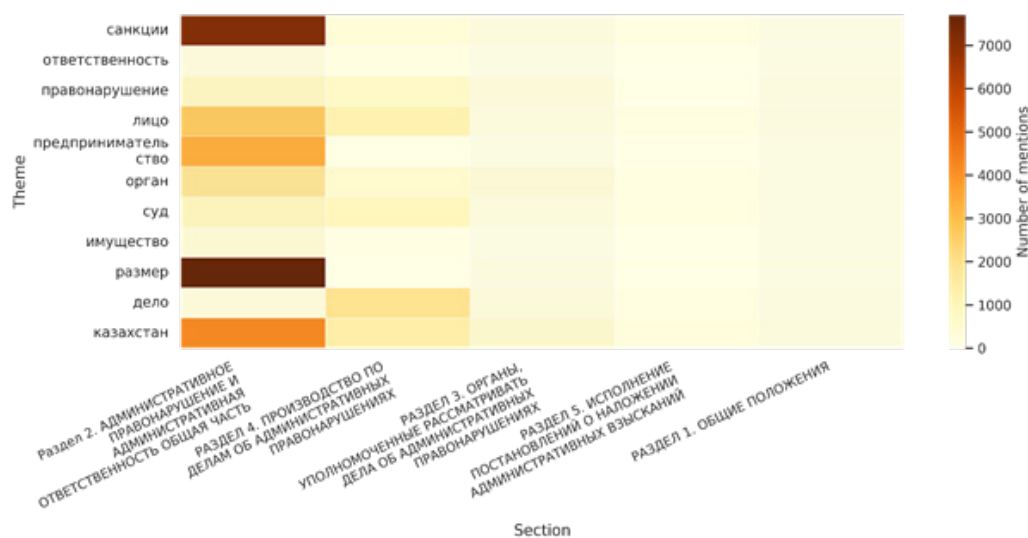


Figure 3 – Thematic distribution of legal categories across sections of the Code of Administrative Offenses of the Republic of Kazakhstan (CAO RK)

Dominant Chapter–Term Intersection

To highlight the primary lexical orientation of the most content-dense chapters, a “chapter × term” intersection analysis was conducted.

For each chapter, the thematic category with the highest frequency was identified as its dominant lexical field. Among the ten chapters exhibiting the greatest overall lexical density, the field размер (“measure” or “amount”) emerges as the most recurrent leading term, especially in sections regulating trade, finance, and administrative enforcement. This prevalence reflects the quantitative character of sanction-related provisions, where penalties are systematically defined through monetary or procedural measures.

Other dominant categories include санкции (sanctions), лицо (person), and казахстан, the latter appearing primarily in chapters devoted to procedural jurisdiction and institutional authority.

The pattern illustrates that the CAO RK emphasizes proportionality and standardization of sanctions, embedding numerical evaluation directly into the language of legal accountability (Figure 4).



Figure 4 – Dominant thematic category in the ten most content-dense chapters of the Code of Administrative Offenses of the Republic of Kazakhstan (CAO RK)

Article Length Distribution

The dual-panel histogram (Figure 5) provides a detailed view of the lexical structure of the Code of Administrative Offenses of the Republic of Kazakhstan (CAO RK).

The left panel displays the overall distribution of article lengths, illustrating the extreme right-skew typical of codified legal texts.

The right panel zooms into the 0–500 word range, revealing that the majority of articles are compact, generally containing fewer than 250–300 words.

This pattern indicates that while most provisions are concise and normatively focused, a limited number of articles – primarily those describing procedural, sanctioning, or jurisdictional rules – expand substantially in length to accommodate complex legal formulations.

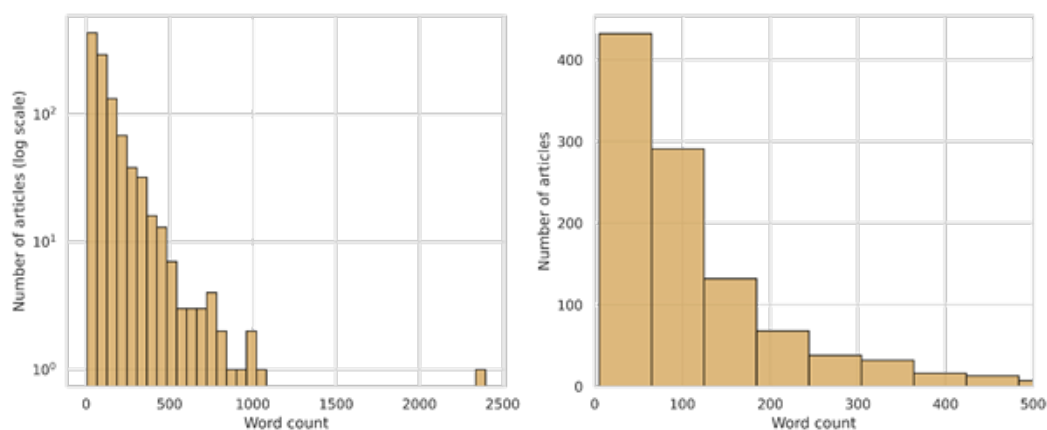


Figure 5 – Distribution of article lengths (in words) in the Code of Administrative Offenses of the Republic of Kazakhstan (CAO RK)

The left panel (logarithmic scale) shows the overall right-skewed distribution, while the right panel highlights the concentration of shorter articles below 500 words.

Amendment Dynamics

The temporal distribution of legislative amendments (Figure 6) highlights distinct cycles of revision within the Code of Administrative Offenses of the Republic of Kazakhstan.

Amendment metadata were extracted from article-level footnotes (сноски), capturing formal references to legislative acts and their years of adoption.

The histogram reveals a major legislative peak in 2017, marking the most extensive wave of modifications during the observed period. This surge corresponds to a comprehensive administrative reform aimed at refining sanctioning procedures and institutional responsibilities across multiple chapters of the Code. Moderate amendment activity is also observed in 2014–2015, reflecting the transitional phase following the enactment of the CAO RK, when the legal system was adjusting to the new codified framework. Subsequent years (2018–2022) display more selective revision patterns, including targeted changes to procedural and jurisdictional provisions.

Although the COVID-19 pandemic (2020–2021) did not produce a large quantitative spike in amendments, the updates introduced during this period often addressed public health enforcement, mobility regulation, and emergency governance, signaling a functional adaptation of administrative law to crisis conditions.

The mild uptick in 2025 suggests the beginning of another adjustment cycle aligned with ongoing modernization efforts.

Overall, the amendment dynamics illustrate a nonlinear but adaptive evolution of Kazakhstan’s administrative legislation, alternating between large-scale reform phases and incremental fine-tuning periods.

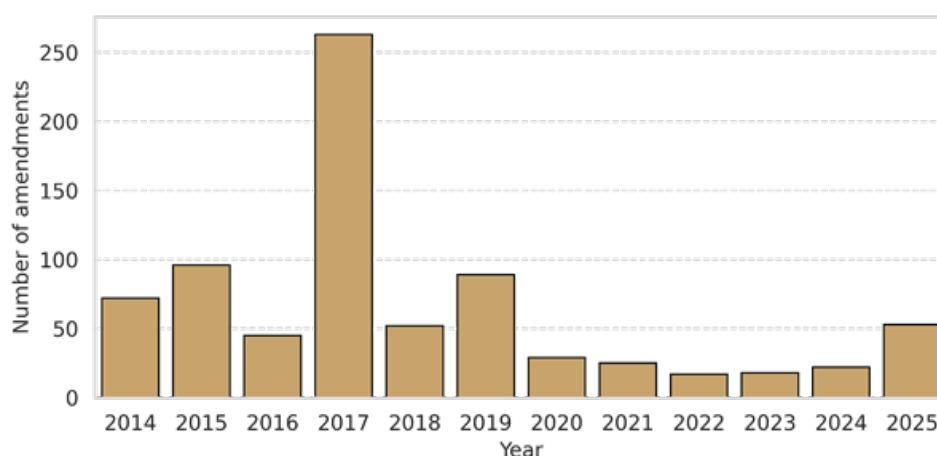


Figure 6 – Distribution of amendment years across articles of the Code of Administrative Offenses of the Republic of Kazakhstan (CAO RK)

Topic Modeling of Legal Texts

To uncover latent semantic structures within the Code of Administrative Offenses of the Republic of Kazakhstan (CAO RK), a topic modeling approach was applied using the Latent Dirichlet Allocation (LDA) algorithm. This probabilistic model identifies recurring lexical patterns by representing each article as a mixture of topics and each topic as a distribution over words. The goal was to capture underlying conceptual domains of administrative regulation that may not be explicitly defined by the Code’s formal hierarchy.

The model was trained with five latent topics ($n = 5$), selected empirically to optimize semantic coherence and interpretability. The ten most probable terms for each topic, together with their domain-specific interpretations, are presented in Table 2. These representative term clusters served as the basis for qualitative labeling of topics according to their legal and contextual significance.

The five topics collectively reflect the dual nature of the CAO RK’s discourse – a balance between quantitative sanctioning logic (Topics 1, 2, and 4) and institutional-procedural governance (Topics 3 and 5).

Table 2 – Topical structure of the CAO RK identified by LDA (Five-Topic Model)

Topic	Representative Terms	Interpretation
1	месячный, расчётный, показатель, штраф, влечь, организация, размер, лицо, республика, казахстан	Quantitative and procedural norms, focusing on the formal specification of sanctions and fiscal measures.
2	крупный, влечь, штраф, предпринимательство, субъект, размер	Economic and entrepreneurial regulation, emphasizing business responsibility and liability.
3	правонарушение, закон, действие, административный, орган, статья, часть	General principles and institutional structure of administrative law, describing legal actions and authorities.
4	расчётный, месячный, предусмотреть, статья, настоящий, штраф	Legislative and prescriptive formulations typical of codified sanctioning clauses.
5	кодекс, суд, должностной, орган, постановление, производство, правонарушение, дело	Judicial and procedural administration, covering courts, officials, and adjudicatory processes.

This thematic composition suggests that the Code’s linguistic architecture is shaped by a systematic concern with measurement, enforcement, and procedural formality, rather than narrative or interpretive exposition.

In other words, the CAO RK encodes legal responsibility through a lexically rigid framework of proportionality and administrative control.

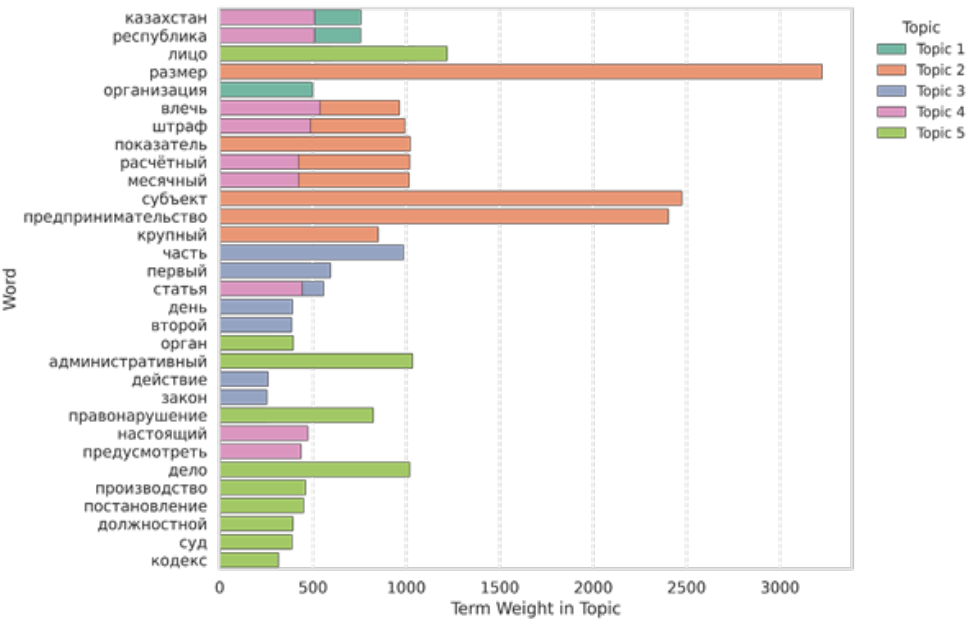


Figure 7 – Distribution of key terms across topics identified by Latent Dirichlet Allocation (LDA) in the Code of Administrative Offenses of the Republic of Kazakhstan (CAO RK)

Interactive Topic Visualization

To further explore the semantic topology of the Code of Administrative Offenses of the Republic of Kazakhstan (CAO RK), an interactive intertopic distance map was generated using pyLDavis (Figure 8). This visualization projects the latent topics produced by the LDA model into a two-

dimensional semantic space via multidimensional scaling, allowing an intuitive assessment of topic proximity, distinctiveness, and lexical dominance.

The left panel displays the spatial arrangement of the five inferred topics, where the size of each circle represents its marginal proportion in the corpus.

The relatively large and well-separated clusters indicate that the LDA model captured thematically coherent and linguistically distinct domains of administrative regulation.

Topic 2 (highlighted in red) occupies the greatest share—approximately 29% of the total token space – and centers on vocabulary related to *предпринимательство* (“entrepreneurship”), *субъект* (“subject”), and *размер* (“measure”), corresponding to the economic and sanction-quantification dimension of the Code.

The right panel ranks the top-30 most relevant terms for the selected topic according to the relevance metric $\lambda = 1.0$, balancing frequency and exclusivity.

The prominence of terms such as *предпринимательство*, *штраф*, and *организация* confirms that the model effectively isolates the language of regulatory enforcement and business accountability.

Taken together, the intertopic map reveals a structured semantic architecture, where distinct lexical clusters correspond to sanctioning mechanisms, procedural norms, institutional roles, and economic regulation.

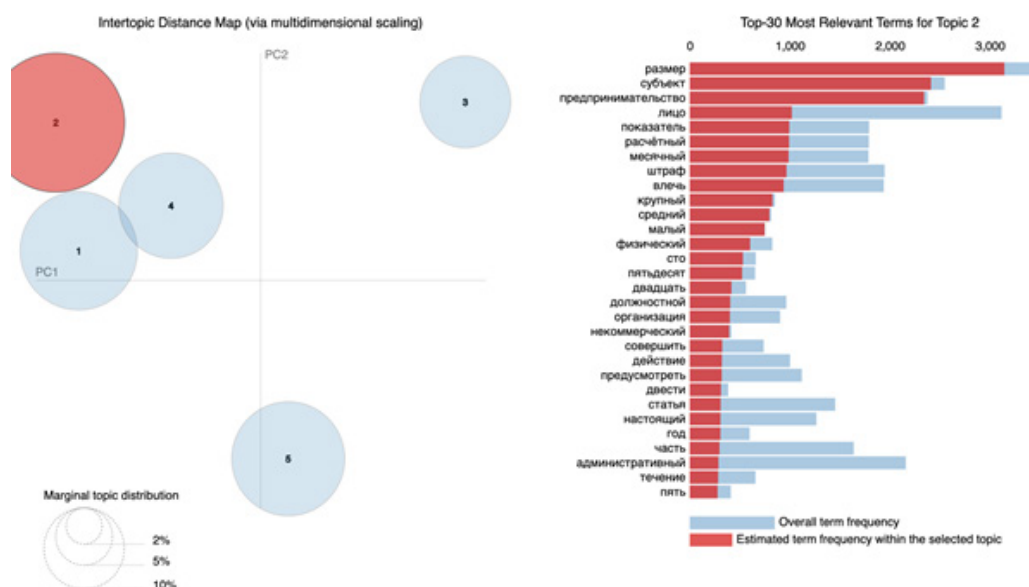


Figure 8 – Intertopic distance map and term relevance visualization for the Code of Administrative Offenses of the Republic of Kazakhstan (CAO RK), generated via Latent Dirichlet Allocation (LDA) and pyLDAvis

Discussion

The linguistic analysis of the Code of Administrative Offenses of the Republic of Kazakhstan (CAO RK) reveals that its textual architecture distinctly embodies the normative-punitive nature of administrative law.

The overwhelming prevalence of vocabulary related to sanctions (*санкции*) and responsibility (*ответственность*) indicates that the Code primarily operates as an instrument for the definition, categorization, and enforcement of administrative penalties.

At the same time, the frequent occurrence of terms such as *лицо* (“person”) and *субъект* (“subject”) highlights the Code’s focus on the individualization of legal liability, consistent with the modern principle of personal accountability in administrative justice.

In contrast, the relative scarcity of judicial lexemes – such as *суд* (“court”) and *жалоба* (“appeal”) – underscores the subsidiary role of judicial institutions, where administrative enforcement and executive oversight remain dominant.

From a broader linguistic perspective, this quantitative profile confirms that the CAO RK is not merely a collection of prohibitive norms but a codified linguistic framework through which state authority is operationalized via structured legal terminology.

The findings therefore complement traditional doctrinal interpretations by showing how lexical hierarchies mirror institutional hierarchies within Kazakhstan’s system of administrative governance.

Recent advances in Kazakh computational linguistics reinforce the relevance of this approach. The development of large-scale instruction-tuning datasets for government and legal domains, together with corpora such as KazQAD and the Aligned Kazakh–Russian Parallel Corpus, demonstrates growing national capacity for data-driven linguistic analysis [6–8]. However, these initiatives primarily address general or bilingual text processing, whereas the present study focuses specifically on formal legislative discourse, providing a complementary structural and thematic dimension to the evolving landscape of Kazakh-language NLP. This alignment between legal linguistics and computational modeling marks an important step toward creating comprehensive resources for automated legal analytics and low-resource language modeling in Kazakhstan.

Conclusion

This study employed quantitative linguistic analysis to examine the Code of Administrative Offenses of the Republic of Kazakhstan (CAO RK), uncovering both structural and thematic regularities within its legal discourse. The findings reveal a marked predominance of sanction- and responsibility-related vocabulary, underscoring the punitive and regulatory orientation of the Code and its function as a linguistic instrument of state authority.

Beyond descriptive insights, the research demonstrates the applicability of computational text analysis to the study of legal language – offering a scalable, data-driven framework for comparative legal linguistics, corpus-based jurisprudence, and the modernization of legislative analytics in Kazakhstan.

These results complement emerging initiatives in Kazakh natural language processing, such as the development of instruction-tuned datasets for legal and governmental text, the KazQAD question-answering corpus, and the Kazakh–Russian parallel legal corpus [6–8].

Future research will extend this approach to other codified acts—such as the Civil, Criminal, and Tax Codes – to explore differences in lexical density, thematic diversity, and cross-referential complexity, thereby advancing a data-driven understanding of national legal language systems.

In this broader context, the present work contributes a structured analytical pipeline and an annotated legislative corpus that can support future efforts toward Kazakh legal language modeling and semantic text alignment across legal codes.

Looking ahead, the framework developed here remains fully compatible with LLM-based semantic enrichment – including contextual classification, summarization, and translation consistency evaluation.

Future research will extend this approach to other codified acts – such as the Civil, Criminal, and Tax Codes – to explore differences in lexical density, thematic diversity, and cross-referential complexity, thereby advancing a data-driven understanding of national legal language systems and their evolution.

Limitations and Future Directions

While the present analysis provides a comprehensive quantitative overview of the Code of Administrative Offenses of the Republic of Kazakhstan (CAO RK), several methodological and interpretive limitations should be acknowledged. First, the study is confined to a single legal corpus, which constrains cross-institutional generalization. Extending the approach to other Kazakhstani codes and to comparative datasets from neighboring jurisdictions would enable a broader understanding of regional legal-linguistic patterns.

Second, the analysis relies primarily on surface-level lexical statistics and topic modeling, which, although informative, do not fully capture syntactic dependencies, modal constructions, or intertextual citation structures that play a central role in legal semantics. Incorporating dependency parsing, named-entity recognition, and semantic role labeling in future work would enhance interpretive depth.

Third, while the LDA-based topic modeling revealed coherent thematic clusters, it remains probabilistic rather than interpretive; integrating these findings with expert legal annotation could bridge the gap between computational and doctrinal perspectives.

Future research should also explore temporal evolution by tracking linguistic change across successive amendments, and should consider applying embedding-based language models (e.g., BERT or RoBERTa trained on legal corpora) to quantify semantic shifts within Kazakhstan's evolving legal framework.

REFERENCES

- 1 Theory and Methodology of the World's National Linguistic Corpora. *Linguistics Journal of Eurasia*, 14(3), 33–45 (2022).
- 2 Tokatov, R.A., Akimzhanova, M.T. On the accuracy of the texts of the Civil Code of the Republic of Kazakhstan (General Part) in the Kazakh and Russian languages. *Bulletin of L.N. Gumilyov Eurasian National University. Law Series*, 3, 135–141 (2021). <https://doi.org/10.31489/2021i3/135-141>.
- 3 Ilyassova, G.A. Problems of ensuring authenticity of texts in Kazakh and Russian in the Civil Procedure Code of the Republic of Kazakhstan. *Bulletin of L.N. Gumilyov Eurasian National University. Law Series*, 3, 71–78 (2022). <https://doi.org/10.31489/2022i3/71-78>.
- 4 Ilyassova, G.A. Issues of application of terms in the state language in civil legislation (according to the text of the special part of the Civil Code of the Republic of Kazakhstan). *Vestnik Akademii Upravleniya*, 69(2), 123–134 (2023). <https://doi.org/10.47649/vau.2023.v69.i2.14>.
- 5 Zhanzhigitov, S.Zh. Linguistic strategies of legal communication in digital environments: The case of the PravMedia online forum. *Bulletin of L.N. Gumilyov Eurasian National University. Philology Series*, 3, 41–49 (2024). <https://doi.org/10.31489/2024ph3/41-49>.
- 6 Yeshpanov, R., Efimov, P., Boytsov, L., Shalkarbayuli, A., Braslavski, P. KazQAD: Kazakh open-domain question answering dataset. *Proceedings of the 2024 Joint International Conference on Computational Linguistics, Language Resources and Evaluation (LREC–COLING 2024)*. Torino, Italy: European Language Resources Association (ELRA), 2024, pp. 9645–9656.
- 7 Khairova, N., Kolesnyk, A., Mamyrbayev, O., Mukhsina, K. The Aligned Kazakh–Russian Parallel Corpus Focused on the Criminal Theme. *Proceedings of the International Conference on Computational Linguistics and Intelligent Systems (COLINS)*, 2019, pp. 116–125.
- 8 Baisalov, A., Kenzhegulov, Y., Alimzhanova, Z. Instruction tuning on public government and cultural data for low-resource language: A case study in Kazakh. *Proceedings of the 2024 Conference on Computational Linguistics for Low-Resource Languages*, 2024 (Preprint available on arXiv).
- 9 Formation of the State Language as the Language of the Law. *Bulletin of Law and State*, 2, 56–68 (2022).
- 10 Kolesnik, A., Khairova, N. Use of linguistic criteria for estimating the quality of Wikipedia articles. *Proceedings of the 1st International Conference on Computational Linguistics and Intelligent Systems (COLINS)*, 2017, pp. 207–215.
- 11 Khairova, N., Mamyrbayev, O., Rizun, N., Razno, M., Ybytayeva, G. A parallel corpus-based approach to crime event extraction for low-resource languages. *IEEE Access*, 11, 54093–54111 (2023).
- 12 Yeshpanov, R., Varol, H.A. KazSAnDRA: Kazakh sentiment analysis dataset of reviews and attitudes. *Proceedings of the 2024 Joint International Conference on Computational Linguistics, Language Resources and Evaluation (LREC–COLING 2024)*, Torino, Italy: European Language Resources Association (ELRA), 2024.
- 13 Karipova, A., Serikbekova, S., Aralbekov, G., Tulegaliyeva, Zh., Sarsenova, A. Comparative analysis of administrative liability for driving while intoxicated in the Commonwealth of Independent States. *Hrvatska i komparativna javna uprava. Croatian and Comparative Public Administration*, 24(4), 889–910 (2024).

14 Drápal, J., Westermann, H., Savelka, J. Using large language models to support thematic analysis in empirical legal studies. Proceedings of the Thirty-sixth Annual Conference on Legal Knowledge and Information Systems (JURIX 2023). Maastricht, The Netherlands: IOS Press, 2023, pp. 65–74.

15 Malik, V., Sanjay, R., Guha, S.K., Hazarika, A., Nigam, S.K., Bhattacharya A., Modi A. Semantic segmentation of legal documents via rhetorical roles. Proceedings of the Natural Legal Language Processing Workshop (NLLP 2022). Abu Dhabi, United Arab Emirates (Hybrid): Association for Computational Linguistics, 2022, pp. 132–142.

16 Niekler, A., Wiedemann, G., Heyer, G. Leipzig Corpus Miner: A text mining infrastructure for qualitative data analysis. Proceedings of the Terminology and Knowledge Engineering Conference (TKE 2014). Berlin, Germany, 2014, pp. 441–450.

¹*Мұхсимбаев Б.,

докторант, ORCID ID: 0009-0008-4606-3628,

*e-mail: b.mukhsimbaev@kbtu.kz

¹Пак А.,

PhD, профессор, ORCID ID: 0000-0002-8685-9355,

e-mail: a.pak@kbtu.kz

¹Куралбаев А.,

докторант, ORCID ID: 0009-0001-0811-5385,

e-mail: a.kuralbaev@kbtu.kz

¹Қазақстан-Британ техникалық университеті, Алматы қ., Қазақстан

ҚАЗАҚСТАН РЕСПУБЛИКАСЫНЫҢ ӘКІМШІЛІК ҚҰҚЫҚ БҰЗУШЫЛЫҚ ТУРАЛЫ КОДЕКСІН ЛЕКСИКАЛЫҚ ЖӘНЕ ТАҚЫРЫПТЫҚ ТАЛДАУҒА АРНАЛҒАН ЕСЕПТЕУ КОНВЕЙЕРІ

Аңдатпа

Бұл зерттеу Қазақстан Республикасының Әкімшілік құқық бұзушылықтар туралы кодексі (ӨҚБТК, K1400000235) мәтінінің автоматтандырылған лингвистикалық және құрылымдық талдауына арналған есептеуіш талдау жолын ұсынады. Ұсынылған жұмыс процесі деректерді жинау, мәтінді алдын ала өңдеу, токенизация, кілт сөздерді анықтау, семантикалық топтастыру және визуализацияны қамтиды. Бұл кезеңдер Python тіліндегі табиғи тілдерді өңдеу (NLP) және статистикалық әдістерді біріктіреді. Ұсынылған жүйе лексикалық, тақырыптық және сандық лингвистикалық талдауларды бірізді тізбекке біріктіріп, Кодекстің иерархиялық құрылымы (бөлімдер, тараулар және баптар) бойынша жиілік үлестірімдерін, семантикалық өрістер мен жасырын тақырыптарды анықтауға мүмкіндік береді. ӨҚБТК корпусын талдау бірнеше ерекше тілдік заңдылықтарды анықтады: санкциялар мен жауапкершілікке қатысты сөздердің (айыппұл, жауапкершілік, құқық бұзушылық) басым болуы, экономикалық және рәсімдік құқық бұзушылықтарға арналған тарауларда жоғары лексикалық тығыздықтың байқалуы, сондай-ақ әкімшілік құқықтың нормативтік-жазалаушылық сипатын бейнелейтін тақырыптық шоғырланулар. Жиілік гистограммалары, тақырыптық жылу карталары және тақырыптық карталар сияқты визуализация әдістері заң мәтіндерін сандық тұрғыдан зерттеудің әлеуетін көрсетеді. Жалпы алғанда, ұсынылған әдістеме салыстырмалы құқықтық лингвистика, заңнаманы автоматты түрде мониторингілеу және Қазақстандағы құқықтық аналитиканы жаңғырту үшін ауқымды негіз қалайды.

Тірек сөздер: әкімшілік құқық, заң мәтінін талдау, табиғи тілді өңдеу, есептеуіш құқықтық лингвистика, жиілік талдауы, тақырыптық модельдеу, құқықтық информатика.

¹Мухсимбаев Б.,

докторант, ORCID ID: 0009-0008-4606-3628,

e-mail: b.mukhsimbaev@kbtu.kz

¹Пак А.,

PhD, профессор, ORCID ID: 0000-0002-8685-9355,

e-mail: a.pak@kbtu.kz

¹Куралбаев А.,

докторант, ORCID ID: 0009-0001-0811-5385,

e-mail: a.kuralbaev@kbtu.kz

¹Казахстанско-Британский технический университет, г. Алматы, Казахстан

ВЫЧИСЛИТЕЛЬНЫЙ КОНВЕЙЕР ДЛЯ ЛЕКСИЧЕСКОГО И ТЕМАТИЧЕСКОГО АНАЛИЗА КОДЕКСА ОБ АДМИНИСТРАТИВНЫХ ПРАВОНАРУШЕНИЯХ РЕСПУБЛИКИ КАЗАХСТАН

Аннотация

В статье представлен вычислительный конвейер для автоматизированного лингвистического и структурного анализа юридических текстов на примере Кодекса Республики Казахстан об административных правонарушениях (КоАП РК, К1400000235). Предложенный рабочий процесс объединяет этапы сбора данных, предобработки текста, токенизации, извлечения ключевых слов, семантической кластеризации и визуализации с применением методов обработки естественного языка (NLP) и статистического анализа на Python. Разработанный конвейер сочетает лексический, тематический и количественный лингвистический анализ в единую последовательную систему, что позволяет выявлять частотные распределения, семантические поля и скрытые темы в иерархической структуре Кодекса (разделы, главы, статьи). Анализ корпуса КоАП РК выявил ряд характерных языковых закономерностей: преобладание лексики, связанной с санкциями и ответственностью (штраф, ответственность, правонарушение), повышенную лексическую плотность в главах, регулирующих экономические и процессуальные правонарушения, а также тематические кластеры, отражающие нормативно-карательную направленность административного права. Визуализационные методы, такие как частотные гистограммы, тематические тепловые карты и топик-карты, демонстрируют потенциал конвейера для количественного исследования законодательного языка. В целом представленная методология формирует масштабируемую основу для сравнительной юридической лингвистики, автоматизированного мониторинга законодательства и модернизации правовой аналитики в Казахстане.

Ключевые слова: административное право, анализ юридических текстов, обработка естественного языка, вычислительная юридическая лингвистика, частотный анализ, тематическое моделирование, правовая информатика.

Article submission date: 07.11.2025

UDK 004.9; 005.8; 332.1
IRSTI 20.18.15; 20.15.12; 06.61.53

<https://doi.org/10.55452/1998-6688-2025-22-4-244-253>

¹***Satybaldiyeva R.Zh.**,

Cand. Tech. Sc., Associate Professor, ORCID ID: 0000-0002-0678-7583,

*e-mail: r.satybaldiyeva@satbayev.university

¹**Kalpeyeva Z.B.**,

PhD, Associate Professor, ORCID ID: 0000-0002-4970-3095,

e-mail: zh.kalpeyeva@satbayev.university

¹**Uskenbayeva R.K.**,

Cand. Tech. Sc., Professor, ORCID ID: 0000-0002-8499-2101,

e-mail: r.k.uskenbayeva@satbayev.university

¹**Sarkambayeva Sh.G.**,

PhD, Associate Professor, ORCID ID: 0000-0001-8509-3688,

e-mail: sh.sarkamabyeva@satbayev.university

¹**Razaque A.**,

PhD, Professor, ORCID ID: 0000-0003-0409-3526,

e-mail: r.abdul@satbayev.university

¹**Elle V.Zh.**,

MSc, ORCID ID: 0000-0002-3863-942X,

e-mail: v.elle@satbayev.university

¹Satbayev University, Almaty, Kazakhstan

BENCHMARKING BEST PRACTICES OF TOP-RANKED CITIES FOR HIGH LIVABILITY AND CITIZEN PARTICIPATION

Abstract

This article presents a comparative analysis of international benchmarking systems and their application to assessing urban livability and citizen engagement. The research examines key global indices – Economist Intelligence Unit (EIU), Mercer, PwC, TUWIEN Smart City Model, MIT Treepedia, and the National League of Cities (NLC) – to identify dominant trends and classification models of city benchmarking. Four major types of benchmarking practices were defined: multi-factor indices, single-indicator rankings, thematic analytical reviews, and diagnostic metrics. An empirical case study of Almaty, Kazakhstan, demonstrates the adaptation of international practices to the local context. Between 2020 and 2023, over 1,500 participatory projects were implemented under the Participatory Budget program, primarily in urban greening, infrastructure, and public safety. The findings show that digital governance platforms (Open Almaty, iKomek, Almaty Urban Center) enhance civic participation but remain limited by unequal digital access. The study concludes that benchmarking serves as an effective governance and evaluation tool for improving urban livability and inclusiveness. The Almaty case illustrates the potential of combining global best practices, data-driven governance, and participatory approaches to foster sustainable urban transformation.

Keywords: city benchmarking, livability, citizen participation, smart city, sustainability, digital governance, digital services.

Introduction

With the development of technology and artificial intelligence, methods for assessing the comfort of living in an urban environment are becoming more accurate and diverse. In these conditions, issues of ensuring a comfortable urban environment and public participation in governance are becoming a

priority for local authorities, researchers, and residents themselves. The emergence of global indices and rankings, such as the Global Liveability Index from the Economist Intelligence Unit, the Quality of Living index from Mercer, and the Smart City Index, has intensified the desire of cities not only to improve their internal indicators but also to compete for positions in international comparisons. City rankings are interpreted as benchmark assessments and represent a comparative analysis based on benchmark indicators.

Benchmarking, as a tool for analyzing and comparing urban strategies, has become widespread in urban planning practice. It allows for the identification of best practices, the adaptation of successful models, and the assessment of progress in key areas: infrastructure, digitalization, environmental sustainability, and social inclusion. Alongside this, the importance of citizen involvement in decision-making processes is growing, especially through digital interaction mechanisms.

For Kazakhstan, these issues are particularly relevant in the context of modernizing urban management and transitioning to the principles of a «smart city.» This article examines international approaches to assessing the quality of the urban environment and citizen participation, and also presents an empirical case study of the city of Almaty as an example of the local adaptation of global practices. It analyzes how the city applies benchmarking tools, implements digital services, and develops feedback mechanisms with the population to improve the quality of life and form an inclusive urban policy.

Materials and methods

Benchmarking is interpreted as a tool for comparative assessment of urban efficiency. More than 500 urban indices are actively used in policy and academic practice [1]. The works analyze digital technologies as a means of increasing citizen participation [2–4]. The most common are multi-factor indices that take into account a combination of parameters related to sustainability and quality of life. Local studies in Kazakhstan emphasize the importance of institutional conditions for the successful localization of global practices [5, 6].

The study uses a mixed approach: theoretical analysis of international indices as EIU, Mercer, GVI, etc., a comparative review of benchmarking methodologies, and a case study of the city of Almaty. Empirical data on Almaty includes an analysis of «Participatory Budget» projects, digitalization initiatives, and environmental transformations of the urban environment during the period 2020–2023 [17].

Modern approaches to urban benchmarking and its applied significance

The comfort of the urban environment is a key category of modern urban planning. The report defines comfort as the integration of infrastructure solutions, housing accessibility, ecology, and the quality of public spaces [7].

Big data, «smart» city technologies, GIS, and crowdsourcing provide new opportunities for urban planners and authorities to create more comfortable living conditions. The use of this knowledge base is necessary for smart management. However, to fully realize these opportunities, it is necessary to improve data accessibility and develop analytical tools that will consider the individual needs of each city. Data analytics, along with a flexible and futuristic strategy, should play a very important role in the formation and development of a metropolis. The participation of local communities in these processes is the basis of their management in smart cities. The National League of Cities (NLC) report presents the concept of «livable communities.»

The literature review presents the following types of benchmarking (see Table 1):

1. Indices and index reports: These are publicly available multi-factor assessment schemes that rank cities from best to worst based on a comprehensive assessment of several interrelated dimensions. Examples include the Economist Intelligence Unit index, the Z-Yen Global Financial Centres Index, and the Sportcal Global Sports Cities Index. They account for approximately 64% of all benchmarks and are widely used for strategic planning, comparative analysis, and positioning cities on the international stage [10].

2. Rankings based on a single indicator: These record the performance of each city, usually from best to worst, based on a single indicator. Examples include the Rome2Rio Global Connectivity Ranking (direct international flights) and the IQAir Air Quality and City Pollution Ranking (particulate matter density). They account for approximately 23% of the total number of benchmarks. They are easy to interpret but do not provide a comprehensive picture of the quality of life [11].

3. Less common are analytical reviews and thematic reports. These reports assess the performance and readiness of cities on specific urban themes. They are less visible, may require a subscription, and contain comparative components as an additional layer of analysis rather than the main focus. They may group or typologize cities instead of strictly ranking them. They represent 9% of benchmarks, are valuable for their depth of analysis, and are used by specialists in developing urban policy programs Best Performing Cities 2024, World Best Cities Report 2024 [12, 13].

4. Measurement metrics: These consider a city's performance in one specific dimension, usually based on a single or primary data point. They are the least common, accounting for 4% of the indicators. Such metrics serve as diagnostic tools for monitoring the effectiveness of individual areas of urban infrastructure (for example, energy efficiency or accessibility of public transport) Resilient Cities Index 2023 [14].

Table 1 – Types of Comparative Analysis and Their Practical Application

No.	Type of Comparative Analysis	Characteristics and Content	Examples	Percentage share	Practical Purpose
1	Indices and Index Reports	Multifactorial complex assessment schemes that rank cities based on a combination of indicators: economy, environment, safety, culture.	Economist Intelligence Unit, Z-Yen Global Financial Centers Index, Sportcal Global Sports Cities Index	≈64%	Strategic planning, diagnosis of strengths and weaknesses, intercity comparisons.
2	Rankings based on a single indicator	Measure performance based on one factor (transport, pollution, cost of living, etc.).	Rome2Rio Global Connectivity, IQAir Air Quality Ranking	≈23%	Monitoring of individual aspects of the urban environment (ecology, transport, etc.).
3	Analytical reviews and thematic reports	Studies of groups of cities with elements of comparison and typification. Often limited access.	Urban Sustainability Reports, City Readiness Reports	≈9%	Assessment of cities' readiness for innovation and sustainable development.
4	Evaluation measures (metrics)	Assessment of effectiveness based on individual parameters without a final ranking.	Urban Mobility Index, Green View Index (MIT Treepedia)	≈4%	Used as tools for evaluating individual areas of urban policy.

Approximately 20% of all benchmarks are produced by research, analytical, and academic organizations, while in other cases, academics contribute. Notable examples include: The Globalization and World Cities Research Network (GaWC), MIT's Senseable City Lab.

Since 1998, GaWC has mapped and ranked cities based on their «network connectivity,» meticulously studying the centrality of cities based on the concentration of advanced producer services. Their ranking of global cities (alpha, beta, gamma categories) has significantly influenced the strategic plans of local governments, multilateral agencies, and the private sector, demonstrating the ability of academic analysis to inform global urban practice. However, GaWC emphasizes that the ranking is a small part of their model, often oversimplified by the media and politics focused on «best» and «worst» lists.

MIT's Senseable City Lab developed the Green View Index (GVI), later named «Treepedia,» which assesses the greenness of a city using Google Street View panoramas. Its partnership with the World Economic Forum for dissemination attracted significant attention, as cities used its comparative overview to strengthen advocacy for nature-based solutions. The GVI/Treepedia results have undergone academic peer review, and the team provides open access for reproducibility.

In addition, there are benchmarks developed by private organizations, such as the Arup-Rockefeller «Resilience Index» and the PwC «Cities of Opportunity» ranking.

Thus, the presented classification demonstrates that urban benchmarking is gradually evolving from narrowly focused ratings to comprehensive and interdisciplinary management tools, in which digital data, citizen participation, and sustainable development play an increasingly important role.

2.2 Case analysis of improving the comfort of the urban environment and citizen engagement.

In recent years, Almaty has been actively developing approaches to a «human-centered» and sustainable city, using both elements of international experience and local initiatives. Almaty is the largest city in Kazakhstan by population (more than 2 million people), the former capital and a key economic and cultural center of the country. Almaty is actively developing «smart city» components and citizen participation programs. Since 2020, the «People's Participation Budget» mechanism has been introduced, within the framework of which more than 1,500 projects have been implemented [15]. Digital services are also being developed: Open Almaty, iKomek, Almaty Urban Center.

Since 2016, Almaty has been implementing the «Development of a Comfortable Urban Environment» program, within the framework of which the city authorities focus on: the development of pedestrian infrastructure (reconstruction of Panfilov, Zhibek Zholy, and Abay streets); expanding the public transport network and dedicated lanes for buses; the introduction of electric buses and charging stations; improvement of public spaces and parks (Gorky Central Park, district-level squares); greening the city and restoring irrigation canals (surface water drainage system – a unique historical element).

Studies by local urbanists show that the quality of the urban environment in Almaty is improving, especially in the central districts [7, 8]. However, the problem of unevenness remains – in residential areas (Alatau, Nauryzbay districts), the infrastructure lags significantly behind. The «People's Participation Budget» (PPB) allows residents to participate in the distribution of a part of the district budget (more than 11 billion tenge in 2022).

Figure 1 demonstrates the growth in the number of projects implemented within the framework of the PPB. A clear positive trend is visible, reflecting both increased budget funding at the district level and increased awareness and activity among the population. A particularly noticeable increase occurred between 2020 and 2021 – likely due to an intensified information campaign and the digitalization of application submissions.

Data on PPB projects in Almaty for 2020–2023 were collected from open sources: the Open Almaty portal, official reports of the akimats (local administrations), and independent publications of the Urban Forum Kazakhstan [21].

Categories were determined based on the descriptions of applications posted on Open Almaty and reports from district akimats. The conditional distribution of projects by type is shown in Figure 2.

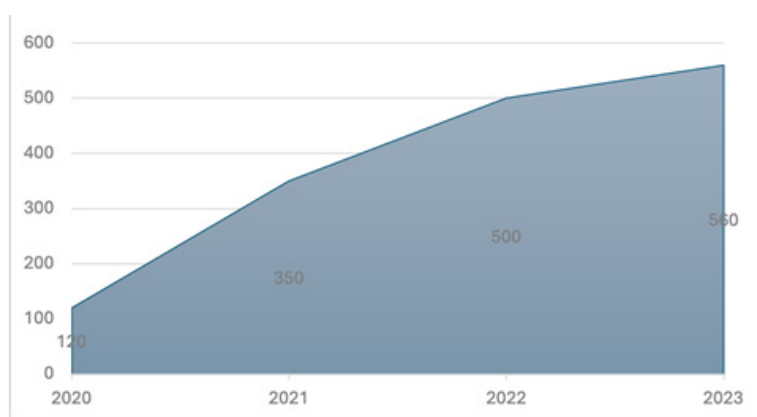


Figure 1 – Dynamics of implemented projects within the framework of the «People's Participation Budget» in Almaty, 2020–2023

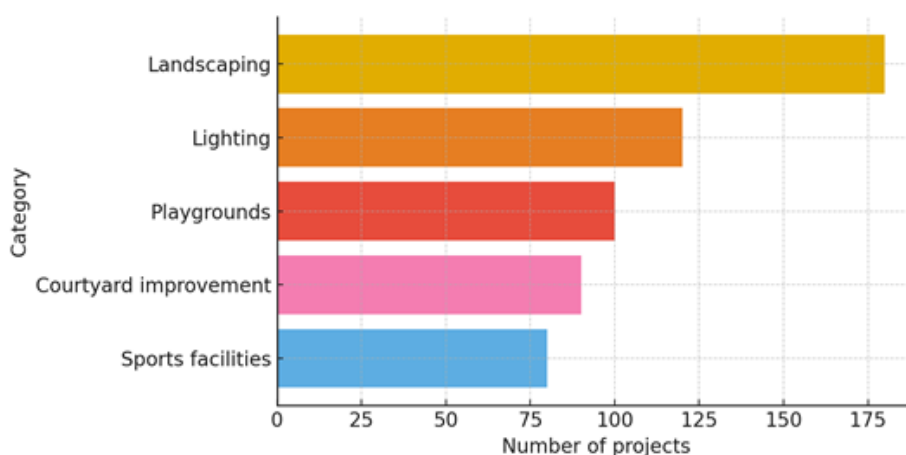


Figure 2 – Distribution of participatory budgeting projects in Almaty by category

The most popular category turned out to be landscaping projects (planting trees, installing flowerbeds, restoring irrigation canals). Lighting and playgrounds also account for a significant share – reflecting the priority of safety and accessible infrastructure in citizens' requests. The share of sports infrastructure is growing as initiatives promoting a healthy lifestyle and urban fitness in residential areas are supported.

The main features to highlight are: the Open Almaty platform is used for submitting initiatives and voting; any resident can submit a beautification project (playground, lighting, landscaping, etc.); decisions are made based on online voting among citizens. Despite a number of advantages, there are challenges. Limited public participation: mainly users aged 30–50 with a high level of digital literacy are active. Insufficient public awareness in areas with low levels of ICT access. Limited project support after implementation (for example, lack of maintenance of small architectural forms and green spaces). Nevertheless, according to the Almaty city administration, more than 1,500 projects have been implemented over the three years of the program, most of which are improvements to residential areas and landscaping [20].

Almaty is actively implementing digital services such as: Open Almaty – a single window for citizens' appeals, suggestions and complaints; iKomek – a system for rapid response to incidents (similar to «SmartCity» platforms in other countries); Almaty Urban Center – an analytical center

engaged in the analysis of the urban environment and public opinion [16–17]. According to a survey by the Urban Forum Kazakhstan, about 47% of surveyed citizens positively assess interaction with the city’s digital services, but note a lack of real feedback [18].

The implementation of «green building» principles within the framework of the «Green Almaty» program reflects the city’s commitment to sustainable urban development. A key focus has been the expansion of cycling infrastructure, including the construction of dedicated bicycle lanes and the operation of the AlmatyBike bicycle rental system. However, the level of use remains relatively low due to poor route connectivity and limited integration with other modes of transport [19]. Positive steps have also been taken in the area of urban mobility: electric buses have been introduced and an air quality monitoring system has been deployed [20]. Despite these efforts, environmental risks persist – Almaty remains among the most polluted cities in the CIS in terms of PM2.5 particulate matter content, which underscores the need for comprehensive, multi-level measures at the national level [21–23].

Almaty could become a model for other Kazakhstani cities, provided that successful practices are systematically scaled up and citizen involvement in governance processes is expanded. The Almaty case demonstrates that the implementation of digital solutions and public participation through participatory budgeting increase the perception of environmental comfort. However, challenges remain: environmental risks, uneven regional development, and limited involvement of socially vulnerable groups. Compared to Copenhagen, Singapore, and Seoul, Almaty is in the formative stage of urban sustainability and requires a systemic approach.

Results and discussion

The comparative analysis of global city rankings and indices, including the Economist Intelligence Unit (EIU), Mercer, PwC, TUWIEN Smart City Model, Treepedia (MIT), and the National League of Cities (NLC), revealed that modern benchmarking systems are evolving from purely comparative rating tools toward integrated, multi-criteria analytical frameworks. These systems combine quantitative data with qualitative dimensions such as citizen satisfaction, accessibility, and governance transparency.

Trends and classification of city benchmarking

The study identified four main types of benchmarking practices commonly applied in urban governance (Table 2).

Table 2 – Classification of urban benchmarking approaches and their practical application

No.	Type of Benchmarking	Description	Examples	Practical Purpose
1	Multi-factor indices	Comprehensive evaluations integrating economic, environmental, and social parameters	EIU, Mercer, PwC	Strategic planning and diagnostics of strengths and weaknesses
2	Single-indicator rankings	Performance measurement based on a single factor	IQAir, Rome2Rio	Monitoring of specific aspects of urban systems
3	Thematic and analytical reviews	In-depth studies assessing readiness for innovation and sustainability	Milken Institute, Resonance Consultancy	Policy design and benchmarking-based recommendations
4	Metrics and assessment tools	Diagnostic indicators without final ranking outcomes	Treepedia (MIT), Urban Mobility Index	Evaluation of specific domains of urban policy

This typology corresponds with findings that benchmarking has become a strategic governance mechanism, helping cities assess policy outcomes, strengthen accountability, and identify transferable practices [1].

Empirical analysis: the case of the city of Almaty.

The empirical analysis of Almaty demonstrates the localization of international benchmarking principles in a Central Asian urban context. Between 2020 and 2023, more than 1,500 citizen-led projects were implemented through the Participatory Budget (PB) mechanism [17]. The number of approved initiatives increased from 120 in 2020 to 560 in 2023, reflecting growing civic awareness, digital literacy, and community engagement.

The most common PB categories were urban greening, playgrounds, and lighting infrastructure, representing over 60% of all proposals. These preferences align with global trends emphasizing environmental comfort, accessibility, and safety National League of Cities.

The adoption of digital governance platforms such as Open Almaty, iKomek, and Almaty Urban Center has strengthened citizen feedback channels. According to Urban Forum Kazakhstan, 47% of respondents expressed satisfaction with the performance of these platforms, while some noted the lack of detailed follow-up communication and limited inclusion of vulnerable groups [23].

Overall, Almaty is gradually implementing the “smart and human-centered city” model, combining digitalization, public participation, and sustainability initiatives. These findings are consistent, where emphasize the link between digital participation and improved urban governance quality [24, 25].

Comparison with International Best Practices

In comparison with leading cities such as Copenhagen, Singapore, and Seoul, Almaty remains at the stage of developing comprehensive urban resilience. While substantial progress has been achieved in citizen participation and digital services, persistent challenges include air pollution (PM2.5) and unequal infrastructure development across districts.

Initiatives under the Green Almaty program – including the deployment of electric buses and the expansion of the AlmatyBike network – indicate the city’s gradual alignment with international sustainable urban mobility trends. These developments reflect principles highlighted in the Resilient Cities Index and Sustainable Cities Index, emphasizing the integration of environmental and social dimensions in city performance evaluation [10, 14].

The conducted research allows the following general conclusions:

- ♦ Benchmarking is an effective diagnostic and governance tool for evaluating urban livability and policy efficiency;
- ♦ Digital governance and participatory budgeting strengthen communication between authorities and citizens, increasing trust and policy transparency;
- ♦ The adaptation of global benchmarking practices requires contextualization to local socio-economic and cultural conditions.

The case of Almaty demonstrates that the combination of digital transformation, participatory governance, and environmental modernization forms a solid foundation for developing a sustainable, inclusive, and livable urban environment in Kazakhstan.

Conclusion

City rankings have dramatically increased in popularity in media reports and are widely used in municipalities and the consulting world. They are used to justify new policies. Attention from rankings can stimulate public interest in urban policy and assist in city marketing by highlighting positive aspects.

Many scholars argue that urban benchmarking provides urban management professionals with opportunities to:

1. Assess their current situation.
2. Compare their situation with other cities.

3. Learn from the experiences of other cities in terms of policy strategies and goals.
4. More effectively prioritize funding for urban infrastructure projects.
5. Create collaborative urban networks.

The analysis showed that international and national rankings allow for identifying the strengths and weaknesses of cities, as well as developing strategies to improve the quality of life. In Kazakhstan, digitalization and citizen participation initiatives contribute to the formation of more comfortable and sustainable urban ecosystems. In the future, it is important to develop local indices that take into account the specifics of the regions and the cultural context.

Funding information

The manuscript presents the results of the research conducted within the framework of funding targeted program financing of scientific research for 2024–2026 under the IRN project BR24993051-OT-25 “Development of an intelligent city system based on IoT and data analysis”. The study was carried out with the financial support of the Science Committee of the Ministry of Science and Higher Education of the Republic of Kazakhstan.

REFERENCES

- 1 Acuto, M. Taking urban rankings seriously: Engaging with the practice of city benchmarking in global urbanism. *IJURR*. (2021).
- 2 Shin, Y., Park, J., Lee, S. A systematic analysis of digital tools for citizen participation. *Government Information Quarterly*, 41(3), 102092 (2024). <https://doi.org/10.1016/j.giq.2024.102092>. URL: <https://www.sciencedirect.com/science/article/pii/S0740624X24000467>.
- 3 Asimakopoulos, G., Chatzidimitriou, K., Theodoridis, P. Impact of Information and Communication Technologies on Democratic Processes and Citizen Participation. *Societies*, 15(2), 40 (2025). <https://doi.org/10.3390/soc15020040>. URL: <https://www.mdpi.com/2075-4698/15/2/40>.
- 4 Leclercq, B., Rijshouwer E. Enabling citizens' right to the smart city through the co-creation of digital platforms. *Urban Transformations*, 4(1), 10 (2022). <https://doi.org/10.1186/s42854-022-00030-y>. URL: <https://urbantransformations.biomedcentral.com/articles/10.1186/s42854-022-00030-y>.
- 5 E-government mechanisms to enhance the participation of citizens. *Technology in Society*, 70, 101988 (2022) <https://doi.org/10.1016/j.techsoc.2022.101988>. URL: <https://www.sciencedirect.com/science/article/pii/S0160791X22001191>.
- 6 Santini, R. Online platforms for citizen participation: meta-synthesis of case studies. *Communications*, 41(2) (2019). URL: <https://journals.openedition.org/cs/1917>.
- 7 Tleppayev, A. Digitalisation and energy: world experience and evidence of correlation from Kazakhstan. *Economic Annals-XXI*, 176(3–4), 56–64 (2019). <https://doi.org/10.21003/ea.V176-06>.
- 8 Digel, I., Mussabalina, D., Urdabayev, M., Nurmukhametov, N., Akparova, A. Evaluating development prospects of smart cities: cluster analysis of Kazakhstan's regions. *Problems and Perspectives in Management*, 20(4), 76–87 (2022). [https://doi.org/10.21511/ppm.20\(4\).2022.07](https://doi.org/10.21511/ppm.20(4).2022.07).
- 9 UN-Habitat. *World Cities Report 2020: The Value of Sustainable Urbanization*. Nairobi: United Nations Human Settlements Programme, 2020.
- 10 Corporate Knights. *2023 Sustainable Cities Index Report*. Toronto: Corporate Knights, 2023. URL: <https://www.corporateknights.com>.
- 11 World Economic Forum. *The Urban Mobility Scorecard Tool 2023*. Geneva: WEF, 2023. URL: <https://www3.weforum.org>.
- 12 Milken Institute. *Best-Performing Cities 2024*. Santa Monica: Milken Institute, 2024. URL: <https://milkeninstitute.org>.
- 13 Resonance Consultancy. *World Best Cities Report 2024*. Vancouver: Resonance Consultancy, 2024. URL: <https://www.baeventures.com>.
- 14 Economist Impact. *Resilient Cities Index 2023*. London: Economist Group, 2023. URL: <https://www.preventionweb.net>.
- 15 City of Almaty Budget Portal URL: <https://budget.open-almaty.kz> (accessed: 02.10.2025).
- 16 Open Almaty. Official portal of the city of Almaty. URL: <https://open-almaty.kz> (accessed: October 28, 2025).

17 Akimat of Almaty. Annual reports on the implementation of the “Participatory Budget” project for 2020–2023. Almaty: Official website of the Akimat, 2020–2023. URL: <https://almaty.gov.kz> (accessed: October 28, 2025).

18 Urban Forum Kazakhstan. Analytical reports on urban development and citizen engagement. URL: <https://urbanforum.kz> (02.10.2025).

19 AlmatyBike. Public Bicycle Sharing System Overview. URL: <https://almatybike.kz>.

20 Almaty City Transport Department. Transition to Electric Public Transport: Strategy and Implementation 2022–2025. Almaty: City Mobility Office. URL: <https://alatransit.kz>.

21 Akimat of Almaty. Green Almaty Program Implementation Report, 2021–2023. Almaty: Department of Natural Resources and Environmental Regulation. URL: <https://almaty.gov.kz>.

22 IQAir. 2023 World Air Quality Report. URL: <https://www.iqair.com/world-air-quality-report>.

23 Urban Forum Kazakhstan. Recommendations for sustainable urban mobility in Almaty. URL: <https://urbanforum.kz>.

24 Rahbariyanz, R. Human-Centric Smart Cities for Inclusive and Ethical Urban Development. Smart Design Policies, 1(1), 15–22 (2024). <https://doi.org/10.38027/smart-v1n1-3>.

25 Noori, N., Hoppe, T., & de Jong, M. (2020). Classifying Pathways for Smart City Development: Comparing Design, Governance and Implementation in Amsterdam, Barcelona, Dubai, and Abu Dhabi. Sustainability, 12(10), 4030. <https://doi.org/10.3390/su12104030>.

¹*Сатыбалдиева Р.Ж.,

т.ғ.к., қауымдастырылған профессор, ORCID ID: 0000-0002-0678-7583,

*e-mail: r.satybaldiyeva@satbayev.university

¹Кальпеева Ж.Б.,

PhD, қауымдастырылған профессор, ORCID ID: 0000-0002-4970-3095,

e-mail: zh.kalpeyeva@satbayev.university

¹Ускенбаева Р.К.,

т.ғ.д., профессор, ORCID ID: 0000-0002-8499-2101,

e-mail: r.k.uskenbayeva@satbayev.university

¹Сарқамбаева Ш.Г.,

PhD, қауымдастырылған профессор, ORCID ID: 0000-0001-8509-3688,

e-mail: sh.sarkamabyeva@satbayev.university

¹Разак А.,

PhD, профессор, ORCID ID: 0000-0003-0409-3526,

e-mail: r.abdul@satbayev.university

¹Элле В.Ж.,

т.ғ.м., ORCID ID: 0000-0002-3863-942X,

e-mail: v.elle@satbayev.university

¹Satbayev University, Алматы қ., Қазақстан

ӨМІР СҮРУ ДЕҢГЕЙІ ЖОҒАРЫ ЖӘНЕ АЗАМАТТАРДЫҢ ҚАТЫСУЫ ЖОҒАРЫ ҚАЛАЛАРДЫҢ ОЗЫҚ ТӘЖІРИБЕЛЕРІН САЛЫСТЫРУ

Андатпа

Бұл мақалада халықаралық эталондық жүйелердің салыстырмалы талдауы және олардың қала тұрғындарының өмір сүру мүмкіндігі мен азаматтардың қатысуын бағалауда қолданылуы қарастырылады. Зерттеу негізгі жаһандық индекстерді – Economist Intelligence Unit (EIU), Mercer, PwC, TUWien Smart City Model, MIT Treepedia және Ұлттық қалалар лигасын (NLC) – қалалардың салыстырмалы бағалауындағы басым тенденциялар мен жіктеу үлгілерін анықтау мақсатында жүргізілді. Бенчмаркинг тәжірибесінің төрт негізгі түрі айқындалды: көпфакторлы индекстер, бір индикаторлы рейтингтер, тақырыптық аналитикалық шолулар және диагностикалық көрсеткіштер. Алматы қаласы бойынша Қазақстанның эмпирикалық жағдайлық зерттеуі халықаралық тәжірибені жергілікті контекстке бейімдеу ерекшеліктерін көрсетеді. 2020–2023 жж. аралығында бюджеттік бағдарлама аясында, ең алдымен, көгалдандыру, инфрақұрылым және қоғамдық қауіпсіздік бағыттарында 1500-ден астам бірлескен жобалар іске асырылды. Нәтижелер

цифрлық басқару платформаларының (Open Almaty, iKomek, Almaty Urban Center) азаматтық белсенділікті арттыратынын, алайда цифрлық қолжетімділік теңсіздігімен шектелетінін көрсетеді. Зерттеу бенчмаркингтің қалалардың өмір сүру мүмкіндігі мен инклюзивтілігін арттыруда тиімді басқару және бағалау құралы ретінде қызмет ететінін дәлелдейді. Алматы мысалы қалалардың тұрақты трансформациясын ынталандыру үшін озық әлемдік тәжірибелерді, деректерге негізделген басқаруды және қатысушылық тәсілдерді біріктіру әлеуетін айқын көрсетеді.

Тірек сөздер: қаланы салыстыру, өмір сүру мүмкіндігі, азаматтардың қатысуы, «ақылды қала», тұрақтылық, цифрлық басқару, цифрлық сервис.

¹*Сатыбалдиева Р.Ж.,

к.т.н., ассоциированный профессор, ORCID ID: 0000-0002-0678-7583,

*e-mail: r.satybaldiyeva@satbayev.university

¹Кальпеева Ж.Б.,

PhD, ассоциированный профессор, ORCID ID: 0000-0002-4970-3095,

e-mail: zh.kalpeyeva@satbayev.university

¹Ускенбаева Р.К.,

д.т.н., профессор, ORCID ID: 0000-0002-8499-2101,

e-mail: r.k.uskenbayeva@satbayev.university

¹Саркамбаева Ш.Г.,

PhD, ассоциированный профессор, ORCID ID: 0000-0001-8509-3688,

e-mail: sh.sarkamabyeva@satbayev.university

¹Разак А.,

PhD, профессор, ORCID ID: 0000-0003-0409-3526,

e-mail: r.abdul@satbayev.university

¹Элле В.Ж.,

м.т.н., ORCID ID: 0000-0002-3863-942X,

e-mail: v.elle@satbayev.university

¹Satbayev University, г. Алматы, Казахстан

БЕНЧМАРКИНГ ГОРОДОВ ПО УРОВНЮ КОМФОРТНОСТИ ПРОЖИВАНИЯ И ВОВЛЕЧЕННОСТИ ГРАЖДАН

Аннотация

В статье представлен сравнительный анализ международных систем бенчмаркинга и их применения для оценки комфортности городской среды и гражданской вовлеченности. Исследование охватывает ключевые мировые индексы – Economist Intelligence Unit (EIU), Mercer, PwC, TUWIEN Smart City Model, MIT Treepedia и National League of Cities (NLC) – с целью выявления современных тенденций и типологии методов сравнительной оценки городов. Определены четыре основных типа бенчмаркинга: многофакторные индексы, рейтинги одного показателя, аналитические обзоры и диагностические метрики. На примере г. Алматы показана адаптация международных практик к местному контексту. В период 2020–2023 гг. в рамках программы «Бюджет народного участия» реализовано свыше 1500 проектов, преимущественно в сферах озеленения, благоустройства и безопасности. Цифровые платформы (Open Almaty, iKomek, Almaty Urban Center) способствуют повышению активности населения, однако сохраняются различия в уровне цифрового доступа. Полученные результаты подтверждают, что бенчмаркинг является эффективным инструментом управления и оценки городской среды. Опыт г. Алматы демонстрирует потенциал сочетания международных практик, цифровых технологий и общественного участия для формирования устойчивых и инклюзивных городов.

Ключевые слова: бенчмаркинг городов, комфортность, вовлеченность граждан, умный город, устойчивое развитие, цифровое управление, цифровой сервис.

Article submission date: 02.11.2025

UDK 004.89
IRSTI 20.23.25

<https://doi.org/10.55452/1998-6688-2025-22-4-254-265>

¹***Kaidullayev M.A.**,

PhD student, ORCID ID: 0009-0000-7727-9095,

e-mail: mi_kaidullayev@kbtu.kz

¹**Akzhalova A.Z.**,

PhD, Professor, ORCID ID: 0000-0002-1141-7595,

e-mail: a.akzhalova@kbtu.kz

¹Kazakh-British Technical University, Almaty, Kazakhstan

A SYSTEM ARCHITECTURE FOR DISASTER MANAGEMENT SYSTEM

Abstract

In the era of accelerating climate change and growing urban populations, the frequency and severity of natural disasters have increased significantly, posing substantial threats to infrastructure, economic stability, and human lives. Disasters, including the likes of earthquakes, floods, and hurricanes usually are the reasons for serious destruction of buildings, requiring rapid and accurate assessment to aid in emergency response and resource allocation. In light of this, the research aims to deliver a deep learning based building damage assessment model, which is a hybrid architecture consisting of Artificial Intelligence and IOT. In this paper we will examine the use of Internet of Things (IoT) and Artificial Intelligence in disaster management systems in order to improve the automation, transparency, and sustainability in smart intelligence systems. The system should collect and analyze pre-disaster and post-disaster aerial imagery to classify buildings into damage categories, i.e. from no damage to destroyed.. Also, we integrate our model into a wide disaster management system in order to make a visualization of damages on a geospatial interface, that helps the decision-makers to get a quick look at priority areas and streamline the response of disaster. This system's plan is to assist public authorities, NGOs, and first responders with quick decision making in post-disaster response times.

Keywords: disaster management, damage detection, natural disasters, system architecture, middleware, AI, IoT.

Introduction

Unforeseen disasters, including earthquakes, storms, floods, and wildfires, are constantly damaging urban centers worldwide, causing serious destruction to buildings and endangering many lives, and infrastructure. According to data from Our World in Data, between 2010 and 2022, the most significant contributors to economic damage from natural disasters as a share of GDP—were storms, earthquakes, and floods [2]. At the same time, the highest disaster-related death rates were caused by earthquakes, floods, and extreme temperatures [1]. These statistics highlight that while storms and high temperatures have notable impacts, earthquakes consistently rank among the top causes of both economic loss and human fatalities. Given their combined impact on infrastructure and human life, earthquakes are a critical area for focused research and technological intervention.

Therefore, this paper prioritizes earthquake-related disasters when developing and evaluating the proposed building damage assessment model. In 2010, January 12, an earthquake with magnitude 7.0 brought estimated 220,000+ casualties, 300,000+ injuries and massive destruction in the capital, Port-au-Prince [3]. This is mainly due to a combination of poor infrastructure, high population density, and lack of disaster preparedness. Many buildings in Port-au-Prince were poorly constructed and could not withstand seismic activity, leading to widespread collapse and entrapment [3]. Additionally, emergency services were overwhelmed and under resourced, slowing rescue efforts

and medical aid. Effective disaster management systems could have significantly reduced casualties by enforcing safer building codes, providing early warnings, training local emergency responders, and establishing clear evacuation and response plans. Investing in disaster preparedness before the earthquake could have saved thousands of lives. By using satellite or drone imagery combined with AI and machine learning, we could assess building conditions, detect structural weaknesses, and map high-risk areas. In the aftermath of the earthquake, such a system could also assist in quickly assessing damage, locating collapsed buildings, and guiding search-and-rescue teams to the hardest-hit areas, ultimately saving more lives and improving the efficiency of relief operations [19].

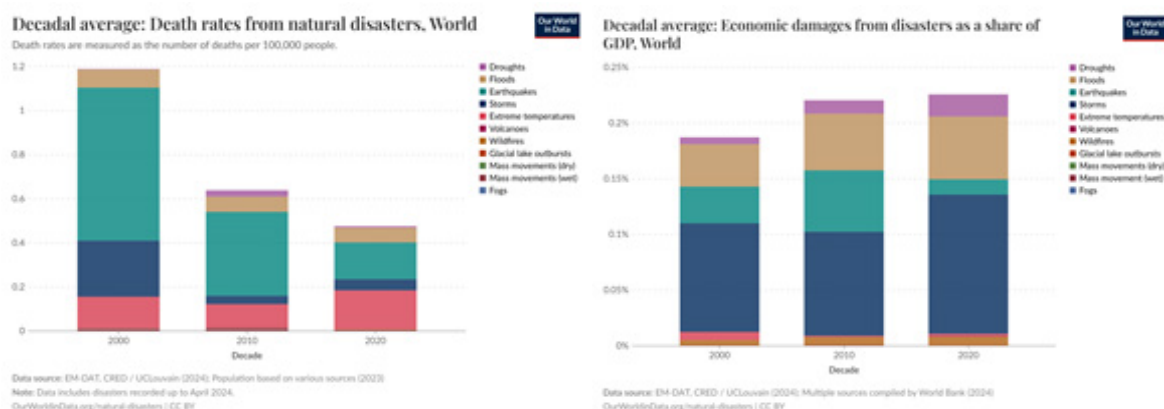


Figure 1 – Death rates and economic damage from disasters

“Disaster Information Management System (DIMS): a mechanism for effectively processing, organizing, storing and disseminating information required for disaster response and recovery, particularly in the immediate aftermath of a natural disaster event”. After the aerial photos of the sites are collected, the damaged site needs to be compared to the site at the beginning of the disaster to be able to assess damage. The Disaster Damage Map can be created with aerial photos before and right after a natural disaster that shows the affected areas. This feature in the system provides initial damage assessments thus helping prioritize search and rescue operations. Moreover, another use of it is to query building damages that happened because of the accident [4].

Materials and methods

Disaster management is commonly divided into four key stages—preparedness, response, recovery, and mitigation—each of which can be significantly enhanced by using IoT technologies [5]. The preparedness stage covers the technical issues of the readiness of measures to be taken before a calamity occurs. This is where the Internet of Things (IoT) comes in to enable continuous environmental monitoring, real-time risk assessment, and effective resource allocation. Networks of sensors, cameras, and other devices give scientists the data necessary to catch early warning signs and inform the population. For example, IoT based weather stations have displayed variables that might indicate dangerous weather, giving people enough time to prepare. In the event of disaster happening, the response phase, comes into play, which is about doing coordinated and quick actions to lower the harm. IoT is beneficial to this stage by providing the channels for quick communication between first responders, increasing situation awareness and directing the overall deployment of search and rescue operations. Drones with IoT sensors can make a rapid map of the affected areas, estimate the damage of the structures, and point the routes where rescue teams can navigate. Wearable devices also help to protect the responders due to their ability to track responders' health metrics and locations in real time [21]. Then, the recovery stage focuses on normalizing the things and redoing the damaged infrastructure. Here, the role of the IoT is to promote recovery via the efficient damage assessment process and identifying rebuilding priorities. For example, IoT-enabled structural sensors

can transmit safety data to central systems, helping officials prioritize which buildings require urgent repair. Smart grids integrated with IoT systems can also speed up energy restoration, ensuring essential services return more quickly. Finally, the mitigation and prevention stage seeks to reduce the severity of future disasters. IoT contributes by collecting and analyzing data from all phases of a disaster, improving forecasting models and guiding infrastructure improvements [5, 18].



Figure 2 – Disaster management cycle

In *Middleware for Internet of Things: A Survey*, the authors define IoT as a network of uniquely addressable objects interconnected through standard protocols and examine its key characteristics from both infrastructure and application perspectives. The architecture we are proposing encompasses both functional and non-functional requirements listed by Mohammad Abdur Razzaque in his work. The representation of the functional part, for instance, requires that middleware facilities must, without human intervention, auto-discover resources and manage resources in the most efficient way by ensuring resource optimality. It also must apply data management tasks like obtaining, filtering and aggregating, while event management that codes raw signals into useful events and enables real-time troubleshooting for analysis. In terms of non-functional requirements, the model emphasizes scalability, allowing the system to handle increasing numbers of devices through techniques such as loose coupling and virtualization. Real-time performance is equally vital, assuring the on-time delivery of services. Reliability and availability are depicted as the critical measures to perform consistently and fast recovery in case of failure [17].

We want to present a system where AI components are integrated with data sources and IoT devices. Swarna Kamal Paul in his paper called “Disaster Management through Integrative AI” demonstrates an innovative framework that combines artificial intelligence, Internet of Things (IoT) devices, and data services into a unified platform [6]. The main idea is that the disaster management architecture should be a module, scalable, and interoperable. This is achieved by an integrative AI platform built on a dataflow graph-based programming model, supported by middleware, and connected to microservices that represent IoT devices, AI services, and data sources. The programming model is particularly significant, as it reduces the complexity of developing disaster management applications. By representing applications as directed acyclic graphs where nodes signify functions and edges represent dataflows, the system allows modularity and parallelization. This approach makes programming less syntactically demanding for users while enabling sophisticated application logic. The middleware layer that positions the API gateway handles communication is the core of the system. It takes responsibility for the routing between the program interpreter and the external services, the unification of the data formats, and the prevention of security threats through access control. This middleware supports both synchronous and asynchronous communication, ensuring that real-time disaster data from IoT devices and AI models can be processed without bottlenecks [6].

Benssam A. (2017) presents middleware which targets decision support improvement in a disaster management environment specifically. This research stands on the idea that disaster

management depends on highly heterogeneous systems, ranging from sensor networks and IoT devices to existing databases and communication services. This middleware is a unifying layer that brings coherence among these different components, thus allowing data to be collected, shared, and processed without any issues. The research question contemplated by this study essentially revolves around whether middleware can improve decision-making in disaster scenarios by bridging the gap between the distributed data sources and the more sophisticated decision-making systems. It is stated that decision-making based on real-time information in critical situations depends not only on the data reliability but also on the architecture that supports timely processing and communication. In a broader perspective, middleware is a bridge connecting distributed data and real-time analysis and is also a tool for system interoperability and adaptive responses. From the results, it has been determined that middleware platforms are in fact effective in regulating the data paths and easing the system heterogeneity through the use of standardized APIs and protocols. The benefits gained in this way should be noted, as the system will share information more quickly and make decisions more swiftly, thus, reducing the time pressure that is normally present in disasters [7].

Work of Pillai. A, focus on building reliable early warning systems (EWS) by studying service-oriented IoT architecture for disaster preparedness and forecasting. The point that separates it from the traditional EWS methods is that the system is based on a layered IoT architecture integrated with a triple modular redundancy (TMR) fault-tolerant mechanism, which ensures the reliability and the absence of faulty sensor data for accurate cloud-based predictions. The paper notes the increasing importance of real-time monitoring and decision support systems in reducing risks and mitigating the socio-economic impacts of disasters, particularly in critical environments like underground mines. The main research question addressed by the paper is how IoT-based service architectures, combined with cloud-driven machine learning models, can provide accurate, fault-tolerant, and real-time disaster prediction capabilities. The authors seek to improve upon existing EWS frameworks by embedding redundancy in sensor networks and integrating machine learning algorithms for predictive analytics. The contributions of the study are demonstrated through the design of a layered IoT architecture composed of perception, middleware, service, and interface layers, which together streamline communication between sensing devices and cloud-based systems. To check the correctness of the framework the research paper analyzed the case of its utilization in underground mines, where reliable and continual environmental monitoring is of utmost importance for the provision of safety to workers and preventing hazardous conditions. The prime outcome suggests that the proposed system is able to generate highly accurate predictions even if its partial sensor fails. The combined work of lightweight IoT communication protocols, fault-tolerant sensor integration, and cloud-based predictive analytics presents a solid means of ensuring non-stop and accurate disaster prediction. This feature renders the system particularly appropriate for safety-critical sectors where downtime or erroneous data can have disastrous effects [8].

Correspondingly, the papers reveal a common aim in overcoming fragmentation in disaster management systems and enabling faster and more reliable decision-making. First paper highlights the role of AI-enhanced microservices for predictive and adaptive responses, while Benssam emphasizes the middleware's ability to harmonize diverse data sources and ensure robust communication. And the IoT architecture study highlights fault-tolerance and predictive accuracy in early warning systems. All these work acknowledge persistent problems such as real-time data constraints, scalability under high load, and the need for security in handling sensitive information. Their findings converge on the idea that middleware, whether AI-enhanced or designed as a decision-support hub, is indispensable in bridging the gap between distributed technologies and effective disaster management practices [6–8].

Artificial intelligence (AI) has emerged as a transformative tool in disaster management, offering advanced capabilities for monitoring, prediction, response, and recovery. It offers an innovative solution that integrates knowledge from diverse data types and sources, enables the simulation of realistic disaster scenarios, and detects emerging patterns with unprecedented speed. The following

literature review examines key approaches and applications of AI in disaster management, emphasizing their effectiveness, challenges, and potential for improving resilience [20].

The formula behind convolution operation in convolutional neural network for multi-class classification is:

$$Y(i, j) = \sum_{m=1}^{k_h} \sum_{n=1}^{k_w} \sum_{c=1}^C X(i+m, j+n, c) * K(m, n, c) + b$$

Where, input image $X \in R^{H*W*C}$, with height H, width W, channels C, and $Y(i, j)$ is a feature map value at location (i,j), with bias b.

Activation function is: (ReLU) is $A(i, j) = \max(0, Y(i, j))$, which introduces non-linearity.

And in order to reduce spatial size and capture dominant features, for a pooling window of pxp, the max pooling layer formula is:

$$P(i, j) = \max_{(m,n) \in p \times p} A(i+m, j+n)$$

The formula for Adam optimizer, which will be later used for building damage classification task is follows, gradient at step t is $g_t = \nabla_{\theta} L(\theta_t)$, where θ_t is model parameters at step t, and L is the loss function. From there exponential moving average of the gradient would as follows: $m_t = B_1 m_{t-1} + (1 - B_1) g_t$, where m_t is biased first momentum, while the second momentum estimate is: $v_t = B_2 v_{t-1} + (1 - B_2) g_t^2$.

And the parameter update rule is given by:

$$\theta_{t+1} = \theta_t - n * \frac{m_t}{\sqrt{v_t + \epsilon}}$$

Table 1 – Building damage classification CNN models

Reference	Model	Dataset/ city	Metric	Accuracy	Size of image	Number of epoch / batch size	Building features
[9]	U-Net, ResNet	Inria Dataset, America	Dice metric Test	97.95 %	(512×512)	200	shapes, structures, textures, and colours
[10]	UNet, ResNet	Xinxing County, China	Accuracy	84.9 %	(900 × 900) to (1024 × 1024)	50 / 2	detecting new, old buildings
[11]	U-Net, ResNet50	Boston	IoU accuracy	82.2 %	(1500×1500)	60	residential buildings in urban settings
[12]	U-Net, ResNet V2	Chicago dataset	F1 score	86 %	(5000 x 5000) cut to (512 x 512)	50 / 4	shadow, man-made non-building features, heterogeneity of roof
[13]	U-Net, ResNet34	Nashik, Maharastra	Accuracy	89 %	(512 x 512)	25 / 16	residential, industrial, holy places
[14]	U-Net, ResNet34	SpaceNet	Accuracy	84.7 %	(650 × 650)		building footprint extraction

Recent literature on CNN-based building damage classification consistently shows that encoder–decoder segmentation architectures, most commonly U-Net paired with ResNet backbones, are effective for extracting building footprints and diagnosing damage from high-resolution aerial and

satellite imagery [15, 16]. Across diverse datasets such as Inria, SpaceNet, and city-scale datasets such as Boston, Chicago, Xinxing County, and Nashik, these hybrid models benefit from the strong feature-extraction capacity of ResNet variants and the spatial-precision of U-Net decoders. Reported performance varies by dataset and task metric, ranging from very high Dice scores on clean building-extraction tasks ($\approx 97.9\%$ on Inria) to more modest IoU/F1 results ($\approx 82\text{--}86\%$) in urban settings where roof heterogeneity, shadows, and surrounding man-made features complicate segmentation. Differences in image resolution, pre-processing, and experimental setup strongly influence outcomes. Studies using large original tiles (e.g., 1500×1500 or 5000×5000) that are cropped to smaller patches (commonly 512×512) enable models to learn both contextual and fine-grained cues; however, varying patch sizes and batch/epoch regimes (examples: 200 epochs at 512×512 vs. 50 epochs with batch size 2 at $900\text{--}1024$ tiles) make direct comparison difficult. Performance also depends on the target labeling: tasks focused on building footprint extraction tend to report higher segmentation metrics, whereas damage classification, especially distinguishing new vs. old or partially damaged buildings, suffers from lower recall/F1 because the visual cues of damage are subtle and dataset imbalance is common [9–14].

Common challenges and recommended directions emerge from these works. First, data heterogeneity between different cities, roof materials, shadows, and non-building objects requires robust augmentation, domain adaptation, or fine-tuning strategies to transfer models between regions. Second, combining spectral, temporal (pre/post-disaster) and contextual metadata often improves damage detection beyond single-image segmentation.

Results and discussion

Building on the insights from existing research, the proposed system architecture is designed to address the challenges of fragmentation, real-time data processing, scalability, and security in disaster management. It is an integrated, AI-driven platform that analyzes pre-disaster and post-disaster imagery to assess structural damage in buildings following an earthquake. This system architecture could be described as modular, cloud-native, and it is organized around microservices to provide flexibility, fault tolerance, and ease of deployment. The pivotal component of the whole system is the Data Ingestion Layer, which aggregates data from lots of sources, such as satellite and drone images, IoT sensor data from in-situ devices, and historical building records. This layer facilitates both batch and real-time data collection through secure APIs and automated upload pipelines. After collection, the data enters the Middleware and Communication Layer, which operates as the platform's central nervous system. Here, the Middleware using Kafka or RabbitMQ integrates message queues and event brokers to transform the data, adapt formatting, and ensure proper, asynchronous module communication, which thus, upon dealing with integration gaps specified in the literature, facilitates the whole process [17]. The AI and ML layer is the system intelligence hub. It has a collection of microservices, which are containerized and can be updated independently. They perform tasks such as building detection, comparing pre- and post-event images, and deep learning models for damage classification. These models are fine-tuned for speed and reliability, supporting both real-time inference and batch processing after the incident.

The Decision Support Layer, for example, gets the outputs from this layer and it is here that the various outputs such as damage scores, heatmaps, and risk assessments that are generated and presented through a web dashboard. The dashboard was created specifically for emergency response teams with visual tools and prioritization recommendations to help them during rapid and informed decision-making. The entire architecture is underpinned by the Storage and Data Management Layer, which is solid and robust. It makes use of distributed object storage for high-resolution images and structured databases like PostgreSQL for metadata and model outputs. Data versioning, indexing, and lineage tracking are implemented to ensure auditability and traceability which are key requirements in disaster scenarios. To prevent breaches of sensitive information, the system incorporates a Security and Access Control Layer that implements role-based access, encrypts data both at rest and during

transit, while also maintaining comprehensive audit logs for accountability. The system, in the end, is built not only for scalability but also for fault tolerance. This is achieved by each microservice being horizontally scalable which enables the system to burst during high traffic moments under a natural disaster. Health checks, load balancing, and redundancy mechanisms will keep the system working even though not all parts are functioning correctly. Simply, all these architectural decisions are behind a strong, intelligent, and capable platform that can deliver timely and reliable post-earthquake building damage assessments.

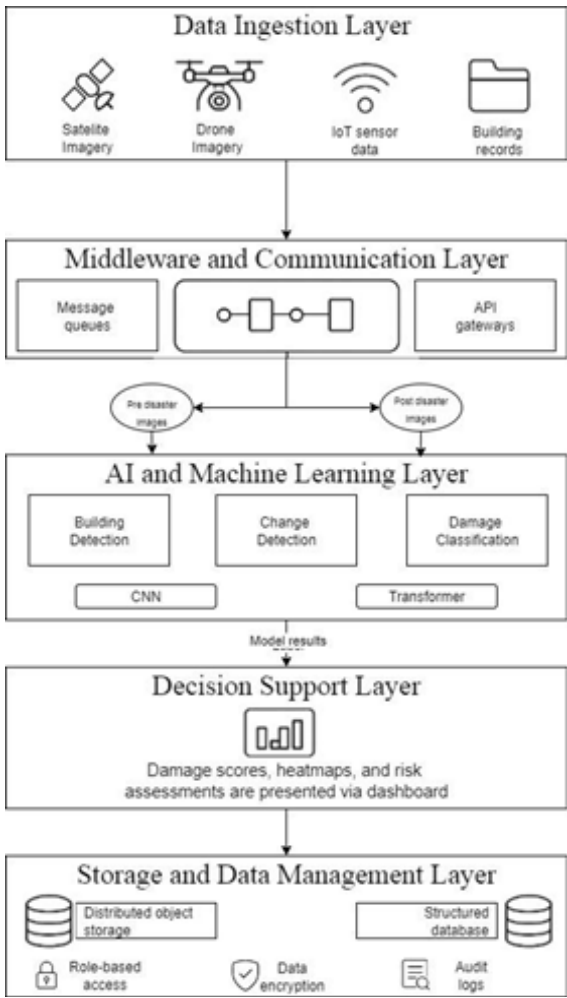


Figure 3 – Proposed system architecture

To evaluate the effectiveness of the proposed architecture, we implemented and tested the AI-based damage classification component, which plays a central role in the system’s analytical layer. This component would be integrated within the broader microservices framework described earlier, enabling seamless data flow from ingestion to analysis and visualization. Approach for building damage assessment using satellite imagery is to pose the problem as a combination of segmentation and classification tasks and to train deep-learning models on pre-disaster and post-disaster satellite images. The problem of damage classification can be thought of as a change detection problem, where the change is to be calculated over a period of time with a pair of pre- and post-disaster images [22]. The model was trained using the xBD dataset by Maxar and Microsoft AI for Earth is a large-scale satellite imagery dataset for building damage assessment, containing pre- and post-disaster imagery with building footprint annotations and damage classifications (No Damage, Minor, Major, Destroyed). The method of identifying damage level is that the model firstly represents a particular

building as a polygon and assigns a damage class to the polygon by comparing the difference between polygon representation of pre- and post-disaster satellite imagery.



Figure 4 – Satellite imagery and polygon representation

This CNN architecture is designed for 4-class building damage classification. It begins with three convolutional blocks (Conv2D → MaxPooling → BatchNorm), progressively reducing spatial dimensions while increasing feature depth. The output is flattened and passed through a dense layer with 128 units, followed by dropout for regularization. Finally, a dense output layer with 4 units predicts the damage class. The subsequent part enumerates the performance statistics of the building damage classification model, underscoring its precision, trustworthiness, and realistic utility in post-earthquake scenarios. These findings also provide information on the effectiveness of the system in actual circumstances thereby serving to prove the practicability of the use of AI in decision support within the system. The training and validation accuracy curves in the top plot show that the CNN model consistently achieves around 75.5% accuracy, with little improvement over 30 epochs, indicating early convergence. Meanwhile, the loss curves in the bottom plot demonstrate that training loss steadily decreases, while validation loss fluctuates, suggesting that the model may be overfitting slightly or struggling to generalize, especially due to class imbalance.

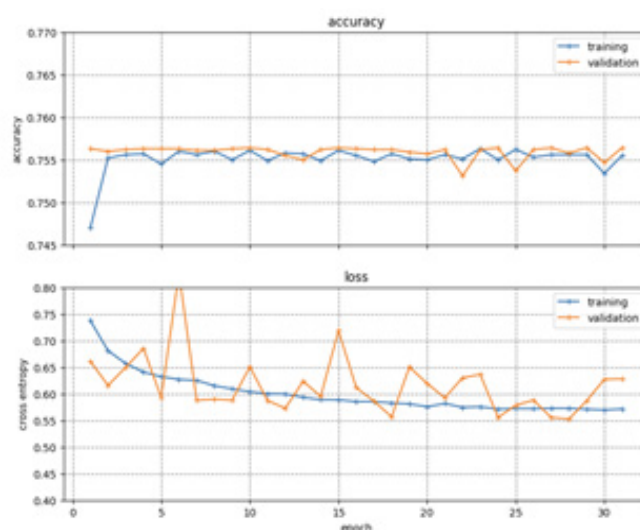


Figure 5 – Accuracy and loss function graph

Table 2 – Confusion matrix

	precision	recall	f1 score	support
class 0	0.76	1.00	0.86	6935
class 1	0.31	0.06	0.10	514
class 2	0.5	0.02	0.04	1675
accuracy			0.77	9124
macro avg	0.52	0.36	0.33	9124
weighted avg	0.68	0.77	0.67	9124

The matrix shows that the model is biased toward class 0, which dominates the dataset. Class 0 achieves strong performance with a precision of 0.76, perfect recall of 1.00, and an F1 score of 0.86, indicating that almost all instances of this class are predicted correctly. In contrast, the minority classes perform poorly. Class 1 and 2 have modest precision, but a low recall, which results in a weak F1 score, meaning most true instances of this class are missed. Overall accuracy is 0.77, but this is misleading because it is mainly driven by the number of correctly classified class 0 samples. The macro average precision 0.52, recall 0.36, F1 0.33 highlight poor balance across classes. The performance on classes 1 and 2 is very likely linked to their much smaller support compared to class 0. As a result, it learns to prioritize predicting the majority class, since doing so maximizes overall accuracy.

After balancing the dataset across classes and adopting the SiamUnet model which is more effective compared to traditional sequential CNN model, we see improvements in results, as shown in table 3. The architecture of the model consists of two identical U-Net encoders and decoders that process the pre- and post-event images separately. Each branch extracts multi-scale features using convolutional and pooling layers, followed by symmetric upsampling and decoding. After individual decoding, the Siamese module computes the difference between feature maps (bottleneck_2 and bottleneck_1) to capture the extent and location of changes. These difference maps are then progressively upsampled and concatenated with encoder differences from earlier layers, enabling the network to reconstruct fine-grained change patterns. The model was trained using the Adam optimizer with a learning rate of $1e-4$ and batch size of 8, training on 20 epoch. Two types of experiments were conducted:

Cross-event training: where model was trained randomly on all 19 different disaster events included in xBD dataset.

Event-specific training: where model was trained and evaluated only on images from Mexico City earthquake, following the 80/10/10 proportional split.

The results demonstrate significant improvements. It can be observed that the model performed better in classes 0 (no damage) and 2 (major damage), than in class 1 (minor damage), even though the dataset was balanced across classes. We can deduce that it might be because classes 0 and 2 have higher and more pronounced visual signs compared to minor-damaged-class. Interestingly, when the model was trained on earthquake specific dataset, classes 0 and 2 improved, whereas class 1 did not. It might be because earthquake-specific training is best at enhancing only distinctive cues, while minor damage class can be visually ambiguous, and poor satellite resolution can make it overlap with class 0.

Table 3 – F1 score for general and earthquake specific training

	Class 0	Class 1	Class 2
Metric	f1 score	f1 score	f1 score
General	0.88	0.25	0.41
Earthquake-specific	0.89	0.21	0.43

Conclusion

The proposed architecture presents a scalable, AI-driven, and cloud-native platform designed to deliver timely and reliable building damage assessments in the aftermath of earthquakes. By integrating diverse data sources, including satellite imagery, drone captures, IoT sensor readings, and historical records, the system ensures a realistic view of disaster impact. Its modular microservices design, supported by middleware for seamless communication and a robust storage layer, provides flexibility, resilience, and high performance. The inclusion of advanced AI and deep learning models enables accurate detection, classification, and visualization of structural damage. As the present work concentrates on the conceptual design of the disaster management architecture, future efforts will focus on completing the full system architecture. The next steps will involve developing the data ingestion layer to enable seamless integration of satellite/drone imagery, IoT sensor feeds, and building records. Additionally, work will extend to building the decision support layer, with an interactive dashboard that translates AI outputs into actionable insights for emergency responders. Efforts will also include establishing the storage and data management layer with distributed storage, indexing, and auditability, along with implementing the security and access control to safeguard sensitive information. Additionally, advancements in edge computing can be leveraged to bring AI inference closer to disaster sites, reducing latency and enabling offline operation in bandwidth-constrained environments.

This study demonstrated the potential of CNN-based models for building damage classification within disaster management systems. The results show that while the model performed strongly on class 0 (no damage), it struggled with classes 1 and 2. This imbalance highlights the challenges posed by unequal class distributions, where the majority class dominates training outcomes while minority classes remain underrepresented. After ensuring more balanced learning across all categories, classes 0 and 2 showed great performance, while class 1 struggled because of low visual cues and poor image quality that make it hard to capture subtle nuances of class 1. To address this, future work will focus on two main directions. First, I plan to incorporate Transformer-based architectures, which can capture richer contextual relationships and long-range dependencies, improving recognition of subtle classes. These improvements are expected to enhance the model's robustness and make it more reliable for real-world disaster management applications.

REFERENCES

- 1 Our World in Data. (n.d.). Decadal average: Death rates from natural disasters*. Our World in Data (access: August 24, 2025). URL: <https://ourworldindata.org/grapher/decadal-average-death-rates-from-natural-disasters>.
- 2 Our World in Data. (n.d.). Decadal average: Economic damages from disasters as a share of GDP. Our World in Data (access: August 24, 2025). URL: <https://ourworldindata.org/grapher/decadal-economic-disasters-type>.
- 3 Koshy, M., & Smith, D. Community resilience implications for institutional response under uncertainty: Cases of the floods in Wayanad, India and the earthquake in Port-au-Prince, Haiti. *Sustainable Development*, 32(2), 1412–1427 (2024).
- 4 Murayama, Y., Scholl, H. J., & Velev, D. Information technology in disaster risk reduction. *Information Systems Frontiers*, 23(5), 1077–1081 (2021).
- 5 Nasar, W., Da Silva Torres, R., Gundersen, O. E., & Karlsen, A. T. The use of decision support in search and rescue: A systematic literature review. *ISPRS International Journal of Geo-Information*, 12(5), 182 (2023).
- 6 Kamal Paul, S., & Bhaumik, P. Disaster management through integrative ai. In *Proceedings of the 23rd International Conference on Distributed Computing and Networking* (2022, January), pp. 290–293.
- 7 Benssam, A., Nouali-Taboudjemat, N., Nouali, O., & Kabou, A. A middleware platform for decision support in disaster management. In *2017 4th International Conference on Information and Communication Technologies for Disaster Management (ICT-DM)*. IEEE (2017, December), pp. 1–8.

- 8 Pillai, A.S., Chandraprasad, G.S., Khwaja, A.S., & Anpalagan, A. A service oriented IoT architecture for disaster preparedness and forecasting system. *Internet of Things*, 14, 100076 (2021).
- 9 Aghayari, S., Hadavand, A., Mohamadnezhad Niazi, S., & Omidalizarandi, M. Building detection from aerial imagery using inception resnet unet and unet architectures. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 10, 9–17 (2023).
- 10 Alsabhan, W., Alotaiby, T., Dudin, B. Detecting buildings and nonbuildings from satellite images using U-Net. *Comput. Intell. Neurosci* (2022). <https://doi.org/10.1155/2022/4831223>.
- 11 Alsabhan, W., & Alotaiby, T. Automatic building extraction on satellite images using Unet and ResNet50. *Computational Intelligence and Neuroscience*, 2022(1), 5008854 (2022).
- 12 Erdem, F., & Avdan, U. Comparison of different U-net models for building extraction from high-resolution aerial imagery. *International Journal of Environment and Geoinformatics*, 7(3), 221–227 (2020).
- 13 Vasavi, S., Somagani, H. S., & Sai, Y. Classification of buildings from VHR satellite images using ensemble of U-Net and ResNet. *The Egyptian Journal of Remote Sensing and Space Sciences*, 26(4), 937–953 (2023).
- 14 Van Etten, A., Lindenbaum, D., & Bacastow, T.M. SpaceNet: A Remote Sensing Dataset and Challenge Series. *ArXiv*, abs/1807.01232 (2018).
- 15 Mohammad, A., Gullapalli, O. S., Vasavi, S., Jayanthi, S., Updating of GIS maps with Change Detection of Buildings using Deep Learning techniques, 2022 International Conference on Futuristic Technologies (INCOFT), Belgaum, India, 2022, pp. 1–6. <https://doi.org/10.1109/INCOFT55651.2022.10094545>.
- 16 Kaku, K. Satellite remote sensing for disaster management support: A holistic and staged approach based on case studies in Sentinel Asia. *International Journal of Disaster Risk Reduction*, 33, 417–432 (2019).
- 17 Razzaque, M.A., Milojevic-Jevric, M., Palade, A., & Clarke, S. Middleware for internet of things: a survey. *IEEE Internet of things journal*, 3(1), 70–95 (2015).
- 18 Sun, W., Bocchini, P., & Davison, B. D. Applications of artificial intelligence for disaster management. *Natural Hazards*, 103(3), 2631–2689 (2020).
- 19 Khan, S.M., Shafi, I., Butt, W.H., Diez, I.D.L.T., Flores, M.A.L., Galán, J.C., & Ashraf, I. A systematic review of disaster management systems: approaches, challenges, and future directions. *Land*, 12(8), 1514 (2023).
- 20 Raj, A., Arora, L., Girija, S.S., Kapoor, S., Pradhan, D., & Shetgaonkar, A. AI and Generative AI Transforming Disaster Management: A Survey of Damage Assessment and Response Techniques. *arXiv preprint arXiv:2505.08202* (2025).
- 21 Pu, F., Li, Z., Wu, Y., Ma, C., & Zhao, R. Recent Advances in Disaster Emergency Response Planning: Integrating Optimization, Machine Learning, and Simulation. *arXiv preprint arXiv:2505.03979* (2025).

¹*Кайдұллаев М.А.,

докторант, ORCID ID: 0009-0000-7727-9095,

*e-mail: mi_kaidullayev@kbtu.kz

¹Акжалова А.Ж.,

PhD, профессор, ORCID ID: 0000-0002-1141-7595,

e-mail: a.akzhalova@kbtu.kz

¹Қазақстан-Британ техникалық университеті, Алматы қ., Қазақстан

АПАТТАРДЫ БАСҚАРУ ЖҮЙЕСІНІҢ ЖҮЙЕЛІК АРХИТЕКТУРАСЫ

Аңдатпа

Климаттың өзгеруінің жеделдеуі және қала халқының өсуі жағдайында табиғи апаттардың жиілігі мен ауырлығы айтарлықтай артып, инфрақұрылымға, экономикалық тұрақтылыққа және адам өміріне елеулі қауіп төндіруде. Жер сілкіністері, су тасқындары және дауылдар сияқты табиғи апаттар көбінесе құрылымдық зақымдануларға әкеледі, бұл төтенше жағдайларды жоюға және ресурстарды тиімді бөлуге жедел әрі дәл бағалаудың қажеттілігін арттырады. Осы сын-қатерлерге жауап ретінде бұл зерттеуде жасанды интеллект пен IoT технологияларын біріктіретін гибриді архитектурға негізделген ғимараттардың зақымдануын бағалау моделі ұсынылады. Зерттеу смарт-интеллект жүйелерінде автоматтандыруды, ашықтықты және тұрақтылықты арттыру үшін заттар интернеті мен жасанды интеллекттің кешенді

интеграциясын пайдаланады. Ұсынылған жүйе апатқа дейінгі және апаттан кейінгі аэрофототүсірілімдерді жинап, талдау арқылы ғимараттардың зақымдану деңгейін «зақымданбаған» күйден бастап «толық қираған» деңгейге дейінгі санаттарға жіктейді. Сонымен қатар, модель кеңейтілген апаттарды басқару жүйесіне біріктіріліп, геокеңістіктік интерфейсте зақымдану көрсеткіштерін көрнекі түрде ұсынады. Бұл шешім қабылдаушыларға зардап шеккен аймақтарды тиімді бағалауға, басым бағыттарды айқындауға және апатқа қарсы іс-қимылдарды оңтайландыруға мүмкіндік береді. Ұсынылып отырған жүйе мемлекеттік органдарға, үкіметтік емес ұйымдарға және алғашқы көмек көрсетушілерге апаттан кейінгі жағдайларда жедел, дәлелді және уақтылы шешім қабылдауды қолдауға бағытталған.

Тірек сөздер: апаттарды басқару, зақымдануды анықтау, табиғи апаттар, жүйелік архитектура, аралық бағдарлама, ЖИ, заттардың интернеті

^{1*}**Кайдуллаев М.А.,**

докторант, ORCID ID: 0009-0000-7727-9095,

*e-mail: mi_kaidullayev@kbtu.kz

¹**Акжалова А.Ж.,**

PhD, профессор, ORCID ID: 0000-0002-1141-7595,

e-mail: a.akzhalova@kbtu.kz

¹Казахстанско-Британский технический университет, г. Алматы, Казахстан

СИСТЕМНАЯ АРХИТЕКТУРА ДЛЯ СИСТЕМЫ УПРАВЛЕНИЯ СТИХИЙНЫМИ БЕДСТВИЯМИ

Аннотация

В эпоху ускоряющегося изменения климата и роста городского населения частота и интенсивность стихийных бедствий значительно возросли, представляя серьезную угрозу инфраструктуре, экономической стабильности и жизни людей. Такие стихийные бедствия, как землетрясения, наводнения и ураганы, часто приводят к обширным разрушениям зданий, требуя быстрой и точной оценки для экстренного реагирования и распределения ресурсов. В ответ на эти вызовы в данной работе представлена модель оценки ущерба зданиям на основе глубокого обучения, использующая гибридную архитектуру, сочетающую искусственный интеллект и интернет вещей. Данное исследование объединяет интернет вещей и искусственный интеллект для повышения автоматизации, прозрачности и устойчивости интеллектуальных систем. Система предназначена для сбора и анализа аэрофотоснимков до и после стихийных бедствий для классификации зданий по категориям ущерба – от неповрежденных до разрушенных. Кроме того, мы интегрируем модель в более широкую систему управления стихийными бедствиями, которая визуализирует оценки ущерба в геопространственном интерфейсе, позволяя лицам, принимающим решения, определять приоритетность пострадавших районов и оптимизировать меры реагирования на стихийные бедствия. Эта система призвана помочь государственным учреждениям, неправительственным организациям и службам быстрого реагирования принимать обоснованные и своевременные решения в ситуациях после стихийных бедствий.

Ключевые слова: борьба со стихийными бедствиями, обнаружение ущерба, стихийные бедствия, системная архитектура, промежуточное программное обеспечение, ИИ, интернет вещей.

Article submission date: 25.08.2025

МАТЕМАТИКАЛЫҚ ҒЫЛЫМДАР
MATHEMATICAL SCIENCES
МАТЕМАТИЧЕСКИЕ НАУКИ

UDC 519.642
IRSTI 27.33.15

<https://doi.org/10.55452/1998-6688-2025-22-4-266-278>

^{1,2}**Tamabay D.,**

PhD student, ORCID ID: 0000-0001-8315-5849

e-mail: dtamabay@gmail.com

^{1,2}**Temirbekov A.,**

PhD, Associate Professor, ORCID ID: 0000-0002-4157-2799

e-mail: almas_tem@mail.ru

^{1,2}**Tleulesova A.,**

PhD, ORCID ID: 0000-0001-9280-1048

e-mail: aigerim1985_06@mail.ru

^{1,2*}**Mukhanova T.**

Bachelor's student, ORCID ID: 0009-0000-8522-3390

*e-mail: tomirismukhanova04@gmail.com

¹Al-Farabi Kazakh National University, Almaty, Kazakhstan

²National Engineering Academy of the Republic of Kazakhstan, Almaty, Kazakhstan

**ALPERT WAVELET-BASED GALERKIN METHOD
FOR FIRST-KIND FREDHOLM INTEGRAL EQUATIONS****Abstract**

This paper presents a numerical approach for solving Fredholm integral equations of the first kind using the Bubnov–Galerkin method with Alpert wavelet bases. These equations are well-known for being ill-posed, meaning that small changes in input data can lead to large deviations in the solution. Therefore, robust and accurate numerical methods are essential. The proposed method utilizes orthonormal and compactly supported Alpert wavelets, which offer excellent localization properties and yield well-conditioned, sparse system matrices when projecting the integral operator. This enhances numerical stability and reduces computational complexity. A series of computational experiments was carried out using various refinement levels and polynomial degrees. The accuracy of the method was evaluated by comparing approximate solutions to the exact analytical solution. The results demonstrate exceptionally small absolute errors, often approaching machine precision. Additionally, a comparative analysis with power polynomial bases confirms the superiority of the Alpert wavelet approach in terms of convergence and approximation quality. Overall, the method proves to be efficient, stable, and suitable for further extension to more complex integral equations, including multidimensional and noisy-data problems. This confirms the potential of Alpert wavelet-based Galerkin schemes as a reliable tool for the numerical treatment of inverse and ill-posed problems in applied sciences.

Keywords: first kind Fredholm Integral equation, Bubnov–Galerkin method, Alpert wavelet, ill-posedness, orthonormal basis.

Introduction

Mathematics and computational methods play a key role in solving a wide range of applied problems arising in engineering, physics, mechanics, atmospheric sciences, geophysics, geochemistry, geological exploration, medical imaging, remote sensing, military, and other high-tech disciplines. In many of these fields, a significant number of applied and theoretical problems can be reduced to the solution of Fredholm integral equations of the first kind [1–3].

It is worth noting that initial and boundary value problems for differential equations can be transformed into integral form, reducing the original problem to solving a single integral equation [4, 5]. In this case, computations are performed based on boundary data, which is especially important for problems with high dimensionality or complex domain geometries. However, a crucial theoretical and numerical aspect in dealing with Fredholm integral equations of the first kind is their ill-posedness: the solution strongly depends on the smoothness of the kernel and the right-hand side, and even small perturbations in the data may lead to large deviations in the solution. This makes it impossible to obtain an exact analytical solution in most cases and necessitates the use of stable and highly accurate numerical methods [6, 7].

The development of numerical methods for solving Fredholm integral equations of the first kind is particularly relevant due to the need to ensure convergence, stability, and approximation accuracy of the computed solution [8–10]. At the same time, an important aspect of theoretical analysis is the study of the conditions of existence and uniqueness of the solution, since for ill-posed problems, which include Fredholm integral equations of the first kind, such properties are not guaranteed without introducing additional restrictions or using special regularization methods [11–14]. In recent decades, various approaches have been proposed: direct numerical integration, regularization methods [15, 16] including the modified Tikhonov method [17], projection methods [18–20], smoothing techniques, GMRES-type algorithms, multilevel iterative schemes [21], as well as modern wavelet and wavelet-Galerkin methods [22–24], among others [25–27].

Wavelet-based methods have attracted increasing interest in recent years as a versatile tool for constructing adaptive bases with strong localization and high approximation order [28–29]. Since the first application of wavelets to the numerical solution of integral equations in 1991, their use has expanded considerably. Different types of wavelets—including Haar, Coifman, Legendre, and others – have demonstrated good performance in solving problems with various characteristics, confirming the high accuracy and stability of the approach [30–33]. Each wavelet type possesses specific properties—such as compact support, orthogonality, vanishing moments, symmetry, and smoothness – which determine its effectiveness in projection-based numerical methods.

Among the most promising directions in numerical modeling are wavelets constructed from classical orthogonal polynomials [30]. These constructions are actively used in spectral analysis, numerical solution of differential equations, signal and image processing, and high-accuracy approximations with controlled localization. In the context of projection methods for solving Fredholm integral equations of the first kind, the approximation properties of the basis functions are of particular importance. Effective choices include wavelet systems based on classical orthogonal polynomials such as Legendre, Chebyshev, Hermite, and Laguerre. These functions form orthonormal bases that satisfy requirements of localization, smoothness, and stability.

Legendre wavelets, orthogonal on the interval $[-1, 1]$, have compact support and are well suited for implementing spectral-Galerkin schemes. Hermite wavelets, defined on $\mathbb{R}$, exhibit strong space-frequency localization and are applied in modeling solutions on the infinite domain. Chebyshev wavelets possess minimax properties and offer high approximation accuracy in the uniform norm, while Laguerre wavelets, defined on the half-line $[0, \infty)$, effectively approximate exponentially decaying functions [31–33].

The choice of a specific wavelet system as a basis in projection methods should be guided by both the geometry of the domain and the analytical properties of the solution and kernel of the integral equation. The basis properties of wavelets – such as orthonormality, locality, smoothness,

and exponential decay – directly influence approximation quality, convergence, and the conditioning of the resulting algebraic systems.

Among all wavelets, Alpert wavelets occupy a special place – these are multilevel orthonormal bases constructed from Legendre polynomials [34]. They combine compact support, localized structure, and high smoothness, making them particularly effective within variational numerical methods. Their use in the Galerkin method for solving Fredholm integral equations of the first kind ensures high approximation accuracy and numerical stability, especially for smooth or locally inhomogeneous kernels. Due to the orthonormality of the basis functions, the computation of matrix elements is significantly simplified, reducing the conditioning of the linear system. Alpert wavelets thus represent a powerful tool in the spectrum of modern projection methods for numerical analysis of integral equations.

In this work, we present a numerical approach for solving Fredholm integral equations of the first kind based on the Bubnov–Galerkin method using Alpert wavelets as basis functions. Computational experiments are conducted with varying refinement levels and numbers of polynomial components, with visualizations of the obtained solutions. A comparative study is also carried out using standard basis functions in the form of power polynomials. The results demonstrate a high degree of approximation convergence of the numerical solution to the exact one, confirming the effectiveness of the chosen basis within the proposed numerical method.

Materials and methods

Fredholm integral equation of the first kind

Let $\varphi(x)$ be the unknown function, $f(x)$ the given function, and $K(x, t)$ the kernel. Let $\lambda \in \mathbb{R}$ be a parameter. The Fredholm integral equation of the first kind arises from the general equation when $\lambda = 0$:

$$\int_a^b K(x, t)\varphi(t) dt = f(x), \quad x \in [a, b]. \quad (1)$$

Here $K(x, t), f(x) \in L^2([a, b])$. The main feature of this equation is its ill-posedness. Specifically: solution may not be unique and can be extremely sensitive to small perturbations in the function $f(x)$.

Therefore, to solve such equations, it is necessary to apply regularizing, projection, or approximate methods.

Bubnov–Galerkin method

The Bubnov–Galerkin method is a projection method in which the solution $\varphi(t)$ is approximated by a linear combination of specially chosen basis functions:

$$\varphi(t) \approx \varphi_N(t) = \sum_{j=1}^N c_j \phi_j(t), \quad (2)$$

where $\{\phi_j(t)\}_{j=1}^N$ is a system of basis functions, and c_j are the coefficients to be determined.

Substituting this approximation into equation (1):

$$\int_a^b K(x, t) \left(\sum_{j=1}^N c_j \phi_j(t) \right) dt = f(x). \quad (3)$$

Interchanging the integral and the sum in (3) yields:

$$\sum_{j=1}^N c_j \int_a^b K(x, t) \phi_j(t) dt = f(x).$$

In the Bubnov–Galerkin method, the residual must be orthogonal to the basis functions. That is, for each $i = 1, 2, \dots, N$, we require:

$$\sum_{j=1}^N c_j \int_a^b \int_a^b K(x, t) \phi_j(t) \phi_i(x) dt dx = \int_a^b f(x) \phi_i(x) dx. \quad (4)$$

We introduce notations:

$$A_{ij} = \int_a^b \int_a^b K(x, t) \phi_j(t) \phi_i(x) dt dx, \quad b_i = \int_a^b f(x) \phi_i(x) dx. \quad (5)$$

This leads to the linear system:

$$\sum_{j=1}^N A_{ij} c_j = b_i, \quad i = 1, 2, \dots, N. \quad (6)$$

This is a linear algebraic system of the form $Ac = b$.

The effectiveness of the Bubnov–Galerkin method depends heavily on the choice of basis functions. Traditional bases include polynomials or trigonometric functions (e.g., Legendre or Chebyshev polynomials), which usually lead to dense matrices.

Alpert wavelets, however, are based on generalized Haar-like functions and form a multiscale family. When such bases are used, the matrix often becomes sparse, which greatly simplifies computations.

Alpert Wavelets

Alpert wavelets are defined using scaled, shifted, and rescaled versions of the first k Legendre polynomials on the interval $[0, 1]$ via scaling functions $\{S^j(t)\}_{j=0}^{k-1}$. These functions have the following properties:

1. They are defined on the interval $[0, 1]$;
2. They form an orthonormal basis

$$\int_0^1 S^j(t) S^i(t) dt = \delta_{ij}, \quad (7)$$

3. They span the space of polynomials of degree less than k .

Alpert wavelets are given by:

$$h_{j,m}^n(x) = 2^{m/2} h^j(2^m x - n), \quad n \in \mathbb{N}_0, \quad (8)$$

where m is the level of scaling and n is the shift parameter. These functions form a basis consisting of k wavelet functions localized on subintervals of length 2^{-m} in $[0, 1]$.

For a fixed positive integer k and $m = 0, 1, 2, \dots$, the piecewise polynomial space S_m^k is defined as:

$$S_m^k = \{f: f|_{(2^{-m}n, 2^{-m}(n+1))} \in P_{k-1}, \quad n = 0, 1, \dots, 2^m - 1, \quad f = 0 \text{ elsewhere}\},$$

where P_{k-1} is the space of polynomials of degree less than k . The dimension of this space is $\dim(S_m^k) = 2^m k$, and the following nested structure holds:

$$S_0^k \subset S_1^k \subset \dots \subset S_m^k \subset \dots \quad (9)$$

For each m , define the orthogonal complement:

$$S_m^k \oplus R_m^k = S_{m+1}^k, \quad R_m^k \perp S_m^k, \quad \dim(R_m^k) = 2^m k. \quad (10)$$

By induction, we obtain:

$$S_m^k = S_0^k \oplus R_0^k \oplus R_1^k \oplus \dots \oplus R_{m-1}^k. \quad (11)$$

Let $h_1, \dots, h_k: \mathbb{R} \rightarrow \mathbb{R}$ be an orthonormal basis of R_0^k . Since $R_0^k \perp S_0^k$, the first k moments of h_j vanish:

$$\int_0^1 h_j(x) x^i dx = 0, \quad i = 0, 1, \dots, k-1. \quad (12)$$

To define $h_1, \dots, h_k$, we first construct functions $f_1, \dots, f_k: \mathbb{R} \rightarrow \mathbb{R}$ on $[0, 1]$ with the following properties:

- ♦ Each f_i is a polynomial of degree at most $k - 1$ on $[0, 1]$;
- ♦ It extends to $(-1, 0)$ as an even or odd function depending on the parity of $i + k - 1$;
- ♦ They satisfy orthonormality and normalization:

$$\int_{-1}^1 f_i(x) f_j(x) dx = \delta_{ij}, \quad i, j = 1, \dots, k, \quad (13)$$

- ♦ Each f_j satisfies moment vanishing conditions:

$$\int_{-1}^1 f_j(x) x^i dx = 0, \quad i = 0, 1, \dots, j + k - 2. \quad (14)$$

- ♦ These functions f_j meet all four properties. Then,

$$h_i(x) = \sqrt{2} f_i(2x - 1), \quad i = 1, \dots, k. \quad (15)$$

Thus,

$$R_0^k = \text{span}\{h_i: i = 1, \dots, k\}. \quad (16)$$

In general,

$$R_m^k = \text{span}\{h_{j,m}^n: h_{j,m}^n(x) = 2^{m/2} h_j(2^m x - n), j = 1, \dots, k, n = 0, \dots, 2^m - 1\}.$$

Alpert wavelets when $k = 2$:

$$h_1^0(t) = \begin{cases} -\sqrt{3}(4t - 1), & \text{if } 0 \leq t < \frac{1}{2}, \\ 0, & \text{if } \frac{1}{2} < t \leq 1, \end{cases} \quad (17)$$

$$h_2^0(t) = \begin{cases} 0, & \text{if } 0 \leq t < \frac{1}{2}, \\ \sqrt{3}(4t - 3), & \text{if } \frac{1}{2} < t \leq 1, \end{cases} \quad (18)$$

$$h_1^1(t) = \begin{cases} 6t - 1, & \text{if } 0 \leq t \leq \frac{1}{2}, \\ 0, & \text{if } \frac{1}{2} < t \leq 1, \end{cases} \quad (19)$$

$$h_2^1(t) = \begin{cases} 0, & \text{if } 0 \leq t \leq \frac{1}{2}, \\ 6t - 5, & \text{if } \frac{1}{2} < t \leq 1. \end{cases} \quad (20)$$

when $k = 3$:

$$h_1^0(t) = \begin{cases} -\frac{7}{3} + 24t - 40t^2, & \text{if } 0 \leq t \leq \frac{1}{2}, \\ 0, & \text{if } \frac{1}{2} < t \leq 1, \end{cases} \quad (21)$$

$$h_2^0(t) = \begin{cases} 0, & \text{if } 0 \leq t \leq \frac{1}{2}, \\ \frac{55}{3} - 56t + 40t^2, & \text{if } \frac{1}{2} < t \leq 1, \end{cases} \quad (22)$$

$$h_1^1(t) = \begin{cases} \sqrt{3}(1 - 14t + 30t^2), & \text{if } 0 \leq t < \frac{1}{2} \\ 0, & \text{if } \frac{1}{2} < t \leq 1 \end{cases} \quad (23)$$

$$h_2^1(t) = \begin{cases} 0, & \text{if } 0 \leq t < -\frac{1}{2}, \\ \sqrt{3}(17 - 46t + 30t^2), & \text{if } \frac{1}{2} < t \leq 1, \end{cases} \quad (24)$$

$$h_1^2(t) = \begin{cases} -\sqrt{5}\left(\frac{1}{3} + 6t - 16t^2\right), & \text{if } 0 \leq t \leq \frac{1}{2}, \\ 0, & \text{if } \frac{1}{2} < t \leq 1, \end{cases} \quad (25)$$

$$h_2^2(t) = \begin{cases} 0, & \text{if } 0 \leq t \leq \frac{1}{2}, \\ \sqrt{5}\left(\frac{31}{3} - 26t + 16t^2\right), & \text{if } \frac{1}{2} < t \leq 1. \end{cases} \quad (26)$$

Results and discussion

Test problem 1

Consider the following linear integral equation as a representative applied example of Fredholm integral equations of the first kind:

$$\int_0^1 \sqrt{x^2 + t^2} y(t) dt = \frac{(1+x^2)^{3/2} - x^3}{3}, \quad x \in [0,1].$$

It is known that the exact analytical solution to this equation is given by:

$$y(t) = t.$$

The following Table 1 presents the absolute error values between the exact analytical solution and the approximate solution obtained using the Alpert wavelet-based method at selected nodal points in the interval $t \in [0,1]$:

Table 1 – Absolute errors between the exact and approximate solutions at selected nodes

t	Exact solution	Approximate solution k=2, M=2	Absolute error k=2, M=2	Approximate solution k=2, M=3	Absolute error k=2, M=3
0.0	0.0	6.11715×10^{-10}	6.11715×10^{-10}	2.94681×10^{-5}	2.94681×10^{-5}
0.1	1.0×10^{-1}	1.00000×10^{-1}	2.88965×10^{-10}	1.00003×10^{-1}	3.08759×10^{-6}
0.2	2.0×10^{-1}	2.00000×10^{-1}	3.37850×10^{-11}	1.99989×10^{-1}	1.13691×10^{-5}
0.3	3.0×10^{-1}	2.99999×10^{-1}	3.56535×10^{-10}	2.99986×10^{-1}	1.39021×10^{-5}
0.4	4.0×10^{-1}	3.99999×10^{-1}	6.79285×10^{-10}	3.99995×10^{-1}	4.51122×10^{-6}
0.5	5.0×10^{-1}	4.99999×10^{-1}	8.12233×10^{-10}	4.99998×10^{-1}	2.10066×10^{-6}
0.6	6.0×10^{-1}	6.00000×10^{-1}	6.00417×10^{-10}	6.00003×10^{-1}	2.78820×10^{-6}
0.7	7.0×10^{-1}	7.00000×10^{-1}	3.88602×10^{-10}	7.00005×10^{-1}	4.66764×10^{-6}
0.8	8.0×10^{-1}	8.00000×10^{-1}	1.76788×10^{-10}	8.00004×10^{-1}	3.53763×10^{-6}
0.9	9.0×10^{-1}	9.00000×10^{-1}	3.50280×10^{-11}	9.00000×10^{-1}	6.01806×10^{-7}
1.0	1.0×10^0	1.00000×10^0	2.46844×10^{-10}	9.99992×10^{-1}	7.75068×10^{-6}

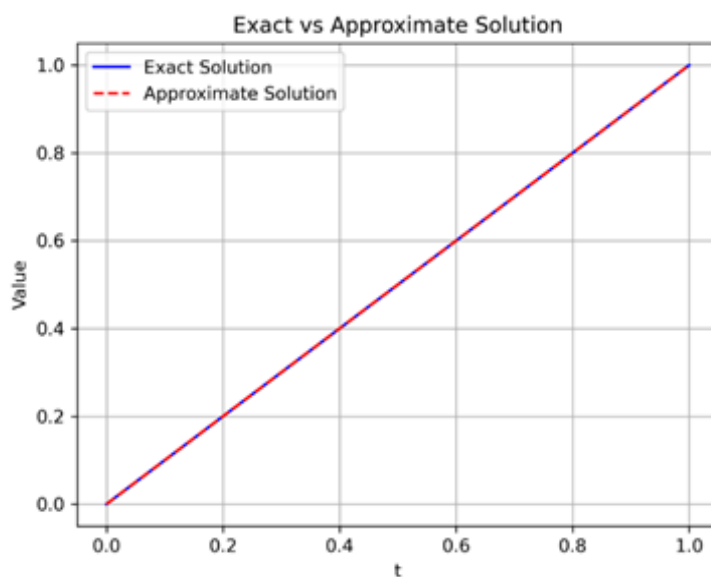


Figure 1 – Graph of exact solution and approximate solution of problem 1 in the case when $k = 2, M = 2$

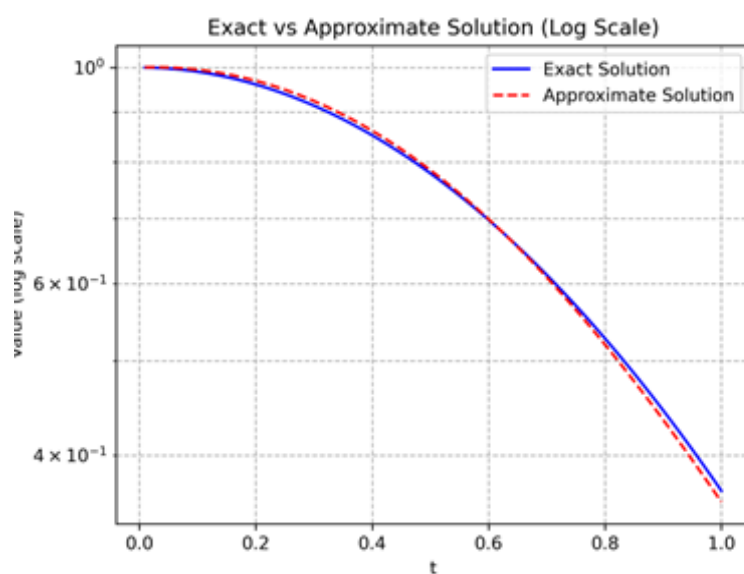


Figure 2 – Graph of exact solution and approximate solution of problem 1 in log scale in the case when $k = 2, M = 3$

Test problem 2

Consider the following linear integral equation as a representative applied example of Fredholm integral equations of the first kind:

$$\int_0^1 e^{xt} y(t) dt = \frac{e^{(x+1)} - 1}{1+x}, x \in [0,1].$$

It is known that the exact analytical solution to this equation is given by:

$$y(t) = e^t.$$

The following table presents the absolute error values between the exact analytical solution $y(t) = t$ and the approximate solution obtained using the Alpert wavelet-based method at selected nodal points in the interval $t \in [0,1]$:

Table 2 – Absolute errors between the exact and approximate solutions at selected nodes

t	Exact solution	Approximate solution $k=2, M=2$	Absolute error $k=2, M=2$	Approximate solution $k=2, M=3$	Absolute error $k=2, M=3$
0	1.00000×10^0	9.74063×10^{-1}	2.59365×10^{-2}	1.00032×10^0	3.19253×10^{-4}
1.0	1.10517×10^0	1.10487×10^0	2.99323×10^{-4}	1.10496×10^0	2.07005×10^{-4}
2.0	1.22140×10^0	1.23568×10^0	1.42770×10^{-2}	1.22163×10^0	2.29455×10^{-4}
3.0	1.34986×10^0	1.36649×10^0	1.66290×10^{-2}	1.35032×10^0	4.65347×10^{-4}
4.0	1.49182×10^0	1.49730×10^0	5.47124×10^{-3}	1.49104×10^0	7.84960×10^{-4}
5.0	1.64872×10^0	1.58505×10^0	6.36731×10^{-2}	1.65490×10^0	6.17632×10^{-3}
6.0	1.82212×10^0	1.80522×10^0	1.68980×10^{-2}	1.82271×10^0	5.88389×10^{-4}
7.0	2.01375×10^0	2.02539×10^0	1.16408×10^{-2}	2.01285×10^0	9.01444×10^{-4}
8.0	2.22554×10^0	2.24557×10^0	2.00252×10^{-2}	2.22533×10^0	2.11119×10^{-4}
9.0	2.45960×10^0	2.46574×10^0	6.13561×10^{-3}	2.46014×10^0	5.39715×10^{-4}
1.0	2.71828×10^0	2.68591×10^0	3.23705×10^{-2}	2.71729×10^0	9.91514×10^{-4}

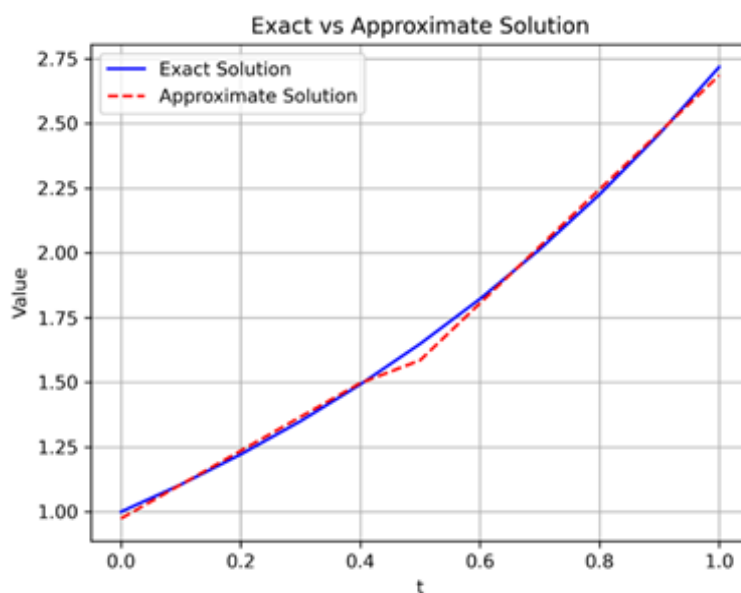


Figure 3 – Graph of exact solution and approximate solution of problem 2 in the case when $k = 2, M = 2$

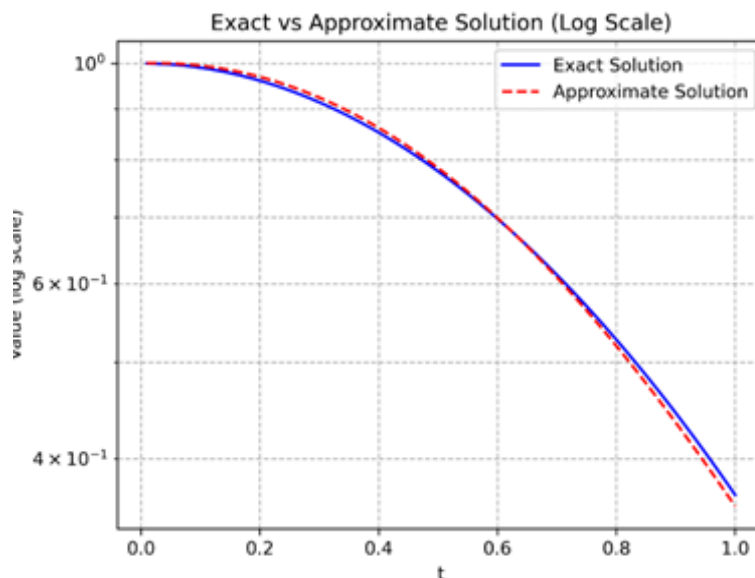


Figure 4 – Graph of exact solution and approximate solution of problem 2 in log scale in the case when $k = 2, M = 3$

When applying the Galerkin method to solve Fredholm integral equations of the first kind, the use of Alpert wavelets as basis functions presents both significant advantages and certain limitations, particularly in problems involving non-smooth kernels or solutions.

One of the main advantages of Alpert wavelets lies in their compact support, orthonormality, and piecewise-polynomial structure, which allow for efficient numerical implementation and sparse matrix representation. Their localized nature ensures that local features of the solution – such as peaks or sharp transitions – can be captured more accurately than with global basis functions like polynomials or trigonometric series. Additionally, the multi-resolution properties of wavelets provide a natural framework for adaptivity, enabling high accuracy with fewer basis functions when the solution is sufficiently smooth. These characteristics make Alpert wavelets particularly well-suited for ill-posed problems where stability and local refinement are essential [34].

However, the effectiveness of Alpert wavelets strongly depends on the smoothness of both the kernel $K(x, t)$ and the exact solution $u(t)$. If either contains discontinuities, singularities, or non-differentiable features, the approximation using a fixed number of wavelets becomes less accurate. In such cases, the method may require an excessive number of basis functions, which leads to the formation of an ill-conditioned linear system. This, in turn, increases the sensitivity of the solution to round-off and discretization errors, especially in the presence of small perturbations in the right-hand side function $f(x)$. Consequently, even minor numerical noise can result in significant deviations in the computed solution.

Furthermore, numerical integration used to compute the matrix coefficients in the Galerkin system may lose accuracy when the kernel is not smooth. The interaction between the wavelet basis and a non-smooth kernel can generate integrals with sharp variations that are difficult to evaluate precisely using standard quadrature techniques. Without adaptive integration or mesh refinement, these errors propagate into the system matrix and degrade the overall accuracy of the solution.

To mitigate these challenges, it is often necessary to introduce regularization techniques such as Tikhonov regularization, or to enhance the wavelet basis by including adaptive refinement or non-uniform scaling. These approaches help stabilize the numerical solution and improve approximation in the presence of singularities or discontinuities.

Conclusion

The numerical results obtained provide strong evidence for the accuracy and robustness of the Bubnov–Galerkin method when implemented with Alpert wavelet bases for solving first-kind Fredholm integral equations. A detailed analysis of the absolute errors between the exact analytical solution and the computed approximate solutions at uniformly distributed nodal points in the interval $t \in [0,1]$ demonstrates that the discrepancies are minimal – often approaching machine precision – thereby confirming the method’s high numerical fidelity.

A key aspect contributing to the method’s efficiency is the use of orthonormal and compactly supported Alpert wavelets. These basis functions ensure localized approximation and allow for the construction of sparse and well-conditioned system matrices upon projection of the integral operator. This, in turn, leads to enhanced numerical stability and reduced computational complexity, particularly beneficial when dealing with ill-posed problems.

Moreover, the structure of the resulting matrices facilitates scalable and efficient implementation in higher-dimensional problems. The localized wavelet approach not only improves accuracy but also enables selective resolution refinement, which is especially valuable in practical applications involving singularities, discontinuities, or noisy data.

The versatility of the proposed method is noteworthy. It exhibits strong potential for extension to more complex and realistic scenarios, including integral equations with discontinuous kernels, noisy right-hand sides, and multidimensional domains commonly encountered in physical, engineering, and geophysical modeling. The method remains stable and accurate under increased problem complexity and dimensionality, making it a promising computational tool for modern inverse and direct problems.

Information on funding

The research is funded by Science Committee of the Ministry of Science and Higher Education of the Republic of Kazakhstan (Grant No. BR27100483 "Development of predictive exploration technologies for identifying ore-prospective areas based on data analysis from the unified subsurface user platform "Minerals.gov.kz" using artificial intelligence and remote sensing methods").

REFERENCES

- 1 Bakushinskii, A.B. A numerical method for solving Fredholm integral equations of the 1st kind. USSR Computational Mathematics and Mathematical Physics, 5 (4), 226–233 (1965). [https://doi.org/10.1016/0041-5553\(65\)90135-7](https://doi.org/10.1016/0041-5553(65)90135-7).
- 2 Denisov, A.M. Approximation of quasi-solutions of Fredholm’s equation of the first kind with a kernel of special form. USSR Computational Mathematics and Mathematical Physics, 11 (5), 269–276 (1971). [https://doi.org/10.1016/0041-5553\(71\)90051-6](https://doi.org/10.1016/0041-5553(71)90051-6).
- 3 Denisov, A.M. Approximation of the quasi-solutions of a Fredholm integral equation of the first kind of a special form. USSR Computational Mathematics and Mathematical Physics, 12 (6), 244–248 (1972). [https://doi.org/10.1016/0041-5553\(72\)90152-8](https://doi.org/10.1016/0041-5553(72)90152-8).
- 4 Denisov, A.M. On the order of approximation when solving a Fredholm equation of the first kind with a kernel of special type. USSR Computational Mathematics and Mathematical Physics, 13 (1), 255–260 (1974). [https://doi.org/10.1016/0041-5553\(74\)90019-6](https://doi.org/10.1016/0041-5553(74)90019-6).
- 5 Groetsch, C.W. Convergence analysis of a regularized degenerate kernel method for Fredholm integral equations of the first kind. Integral Equations and Operator Theory, 13 (1), 67–75 (1990). <https://doi.org/10.1007/BF01195293>.
- 6 Yuan, D., Lu, S., Li, D., Zhang, X. Graph refining via iterative regularization framework. SN Applied Sciences, 1 (387), 1–10 (2019). <https://doi.org/10.1007/s42452-019-0412-9>.
- 7 Maleknejad, K., Mollapourasl, R., Nouri, K. Convergence of numerical solution of the Fredholm integral equation of the first kind with degenerate kernel. Applied Mathematics and Computation, 181 (2), 1000–1007 (2006). <https://doi.org/10.1016/j.amc.2006.01.074>.

- 8 Wei, P. Integral equations and their numerical methods. Beijing: Metallurgical Industry Press, 70–89 (2007).
- 9 Hackbusch, W. Integral equations: theory and numerical treatment. Birkhäuser, 120 (2012). <https://doi.org/10.1007/978-3-0348-9215-5>.
- 10 Yun, T. Representation theorem and one-iteration theorem for Fredholm integral equation of the first kind $ax \doteq y$. Applied Mathematics and Mechanics, 07, 593–597 (1989). <https://doi.org/10.1007/BF02115791>
- 11 Li, J. et al. A rapid solution of a kind of 1D Fredholm oscillatory integral equation. Journal of Computational and Applied Mathematics, 236 (10), 2696–2705 (2012). <https://doi.org/10.1016/j.cam.2012.01.007>.
- 12 Han, H., Li, Y., Yin, D., Chen, Z. The necessary and sufficient condition for the existence and uniqueness of a system of Fredholm integral equations of the first kind. Science China Mathematics, 45 (8), 1231–1248 (2015).
- 13 He, R., Zhang, L. Numerical methods for Fredholm integral equations of the first kind. In: International Conference on Physics, Computing and Mathematical Modeling, 2018, pp. 133–138.
- 14 Lin, F., Yang, S. A two-stage method for piecewise-constant solution for Fredholm integral equations of the first kind. Mathematics, 5 (2), 1–16 (2017). <https://doi.org/10.3390/math5020028>.
- 15 Lee, J.W., Prenter, P.M. An analysis of the numerical solution of Fredholm integral equations of the first kind. Numerische Mathematik, 30 (1), 1–23 (1978). <https://doi.org/10.1007/BF01403903>.
- 16 Khromov, A.A., Khromova, G.V. Finding approximations of continuous solutions to first-kind equations. Computational Mathematics and Mathematical Physics, 49 (2), 217–223 (2009). <https://doi.org/10.1134/S096554250902002X>.
- 17 Tikhonov, A.N., Goncharsky, A.V., Stepanov, V.V., Yagola, A.G. Numerical methods for the solution of ill-posed problems. Dordrecht: Kluwer, 1995. https://doi.org/10.1007/978-94-015-8480-7_3.
- 18 Nair, M.T., Pereverzev, S.V. Regularized collocation method for Fredholm integral equations of the first kind. Journal of Complexity, 23 (4), 454–467 (2007). <https://doi.org/10.1016/j.jco.2006.09.002>.
- 19 Maleknejad, K., Saeedipour, E., Dehbozorgi, R. Legendre wavelets direct method for the numerical solution of Fredholm integral equation of the first kind. In: Proceedings of the World Congress on Engineering, 2016, vol. 1.
- 20 Bechouat, T. A collocation method for Fredholm integral equations of the first kind via iterative regularization scheme. Mathematical Modelling and Analysis, 28 (2), 237–254 (2023).
- 21 Luo, X. et al. Multilevel Jacobi and Gauss–Seidel type iteration methods for solving ill-posed integral equations. Journal of Inverse and Ill-posed Problems, 23 (5), 477–490 (2015). <https://doi.org/10.1515/jiip-2013-0045>.
- 22 Shang, X., Han, D. Numerical solution of Fredholm integral equations of the first kind by using linear Legendre multi-wavelets. Applied Mathematics and Computation, 191 (2), 440–444 (2007). <https://doi.org/10.1016/j.amc.2007.02.108>.
- 23 Maleknejad, K., Sohrabi, S. Numerical solution of Fredholm integral equations of the first kind by using Legendre wavelets. Applied Mathematics and Computation, 186 (1), 836–843 (2007). <https://doi.org/10.1016/j.amc.2006.08.023>.
- 24 Maleknejad, K., Lotfi, T., Mahdiani, K. Numerical solution of first kind Fredholm integral equations with wavelets-Galerkin method (WGM) and wavelets precondition. Applied Mathematics and Computation, 186 (1), 794–800 (2007). <https://doi.org/10.1016/j.amc.2006.08.027>.
- 25 Maleknejad, K., Ostadi, A. Using sinc-collocation method for solving weakly singular Fredholm integral equations of the first kind. Applicable Analysis, 96 (4), 702–713 (2017). <https://doi.org/10.1080/00036811.2016.1153629>.
- 26 Karimi, S., Jozi, M. A new iterative method for solving linear Fredholm integral equations using the least squares method. Applied Mathematics and Computation, 250, 744–758 (2015). <https://doi.org/10.1016/j.amc.2014.10.131>.
- 27 Michel, V., Orzłowski, S. On the null space of a class of Fredholm integral equations of the first kind. Journal of Inverse and Ill-Posed Problems, 24 (6), 687–710 (2016). <https://doi.org/10.1515/jiip-2015-0026>.
- 28 Tamabay, D., Temirbekov, N., Zhumagulov, B. Approximate solution of nonlinear integral Hammerstein equations by projection method using multiwavelets. AIP Conference Proceedings, 2879 (1) (2023). <https://doi.org/10.1063/5.0175541>.
- 29 Maleknejad, K., Karami, M. Numerical solution of non-linear Fredholm integral equations by using multiwavelets in the Petrov–Galerkin method. Applied Mathematics and Computation, 168 (1), 102–110 (2005). <https://doi.org/10.1016/j.amc.2004.08.047>.

- 30 Walnut, D.F. An Introduction to Wavelet Analysis. Springer Science & Business Media, 2013.
- 31 Adibi, H., Assari, P. Chebyshev wavelet method for numerical solution of Fredholm integral equations of the first kind. Mathematical Problems in Engineering, 2010 (1), Article ID 138408 (2010). <https://doi.org/10.1155/2010/138408>.
- 32 Temirbekov, N.M., Temirbekova, L.N., Nurmangaliyeva M.B. Numerical solution of the first kind Fredholm integral equations by projection methods with wavelets as basis functions. TWMS Journal of Pure and Applied Mathematics, 13 (1), 105–118 (2022).
- 33 Angadi, L.M. Laguerre wavelet based Galerkin method for the numerical solution of singular boundary value problems. International Journal of Modern Mathematical Sciences, 19 (1), 34–44 (2021).
- 34 Alpert, N.M. et al. Optimization of dynamic measurement of receptor kinetics by wavelet denoising. Neuroimage, 30 (2), 444–451 (2006). <https://doi.org/10.1016/j.neuroimage.2005.09.031>.

^{1,2}Тамабай Д.,

докторант, ORCID ID: 0000-0001-8315-5849

e-mail: dtamabay@gmail.com

^{1,2}Темирбеков А.,

PhD, қауымдастырылған профессор, ORCID ID: 0000-0002-4157-2799

e-mail: almas_tem@mail.ru

^{1,2}Тлеулесова А.,

PhD, ORCID ID: 0000-0001-9280-1048

e-mail: aigerim1985_06@mail.ru

^{1,2*}Муханова Т.,

студент, ORCID ID: 0009-0000-8522-3390

*e-mail: tomirismukhanova04@gmail.com

¹Өл-Фараби атындағы Қазақ ұлттық университеті, Алматы қ., Қазақстан

²ҚР Ұлттық инженерлік академиясы, Алматы қ., Қазақстан

АЛЬПЕРТ ВЕЙВЛЕТ БАЗИСТЕРІН ҚОЛДАНА ОТЫРЫП ФРЕДГОЛЬМНІҢ БІРІНШІ ТҮРДЕГІ ИНТЕГРАЛДЫҚ ТЕҢДЕУІН БУБНОВ–ГАЛЕРКИН ӘДІСІМЕН САНДЫҚ ШЕШУ

Аңдатпа

Бұл мақалада Фредгольмнің бірінші түрдегі интегралдық теңдеулерін Бубнов–Галеркин әдісімен шешудің сандық тәсілі ұсынылады, мұнда базистік функциялар ретінде Альперт вейвлеттері пайдаланылады. Аталған теңдеулер қойылымының дұрыс еместігімен (иллю-пост) ерекшеленеді, яғни кіріс деректеріндегі шағын өзгерістер шешімге айтарлықтай әсер етуі мүмкін. Сондықтан мұндай теңдеулерді шешу үшін орнықты және дәл сандық әдістер қажет. Ұсынылып отырған әдісте орто-нормаланған және компакты тірегі бар Альперт вейвлеттері қолданылады, олар локализация қасиеттерімен ерекшеленеді және интегралдық операторды проекциялағанда жақсы жағдайланған, сиретілген (разреженные) матрицалар береді. Бұл өз кезегінде сандық орнықтылықты арттырып, есептеу күрделілігін азайтады. Өртүрлі тор тығыздықтары мен полиномдық дәрежелер үшін сандық эксперименттер жүргізілді. Әдістің дәлдігі нақты аналитикалық шешіммен салыстыру арқылы бағаланды. Нәтижелер абсолюттік кателіктердің өте кіші екенін, кейбір жағдайларда машиналық дәлдікке жақындайтынын көрсетті. Сонымен қатар, полиномдық базистермен салыстырмалы талдау Альперт вейвлеттеріне негізделген тәсілдің жуықтау сапасы мен жинақтылық жылдамдығы жағынан басымдығын растады. Жалпы алғанда, әдіс тиімділігімен, орнықтылығымен және көпөлшемді немесе шуды деректер жағдайларына кеңейту мүмкіндігімен ерекшеленеді. Бұл Альперт вейвлеттері негізінде құрастырылған Галеркин сұлбаларының қолданбалы ғылымдардағы дұрыс қойылмаған және кері есептерді сандық шешуге арналған сенімді құрал екенін көрсетеді.

Тірек сөздер: Фредгольмнің бірінші түрдегі интегралдық теңдеуі, Бубнов–Галеркин әдісі, Альперт вейвлеттері, дұрыс қойылмаған есеп, орто-нормаланған базис.

^{1,2}**Тамабай Д.,**

докторант, ORCID ID: 0000-0001-8315-5849

e-mail: dtamabay@gmail.com

^{1,2}**Темирбеков А.,**

PhD, ассоциированный профессор, ORCID ID: 0000-0002-4157-2799

e-mail: almas_tem@mail.ru

^{1,2}**Тлеулесова А.,**

PhD, ORCID ID: 0000-0001-9280-1048

e-mail: aigerim1985_06@mail.ru

^{1,2*}**Муханова Т.,**

студент, ORCID ID: 0009-0000-8522-3390

*e-mail: tomirismukhanova04@gmail.com

¹Казахский национальный университет им. аль-Фараби, г. Алматы, Казахстан²Национальная инженерная академия РК, г. Алматы, Казахстан

ЧИСЛЕННОЕ РЕШЕНИЕ ИНТЕГРАЛЬНОГО УРАВНЕНИЯ ФРЕДГОЛЬМА ПЕРВОГО РОДА МЕТОДОМ БУБНОВА–ГАЛЁРКИНА С ИСПОЛЬЗОВАНИЕМ БАЗИСОВ ВЕЙВЛЕТОВ АЛЬПЕРТА

Аннотация

В данной статье представлен численный подход к решению интегральных уравнений Фредгольма первого рода с применением метода Бубнова–Галёркина и базисных функций в виде вейвлетов Альперта. Эти уравнения известны своей некорректностью: малейшие изменения во входных данных могут вызывать существенные отклонения в решении. В связи с этим необходимы устойчивые и точные численные методы. Предлагаемый метод использует ортонормированные и компактно поддерживаемые вейвлеты Альперта, обладающие отличными локализационными свойствами. Они обеспечивают хорошо обусловленные и разреженные матрицы системы при проецировании интегрального оператора, что повышает численную устойчивость и снижает вычислительную сложность. Серия численных экспериментов была проведена с различными уровнями уточнения и степенями полиномов. Точность метода оценивалась путем сравнения приближенных решений с точным аналитическим решением. Результаты показали исключительно малые абсолютные ошибки, зачастую близкие к машинной точности. Кроме того, сравнительный анализ с базисами из степенных полиномов подтвердил превосходство подхода, основанного на вейвлетах Альперта, как по скорости сходимости, так и по качеству аппроксимации. В целом метод продемонстрировал эффективность, устойчивость и пригодность к дальнейшему расширению на более сложные классы интегральных уравнений, включая многомерные задачи и задачи с зашумленными данными. Это подтверждает потенциал вейвлетных схем Галёркина с базисами Альперта как надежного инструмента для численного решения некорректных и обратных задач в прикладных науках.

Ключевые слова: интегральное уравнение Фредгольма первого рода, метод Бубнова–Галёркина, вейвлет Альперта, некорректность задачи, ортонормированный базис.

Article submission date: 20.06.2025

УДК 517.54
МРНТИ 27.23.25

<https://doi.org/10.55452/1998-6688-2025-22-4-279-294>

¹Майер Ф.Ф.,

канд. физ.-матем. наук, профессор, ORCID ID: 0000-0002-2278-2723,
e-mail: maiyer@mail.ru

¹Тастанов М.Г.,

канд. физ.-матем. наук, профессор, ORCID ID: 0000-0003-1926-8958,
e-mail: tastao@mail.ru

^{1*}Утемисова А.А.,

канд. пед. наук, ассоциированный профессор, ORCID ID: 0000-0001-5143-0260
*e-mail: anar_utemisova@mail.ru

¹Калаков Б.А.

канд. физ.-матем. наук, ORCID ID: 0009-0003-6007-0254
e-mail: kalakov1968@gmail.com

¹Костанайский региональный университет им. Ахмета Байтурсынулы,
г. Костанай, Казахстан

О НЕКОТОРЫХ ПОДКЛАССАХ КЛАССОВ ПОЧТИ ВЫПУКЛЫХ И ДВАЖДЫ ПОЧТИ ВЫПУКЛЫХ ФУНКЦИЙ

Аннотация

Известно, что класс K почти выпуклых функций задается условием положительности функционала $Re \left(\frac{zf'(z)}{g(z)} \right)$ со звездообразной функцией $g(z)$. Замена звездообразной функции $g(z)$ на выпуклую приводит к известному подклассу C' класса K . В настоящей статье вводится обобщение класса C' на случай, когда множество значений $\frac{zf'(z)}{g(z)}$ содержится в области специального вида, которая в частном случае может совпадать и с полуплоскостью. Также обобщение связано с расширением класса C' до определенного подкласса класса нормированных дважды почти выпуклых функций. Многообразие частных случаев области специального вида и переход к дважды почти выпуклым функциям позволяет получить как новые оригинальные результаты, так и обобщения ранее известных результатов. Основные исследования данной статьи направлены на доказательство теорем об искажении, нахождение радиусов выпуклости рассматриваемых классов функций и обоснование точности полученных результатов. Также установлена связь введенного класса функций с некоторым новым классом дважды почти звездообразных функций. Для данного класса и его подклассов также получены новые результаты в виде теорем о росте и радиусе звездообразности и обобщения ранее известных результатов.

Ключевые слова: звездообразные функции, почти звездообразные функции, почти выпуклые функции, теоремы искажения, радиусы выпуклости, теоремы роста, радиусы звездообразности.

Введение

Рассматриваются функции, аналитические в единичном круге $E = \{z: |z| < 1\}$. Класс таких функций с разложением вида $\varphi(z) = 1 + c_n z^n + c_{n+1} z^{n+1} + \dots$, $n \geq 1$, обозначим через A_n , а класс функций $f(z) = z + a_{n+1} z^{n+1} + a_{n+2} z^{n+2} + \dots$, $n \geq 1$, – через N_n . Также положим $A := A_1$, $N := N_1$ и пусть P – класс Каратеодори функций $\varphi(z)$ из A с положительной вещественной частью.

Известными подклассами класса N являются класс C выпуклых функций и класс S^* звездообразных функций, а также их подклассы

$$C(\alpha) = \left\{ f \in N: 1 + \operatorname{Re} \left(z \frac{f''(z)}{f'(z)} \right) > \alpha, 0 \leq \alpha < 1, z \in E \right\},$$

$$S^*(\alpha) = \left\{ f \in N: \operatorname{Re} \left(z \frac{f'(z)}{f(z)} \right) > \alpha, 0 \leq \alpha < 1, z \in E \right\}.$$

При этом $C := C(0)$ и $S^* := S^*(0)$. Опираясь на классы C и S^* , в [1] был представлен класс K почти выпуклых функций.

Напомним [1], что функция $f \in K$ тогда и только тогда, когда для некоторой функции $g \in S^*$ выполняется условие

$$\operatorname{Re} \left(z \frac{f'(z)}{g(z)} \right) > 0, z \in E. \quad (1)$$

Поскольку $zg'(z) \in S^*$ тогда и только тогда, когда $g \in C$, то условие (1) равносильно условию

$$\operatorname{Re} \frac{f'(z)}{g'(z)} > 0, z \in E, \quad (2)$$

в котором $g \in C$.

Рассмотрим также класс $K(\gamma)$ функций $f(z)$, почти выпуклых порядка γ [2], [3], который задается условием

$$\left| \arg \frac{zf'(z)}{g(z)} \right| < \gamma \frac{\pi}{2}, \quad 0 < \gamma \leq 1, g \in S^*, z \in E.$$

Отметим, что $C = K(0) \subset K(\gamma) \subset K(1) = K$ и $S^* \subset K$.

Если в условии (2) вместо выпуклой функции $g(z)$ использовать почти выпуклую функцию, то получим класс CK дважды почти выпуклых функций (doubly close-to-convex functions) $f(z)$. При этом $K \subset CK$ и функции класса CK могут быть и неоднолиственными.

В статье исследуются свойства некоторых подклассов класса почти выпуклых функций и класса дважды почти выпуклых функций, включая теоремы об искажении и о радиусе выпуклости. Устанавливается взаимосвязь данных классов с другими классами аналитических функций, на основе которой выводятся свойства функций этих классов. Представленные здесь результаты обобщают ряд ранее известных работ.

Материалы и методы

Концепция подчиненности аналитических функций имеет плодотворные применения при исследовании экстремальных свойств аналитических функций, различные классы которых также удобно задавать в форме подчиненности.

Всюду ниже будем считать, что $\varphi(z)$ и $\varphi_0(z)$ – аналитические в круге E функции и $\varphi_0(z)$ однолистка в E .

Определение 1. Если выполняется условие согласования нормировок $\varphi(0) = \varphi_0(0)$ и условие вложения областей $\varphi(E) \subset \varphi_0(E)$, то функция $\varphi(z)$ называется подчиненной функции $\varphi_0(z)$.

При этом используется обозначение $\varphi(z) \prec \varphi_0(z)$, а функция $\varphi_0(z)$ называется мажорантой подчинения.

В теории подчиненности важным фактом является то, что соотношение $\varphi(E) \subset \varphi_0(E)$ распространяется на все внутренние круги $|z| < r$, $0 \leq r < 1$, что позволяет многие свойства мажоранты подчинения переносить на подчиненную функцию.

Определение 1. Будем говорить, что аналитическая в E функция $\varphi(z)$ из A_n принадлежит классу $P_n(m, \gamma)$ тогда и только тогда, когда выполняется условие

$$\left| (\varphi(z))^{\frac{1}{\gamma}} - m \right| < m, \quad m > \frac{1}{2}, \quad \gamma \in [-1; 1] \setminus \{0\}, z \in E. \quad (3)$$

При этом будем считать, что $P(m, \gamma) := P_1(m, \gamma)$.

Условие (3) равносильно подчиненности $\varphi(z) \prec \varphi_0(z)$, где мажоранта подчинения $\varphi_0(z)$ задается по формуле

$$\varphi_0(z) = \left(\frac{1 + \varepsilon z}{1 - c\varepsilon z} \right)^{\gamma}, \quad c = 1 - \frac{1}{m}, \quad -1 < c \leq 1, \varepsilon = \text{sign}(\gamma). \quad (4)$$

При разных значениях параметров областью значений $\varphi_0(E)$ может быть круг, полуплоскость и область, ограниченная гиперболой или правой половиной лемнискаты Бернулли.

Лемма 1. Пусть $\varepsilon = \text{sign}(\gamma)$, $c = 1 - \frac{1}{m}$. Тогда для любой функции $\varphi \in P_n(m, \gamma)$ при $|z| = r$, $0 \leq r < 1$, выполняются оценки

$$\begin{aligned} \left(\frac{1 - \varepsilon r^n}{1 + \varepsilon c r^n} \right)^{\gamma} &\leq \text{Re } \varphi(z) \leq |\varphi(z)| \leq \left(\frac{1 + \varepsilon r^n}{1 - \varepsilon c r^n} \right)^{\gamma}, \\ \left| z \frac{\varphi'(z)}{\varphi(z)} \right| &\leq \frac{|\gamma|(1+c)nr^n}{(1-r^n)(1+cr^n)}, \end{aligned} \quad (5)$$

которые достигаются для функции $\varphi(z) = \varphi_0(z^n)$, где $\varphi_0(z)$ задана по формуле (4).

Для случая, когда $0 < \gamma \leq 1$, лемма 1 доказана в [4], а при $-1 \leq \gamma < 0$ лемма 1 вытекает из следствия 1 [4].

В статье [5] были исследованы свойства функций $f(z)$ из N , которые удовлетворяют условию (1), но не со звездобразной функцией $g(z)$, а с выпуклой функцией $g(z)$. В дальнейшем этот класс функций исследовался в [6], где было введено обозначение класса C' . Поскольку $C \subset S^*$, то в силу определения класса C' получаем, что $C' \subset K$.

Опираясь на концепцию подчиненности аналитических функций, в статьях [7, 8] были введены и исследовались подклассы $C'[A, B]$ и $C'(\alpha, \beta)$ класса C' , заданные, соответственно, условиями

$$z \frac{f'(z)}{g(z)} \prec \frac{1 + Az}{1 + Bz}, \quad -1 \leq B < A \leq 1, g \in C$$

и

$$z \frac{f'(z)}{g(z)} \prec \frac{1 + (2\alpha - 1)\beta z}{1 + \beta z}, \quad 0 \leq \alpha < 1, 0 < \beta \leq 1, g \in C.$$

Заметим, что $C'[1, -1] = C'(0, 1) = C'$.

В [9] исследовался класс $K'_C(\alpha)$ функций $f \in N$ таких, что

$$\text{Re} \frac{f'(z)}{g'(z)} > \alpha,$$

где $g \in C'$. Поскольку $g \in C'$ и $C' \subset K$, то класс $K'_C(\alpha)$ является подклассом класса CK дважды почти выпуклых функций.

С почти выпуклыми функциями тесно связаны почти звездобразные функции (close-to-starlike functions), класс K^* которых был введен в [10] с помощью условия

$$\text{Re} \frac{f(z)}{g(z)} > 0, \quad z \in E, \quad (6)$$

где $g \in S^*$. Если в условии (6) $g \in K^*$, то получаем класс CK^* дважды почти звездобразных функций (doubly close-to-starlike functions). При этом, $S^* \subset K^* \subset CK^*$.

К первым исследованиям в этом направлении можно отнести статьи [11–12]. В них найдены радиусы звездобразности классов почти звездобразных функций, заданных условиями

вида $Re \frac{f(z)}{g(z)} > 0$ и $\left| \frac{f(z)}{g(z)} - 1 \right| < 1$, в которых $g \in S^*$. Для нас интересными будут два аналогичных результата из [11–12], в которых $g(z)$ является выпуклой. Именно в [11–12] найдены радиусы звездообразности $r_1^* = 1/3$ и $r_2^* = \sqrt{2} - 1$ классов, заданных условиями $Re \frac{f(z)}{g(z)} > 0$ и $\left| \frac{f(z)}{g(z)} - 1 \right| < 1$, в которых $g \in \mathcal{C}$.

Между классами K и K^* можно установить простую связь, которая выражается соотношением $F \in K \Leftrightarrow f = zF'(z) \in K^*$. Учитывая также, что $F \in \mathcal{C} \Leftrightarrow f = zF'(z) \in S^*$, на основе приведенных выше результатов из [11–12] получаем, что радиусы выпуклости классов функций, заданных условиями $Re \left(z \frac{f'(z)}{g(z)} \right) > 0$ и $\left| z \frac{f'(z)}{g(z)} - 1 \right| < 1$, где $g \in \mathcal{C}$, задается формулами $r_1^o = 1/3$ и $r_2^o = \sqrt{2} - 1$. То есть в [11–12], по сути, найдены радиусы выпуклости класса $\mathcal{C}'$ и его подкласса, заданного условием $\left| z \frac{f'(z)}{a(z)} - 1 \right| < 1$.

Всюду ниже будем считать, что $a, b > \frac{1}{2}$ и $\gamma, \delta \in [-1; 1] \setminus \{0\}$.

Определение 2. Пусть $\mathcal{C}'(a, \gamma)$ – класс функций $f \in N$, для каждой из которых существует функция $g \in \mathcal{C}$ такая, что $\frac{zf'(z)}{g(z)} \in P(a, \gamma)$.

Таким образом,

$$f \in \mathcal{C}'(a, \gamma) \Leftrightarrow \left| \left(\frac{zf'(z)}{g(z)} \right)^{\frac{1}{\gamma}} - a \right| < a, \quad g \in \mathcal{C}, z \in E.$$

Отметим, что класс $\mathcal{C}'(a, \gamma)$, являясь подклассом класса K , содержит только однолистные (почти выпуклые) функции.

Покажем, что класс $\mathcal{C}'(a, \gamma)$ не является пустым множеством.

Пусть $\gamma \in [-1; 1] \setminus \{0\}$, $c = 1 - \frac{1}{a}$, $a > \frac{1}{2}$. Рассмотрим функции

$$f_0(z) = \int_0^z \left(\frac{1+t}{1-ct} \right)^{\gamma} \frac{dt}{1-t}, \quad g_0(z) = \frac{z}{(1-cz)}.$$

Тогда $g_0 \in \mathcal{C}$ и

$$\left(\frac{zf_0'(z)}{g_0(z)} \right)^{\frac{1}{\gamma}} = \frac{1+z}{1-cz}.$$

Так как функция $w_0(z) = \frac{1+z}{1-cz}$ однолистно отображает круг E на круг $|w - a| < a$, то функция $f_0(z)$ удовлетворяет условию

$$\left| \left(\frac{zf'(z)}{g(z)} \right)^{\frac{1}{\gamma}} - a \right| < a, \quad z \in E.$$

Следовательно, $f_0 \in \mathcal{C}'(a, \gamma)$ и $\mathcal{C}'(a, \gamma) \neq \emptyset$.

Рассмотрим другие примеры функций класса $\mathcal{C}'(a, \gamma)$.

Пример 1. Пусть $\gamma \in [-1; 1] \setminus \{0\}$, $c = 1 - \frac{1}{a}$, $0 \leq c < 2 - \sqrt{3}$ и

$$f_1(z) = \frac{1}{(1+\gamma)(1+c)} \left[\left(\frac{1+z}{1-cz} \right)^{1+\gamma} - 1 \right], \quad g_1(z) = \frac{z}{(1-cz)^2}.$$

Поскольку в круге E для $c \in [0; 2 - \sqrt{3}]$ выполняется условие

$$Re \left(\frac{zg_1''(z)}{g_1'(z)} \right) = Re \left(\frac{cz}{1+cz} + \frac{3cz}{1-cz} \right) > -\frac{c}{1-c} - \frac{3c}{1+c} = \frac{c^2 - 4c + 1}{1-c^2} \geq -1,$$

то $g_1 \in S^0$. Так как

$$f_1'(z) = \left(\frac{1+z}{1-cz} \right)^{\gamma} \frac{1}{(1-cz)^2},$$

то

$$\left(\frac{zf_1'(z)}{g_1(z)} \right)^{\frac{1}{\gamma}} = \frac{1+z}{1-cz}.$$

Поэтому выполняется условие

$$\left| \left(\frac{zf_1'(z)}{g_1(z)} \right)^{\frac{1}{\gamma}} - a \right| < a,$$

следовательно, $f_1 \in C'(a, \gamma)$.

При $c = 0$ функция $g_1(z) = z$ и мы получаем более простой пример функции класса $C'(1, \gamma)$:

$$f_1(z) = \frac{1}{(1+\gamma)} [(1+z)^{1+\gamma} - 1].$$

Пример 2. Пусть

$$f_2(z) = \frac{2z}{1-z} + \ln \ln(1-z), \quad g_2(z) = \frac{z}{1-z}.$$

Тогда $g_2 \in S^0$ и $f_2 \in C'(\infty, 1)$, так как

$$\operatorname{Re} \frac{zf_2'(z)}{g_2(z)} = \operatorname{Re} \frac{1+z}{1-z} > 0, z \in E.$$

Пример 3. Пусть $0 \leq k \leq 1/4$ и пусть

$$f_3(z) = -(1+2k)z - \frac{k}{2}z^2 - 2(1+k) \ln(1-z), \quad g_3(z) = z + kz^2,$$

причем, в силу ограничения на параметр k функция $g_3 \in S^0$. Поскольку

$$f_3'(z) = -(1+2k) - kz + \frac{2(1+k)}{1-z} = \frac{(1+z)(1+kz)}{1-z},$$

то

$$\frac{zf_3'(z)}{g_3(z)} = \frac{1+z}{1-z}.$$

Поэтому $f_3 \in C'(\infty, 1)$.

При $k = 0$ получаем функцию

$$f_3(z) = -z - 2 \ln(1-z) \in C'(\infty, 1).$$

Пример 4. Пусть $f_4(z) = z + kz^2$, $0 \leq k \leq 1/2$, и пусть $g_4(z) = z$. Тогда

$$\frac{zf_4'(z)}{g_4(z)} = 1 + 2kz.$$

Поэтому

$$\left| \frac{zf_4'(z)}{g_4(z)} - 1 \right| = 2k|z| < 1.$$

Следовательно, $f_4 \in C'(1, 1)$.

С другой стороны,

$$\left| \arg \arg \frac{zf_4'(z)}{g_4(z)} \right| = |\arg \arg(1+2kz)| \leq \arcsin \arcsin(2k) = \gamma \frac{\pi}{2},$$

где $\gamma = \frac{2}{\pi} \arcsin \arcsin(2k)$. Поэтому $f_4 \in C'(\infty, \gamma)$.

Определение 3. Пусть $K'_C(a, \gamma, b, \delta)$ – класс функций $f \in N$, для каждой из которых существует функция $g \in C'(b, \delta)$ такая, что $\frac{f'(z)}{g'(z)} \in P(a, \gamma)$.

Из определения 3 вытекает, что любая функция $f \in K'_C(a, \gamma, b, \delta)$ должна удовлетворять неравенству

$$\left| \left(\frac{f'(z)}{g'(z)} \right)^{\frac{1}{\gamma}} - a \right| < a, z \in E, \quad (7)$$

причем

$$\left| \left(\frac{zg'(z)}{h(z)} \right)^{\frac{1}{\delta}} - b \right| < b, h \in C, z \in E. \quad (8)$$

Из условия (8) следует, что функция $g(z)$ является почти выпуклой. Поэтому в силу условия (7) функция $f(z)$ является дважды почти выпуклой.

Заметим, что $C'(a, \gamma) \subset K'_C(a, \gamma)$ и $C'(a, \gamma) \subset C'$. Кроме того, $C'(a, \gamma) := K'_C(a, \gamma, \infty, 0)$, $C'(b, \delta) := K'_C(\infty, 0, b, \delta)$ и $K'_C(a) := K'_C(\frac{1}{2a}, -1, \infty, 1)$, причем $K'_C(a, \gamma, b, \delta) \subset K'_C(0)$.

Рассмотрим примеры функций класса $K'_C(a, \gamma, b, \delta)$. Кроме того, покажем, что класс $K'_C(a, \gamma, b, \delta)$ может содержать и неоднолистные функции.

Пример 5. Пусть $0 < \delta \leq 1$, $0 < \sigma \leq 2$, $0 < \gamma = \sigma - \delta \leq 1$, $c = 1 - \frac{1}{a}$, $a > \frac{1}{2}$ и

$$f_5(z) = \frac{1}{(1+\sigma)(1+c)} \left[\left(\frac{1+z}{1-cz} \right)^{1+\sigma} - 1 \right],$$

$$g_5(z) = \frac{1}{(1+\delta)(1+c)} \left[\left(\frac{1+z}{1-cz} \right)^{1+\delta} - 1 \right], h_5(z) = \frac{z}{(1-cz)^2}.$$

Тогда

$$\frac{zg'_5(z)}{h_5(z)} = \left(\frac{1+z}{1-cz} \right)^{\delta}.$$

Поэтому $g_5 \in C'(a, \delta)$.

Поскольку

$$\frac{f'_5(z)}{g'_5(z)} = \left(\frac{1+z}{1-cz} \right)^{\gamma},$$

то

$$\left| \left(\frac{f'_5(z)}{g'_5(z)} \right)^{\frac{1}{\gamma}} - a \right| < a, z \in E.$$

В силу этого, с учетом того, что $g_5 \in C'(a, \delta)$, по определению 3 получаем, что $f_5 \in K'_C(a, \gamma, a, \delta)$.

Рассмотрим частный случай, когда $\sigma = 2$, $\delta = 1$ и $c = 0$.

Тогда

$$f_5(z) = \frac{1}{3} [(1+z)^3 - 1], g_5(z) = \frac{1}{2} [(1+z)^2 - 1], h_5(z) = z.$$

Поскольку $\frac{zg'_5(z)}{h_5(z)} = 1+z$, то $\left| \frac{zg'_5(z)}{h_5(z)} - 1 \right| < 1, z \in E$, и $g_5 \in C'(1, 1)$. Аналогично $\frac{f'_5(z)}{g'_5(z)} = 1+z$, поэтому $f_5 \in K'_C(1, 1, 1, 1)$.

Заметим, что функция $f_5(z) = [(1+z)^3 - 1]/3$ не является однолистной в круге E .

Для доказательства этого возьмем две точки $z_1 \neq z_2$ из круга E и покажем, что $f_5(z_1) = f_5(z_2)$. Пусть

$$z_1 = -\frac{3}{4} - \frac{\sqrt{3}}{4}i, z_2 = \underline{z_1} = -\frac{3}{4} + \frac{\sqrt{3}}{4}i.$$

Обозначим $\tau = e^{2\pi i/3} = -\frac{1}{2} + \frac{\sqrt{3}}{2}i$. Тогда $1+z_1 = -\tau/2$, $1+z_2 = -\tau^2/2$. Поэтому $(1+z_1)^3 = -1/8$, $(1+z_2)^3 = -1/8$. Следовательно, $f_5(z_1) = -3/8 = f_5(z_2)$, то есть функ-

ция $f_5(z)$ не является однолистной в круге E .

Таким образом, мы доказали, что класс $K'_C(1,1,1,1)$, а значит, и класс $K'_C(a, \gamma, a, \delta)$ содержит и неоднолистные функции.

Пример 6. Пусть $f_6(z) = \frac{z(5-z)}{1-z} + 4 \ln(1-z)$, $g_6(z) = -z - 2 \ln(1-z)$ и $h_6(z) = z \in C$.

Тогда

$$f'_6(z) = \left(\frac{1+z}{1-z} \right)^2, \quad g'_6(z) = \frac{1+z}{1-z}$$

и

$$\frac{z g'_6(z)}{h_6(z)} = \frac{1+z}{1-z}.$$

Поэтому $g_6 \in C'(\infty, 1)$. С учетом этого, так как

$$\frac{f'_6(z)}{g'_6(z)} = \frac{1+z}{1-z},$$

то $f_6 \in K'_C(\infty, 1, \infty, 1)$.

Определение 4. Пусть $C^*(a, \gamma)$ – класс функций $F \in N$, для каждой из которых существует функция $g \in C$ такая, что $\frac{F(z)}{g(z)} \in P(b, \gamma)$.

Таким образом,

$$F \in C^*(a, \gamma) \Leftrightarrow \left| \left(\frac{F(z)}{g(z)} \right)^{\frac{1}{\gamma}} - a \right| < a, \quad g \in C, z \in E.$$

Определение 5. Пусть $K^*_C(a, \gamma, b, \delta)$ – класс функций $F \in N$, для каждой из которых существует функция $G \in C^*(b, \delta)$ такая, что $\frac{F(z)}{G(z)} \in P(a, \gamma)$.

Из определения 5 следует, что любая функция $F \in K^*_C(a, \gamma, b, \delta)$ должна удовлетворять неравенству

$$\left| \left(\frac{F(z)}{G(z)} \right)^{\frac{1}{\gamma}} - a \right| < a, \quad z \in E, \quad (9)$$

где

$$\left| \left(\frac{G(z)}{h(z)} \right)^{\frac{1}{\delta}} - b \right| < b, \quad h \in C, z \in E. \quad (10)$$

Поскольку $h \in C$, то $h \in S^*$. Кроме того, если функция $G(z)$ удовлетворяет условию (10), то она удовлетворяет и условию $\operatorname{Re} \frac{G(z)}{h(z)} > 0$, где $h \in S^*$. Поэтому функции класса $C^*(b, \delta)$ являются почти звездообразными функциями, то есть $C^*(b, \delta) \subset K^*$. Аналогично, в силу (9)–(10) класс $K^*_C(a, \gamma, b, \delta)$ является подклассом класса CK^* дважды почти звездообразных функций.

Пример 7. Пусть $\gamma, \delta \in [-1; 1] \setminus \{0\}$, $c = 1 - \frac{1}{a}$, $d = 1 - \frac{1}{b}$, $a, b > \frac{1}{2}$ и

$$F(z) = z \left(\frac{1+z}{1-cz} \right)^{\gamma} \left(\frac{1+z}{1-dz} \right)^{\delta}, \quad G(z) = z \left(\frac{1+z}{1-dz} \right)^{\delta}, \quad h(z) = z.$$

Тогда $G \in C^*(b, \delta)$, поскольку

$$\left(\frac{G(z)}{h(z)} \right)^{\frac{1}{\delta}} = \frac{1+z}{1-dz}$$

и поэтому выполняется условие (10). С другой стороны,

$$\left(\frac{F(z)}{G(z)}\right)^{\frac{1}{\gamma}} = \frac{1+z}{1-cz}.$$

Поэтому выполняется условие (9), где $G \in C^*(b, \delta)$. Следовательно, по определению 5 $F \in K_C^*(a, \gamma, b, \delta)$.

В частности, если $\gamma = \delta = c = d = 1$, то $G(z) = z^{\frac{1+z}{1-z}} \in C^*(1, 1)$. При этом $G'(z) = 0$ в точке $z = 1 - \sqrt{2}$ круга E . Следовательно, класс $C^*(1, 1)$ почти звездообразных функций содержит и неоднолистные функции.

Между классами $C'(a, \gamma)$, $K_C'(a, \gamma, b, \delta)$ и классами $C^*(a, \gamma)$, $K_C^*(a, \gamma, b, \delta)$ имеется простая связь

$$f \in C'(a, \gamma) \Leftrightarrow F = zf'(z) \in C^*(a, \gamma)$$

и

$$f \in K_C'(a, \gamma, b, \delta) \Leftrightarrow F = zf'(z) \in K_C^*(a, \gamma, b, \delta). \quad (11)$$

Целью статьи является исследование свойств функций классов $C'(a, \gamma)$ и $K_C'(a, \gamma, b, \delta)$, в том числе теорем об искажении и радиусе выпуклости, а также на основе взаимосвязи (11) – свойств функций классов $C^*(a, \gamma)$ и $K_C^*(a, \gamma, b, \delta)$. Тем самым, получить новые результаты и обобщения ряда результатов из ранних работ.

Результаты и обсуждение

1. Теорема искажения в классе $K_C'(a, \gamma, b, \delta)$

Теорема 1. Пусть $a_1 = 1 - \frac{1}{a}$, $b_1 = 1 - \frac{1}{b}$ и $\varepsilon_\gamma = \text{sign}(\gamma)$, $\varepsilon_\delta = \text{sign}(\delta)$. Тогда для любой функции $f \in K_C'(a, \gamma, b, \delta)$ при $|z| = r$, $0 \leq r < 1$, имеет место точная оценка

$$\begin{aligned} \left(\frac{1 - \varepsilon_\gamma r}{1 + \varepsilon_\gamma a_1 r}\right)^\gamma \left(\frac{1 - \varepsilon_\delta r}{1 + \varepsilon_\delta b_1 r}\right)^\delta \frac{1}{1+r} &\leq |f'(z)| \leq \\ &\leq \left(\frac{1 + \varepsilon_\gamma r}{1 - \varepsilon_\gamma a_1 r}\right)^\gamma \left(\frac{1 + \varepsilon_\delta r}{1 - \varepsilon_\delta b_1 r}\right)^\delta \frac{1}{1-r}. \end{aligned} \quad (12)$$

Доказательство. Пусть $\varphi(z) = \frac{f'(z)}{g'(z)}$, $\psi(z) = \frac{zg'(z)}{h(z)}$. Тогда

$$f'(z) = \varphi(z)\psi(z) \frac{h(z)}{z}.$$

Поскольку $\varphi \in P(a, \gamma)$, $\psi \in P(b, \delta)$, то на основании леммы 1 при $n = 1$ для всех z , $|z| = r$, $0 \leq r < 1$, выполняются неравенства

$$\begin{aligned} \left(\frac{1 - \varepsilon_\gamma r}{1 + \varepsilon_\gamma a_1 r}\right)^\gamma &\leq |\varphi(z)| \leq \left(\frac{1 + \varepsilon_\gamma r}{1 - \varepsilon_\gamma a_1 r}\right)^\gamma, \\ \left(\frac{1 - \varepsilon_\delta r}{1 + \varepsilon_\delta b_1 r}\right)^\delta &\leq |\psi(z)| \leq \left(\frac{1 + \varepsilon_\delta r}{1 - \varepsilon_\delta b_1 r}\right)^\delta. \end{aligned}$$

В силу этого

$$\left(\frac{1 - \varepsilon_\gamma r}{1 + \varepsilon_\gamma a_1 r}\right)^\gamma \left(\frac{1 - \varepsilon_\delta r}{1 + \varepsilon_\delta b_1 r}\right)^\delta \left|\frac{h(z)}{z}\right| \leq |f'(z)| \leq \left|\frac{h(z)}{z}\right| \left(\frac{1 + \varepsilon_\gamma r}{1 - \varepsilon_\gamma a_1 r}\right)^\gamma \left(\frac{1 + \varepsilon_\delta r}{1 - \varepsilon_\delta b_1 r}\right)^\delta.$$

По теореме роста для выпуклой функции $h(z)$ (см. [13])

$$\frac{r}{1+r} \leq |h(z)| \leq \frac{r}{1-r}.$$

Поэтому предыдущая оценка вместе с данной оценкой приводят к оценке (12).

Точность оценки (12) следует из того, что для функции

$$f_0(z) = \int_0^z \left(\frac{1 + \varepsilon_\gamma t}{1 - \varepsilon_\gamma a_1 t} \right)^\gamma \left(\frac{1 + \varepsilon_\delta t}{1 - \varepsilon_\delta b_1 t} \right)^\delta \frac{dt}{1-t} \in K'_C(a, \gamma, b, \delta) \quad (13)$$

в точках $z = -r$ и $z = r$ в оценке (12), соответственно, слева и справа достигаются знаки равенства.

2. Радиус выпуклости класса $K'_C(a, \gamma, b, \delta)$

Теорема 2. Пусть $f \in K'_C(a, \gamma, b, \delta)$ и $a_1 = 1 - \frac{1}{a}$, $b_1 = 1 - \frac{1}{b}$. Тогда функция $f(z)$ отображает круг $|z| < r^0$ на выпуклую область, где r^0 – единственный на интервале $(0; 1)$ корень уравнения

$$\frac{1}{1+r} - \frac{r}{1-r} \left(|\gamma| \frac{1+a_1}{1+a_1 r} + |\delta| \frac{1+b_1}{1+b_1 r} \right) = 0. \quad (14)$$

При $0 < \gamma, \delta \leq 1$ результат точный. Экстремальная функция задается по формуле (13).

Доказательство. Если обозначить $\varphi(z) = \frac{f'(z)}{g'(z)}$, $\psi(z) = \frac{zg'(z)}{h(z)}$, то $zf'(z) = \varphi(z)\psi(z)h(z)$ и

$$1 + z \frac{f''(z)}{f'(z)} = z \frac{h'(z)}{h(z)} + z \frac{\varphi'(z)}{\varphi(z)} + z \frac{\psi'(z)}{\psi(z)}.$$

Поэтому в круге $|z| \leq r$ выполняется неравенство

$$1 + \operatorname{Re} \frac{zf''(z)}{f'(z)} \geq \operatorname{Re} \left(z \frac{h'(z)}{h(z)} \right) - \left| z \frac{\varphi'(z)}{\varphi(z)} \right| - \left| z \frac{\psi'(z)}{\psi(z)} \right|.$$

Известно (напр., [14]), что любая выпуклая функция является звездообразной порядка $\frac{1}{2}$, то есть $\operatorname{Re} \left(z \frac{h'(z)}{h(z)} \right) > 1/2$. Поэтому

$$z \frac{h'(z)}{h(z)} < \frac{1}{1+z},$$

откуда

$$\operatorname{Re} \left(z \frac{h'(z)}{h(z)} \right) = \frac{1}{1+r}.$$

Поэтому, применяя оценку (5) при $n = 1$ к функциям $\varphi \in P(a, \gamma)$ и $\psi \in P(b, \delta)$, для всех z , $|z| = r$, $0 \leq r < 1$, получаем неравенство

$$1 + \operatorname{Re} \frac{zf''(z)}{f'(z)} \geq \frac{1}{1+r} - \frac{|\gamma|(1+a_1)r}{(1-r)(1+a_1 r)} - \frac{|\delta|(1+b_1)r}{(1-r)(1+b_1 r)}.$$

Поэтому, если

$$\frac{1}{1+r} - \frac{|\gamma|(1+a_1)r}{(1-r)(1+a_1 r)} - \frac{|\delta|(1+b_1)r}{(1-r)(1+b_1 r)} = 0,$$

то будет выполняться условие выпуклости и функция $f(z)$ будет выпуклой в круге $|z| \leq r$, то есть приходим к уравнению (14). Поскольку функция $\mu_1(r) = \frac{1}{(1+r)}$ убывает на $[0; 1]$ от 1 до $1/2$, а функция

$$\mu_2(r) = \frac{r}{1-r} \left(|\gamma| \frac{1+a_1}{1+a_1 r} + |\delta| \frac{1+b_1}{1+b_1 r} \right)$$

возрастает на $[0; 1)$ от 0 до $+\infty$, то уравнение (14) на $(0; 1)$ имеет единственный корень r_0 .

Для доказательства точности радиуса выпуклости при $0 < \gamma, \delta \leq 1$ рассмотрим функцию (13) с $\varepsilon_\gamma = \varepsilon_\delta = 1$. Тогда

$$1 + z \frac{f_0''(z)}{f_0'(z)} = \frac{1}{1-z} + \frac{\gamma(1+a_1)z}{(1+z)(1-a_1z)} + \frac{\delta(1+b_1)z}{(1+z)(1-b_1z)}.$$

Поэтому в точке $z = -r_0$ получаем

$$1 + \operatorname{Re} \frac{zf_0''(z)}{f_0'(z)} = \frac{1}{1+r_0} - \frac{r_0}{1-r_0} \left(\gamma \frac{1+a_1}{1+a_1r_0} + \delta \frac{1+b_1}{1+b_1r_0} \right) = 0.$$

То есть радиус выпуклости улучшить нельзя.

3. Некоторые следствия

Положим, $\delta = \gamma = -1$, $b = \frac{1}{(2\beta)}$, $a = \frac{1}{(2\alpha)}$. Тогда класс $K'_C(a, \gamma, b, \delta)$ преобразуется в класс функций $f(z)$, удовлетворяющих условию

$$\operatorname{Re} \frac{f'(z)}{g'(z)} > \alpha, \operatorname{Re} \frac{zg'(z)}{h(z)} > \beta, h \in C, \quad (15)$$

где $0 \leq \alpha, \beta < 1$. Поскольку теперь $a_1 = 1 - 2\alpha$, $b_1 = 1 - 2\beta$, $\varepsilon_\gamma = \varepsilon_\delta = -1$, то из теорем 1–2 вытекает следующее утверждение.

Следствие 1. Если выполняется условие (15), то при $|z| = r$, $0 \leq r < 1$, справедливо неравенство

$$\frac{[1 - (1 - 2\alpha)r][1 - (1 - 2\beta)r]}{(1+r)^3} \leq |f'(z)| \leq \frac{[1 + (1 - 2\alpha)r][1 + (1 - 2\beta)r]}{(1-r)^3}$$

и функция $f(z)$ отображает круг $|z| \leq r^0$ на выпуклую область, где r^0 – единственный на интервале $(0; 1)$ корень уравнения

$$\frac{1}{1+r} - \frac{2r}{1-r} \left(\frac{1-\alpha}{1+(1-2\alpha)r} + \frac{1-\beta}{1+(1-2\beta)r} \right) = 0.$$

При $\beta = 0$ следствие 1 дает следующие результаты статьи [9].

Следствие 2 [9]. В классе $K'_C(\alpha)$ имеет место точная оценка

$$\frac{[1 - (1 - 2\alpha)r](1-r)}{(1+r)^3} \leq |f'(z)| \leq \frac{[1 + (1 - 2\alpha)r](1+r)}{(1-r)^3}$$

и функция $f(z)$ отображает круг $|z| \leq r^0$ на выпуклую область, где

$$r^0 = \frac{(2 - \sqrt{9 - 2\alpha})}{(2\alpha - 5)}.$$

Пусть $\delta = 0$, $b \rightarrow \infty$. Тогда класс $K'_C(a, \gamma, b, \delta)$ преобразуется в класс $C'(a, \gamma)$ функций $f(z)$, удовлетворяющих условию

$$\left| \left(\frac{zf'(z)}{g(z)} \right)^{\frac{1}{\gamma}} - a \right| < a, \quad g \in C.$$

Поэтому прямым следствием теорем 1 и 2 является

Следствие 3. Пусть $f \in C'(a, \gamma)$, где $0 < \gamma \leq 1$, $a > \frac{1}{2}$, и пусть $a_1 = 1 - \frac{1}{a}$ и r^0 – корень уравнения

$$\frac{1}{1+r} - \frac{\gamma r}{1-r} \frac{1+a_1}{1+a_1r} = 0,$$

принадлежащий интервалу $(0; 1)$. Тогда при $|z| = r$, $0 \leq r < 1$ справедлива оценка

$$\frac{(1-r)^{\gamma}}{(1+a_1r)^{\gamma}} \frac{1}{1+r} \leq |f'(z)| \leq \frac{(1+r)^{\gamma}}{(1-a_1r)^{\gamma}} \frac{1}{1-r}$$

и функция $f(z)$ является выпуклой в круге $|z| < r^0$. Теорема искажения и радиус выпуклости являются точными.

Учитывая, что $C' := C'(\infty, 1)$, из следствия 3 при $\gamma = 1, a \rightarrow \infty$ получаем оценку [5]

$$\frac{1-r}{(1+r)^2} \leq |f'(z)| \leq \frac{1+r}{(1-r)^2}$$

и радиус выпуклости $r_0 = 1/3$ класса C' , ранее найденный в [5], [12].

При $\gamma = 1, a = 1$ из следствия 2 вытекает точная оценка

$$\frac{1-r}{1+r} \leq |f'(z)| \leq \frac{1+r}{1-r}$$

и точный радиус выпуклости $r_0 = \sqrt{2} - 1$ класса $C'(1, 1)$ функций $f(z)$, удовлетворяющих условию $\left| \frac{zf'(z)}{g(z)} - 1 \right| < 1$, где $g \in C$. Заметим, что радиус выпуклости получен ранее в [13].

Положим $\delta = \gamma = 1, b = a = 1$. Тогда $\varepsilon_{\gamma} = \varepsilon_{\delta} = 1, a_1 = b_1 = 0$ и получаем класс $K_C'(1, 1, 1, 1)$ функций $f(z)$, заданный условием

$$\left| \frac{f'(z)}{g'(z)} - 1 \right| < 1, \left| \frac{zg'(z)}{h(z)} - 1 \right| < 1, h \in C.$$

Поэтому имеет место

Следствие 4. В классе $K_C'(1, 1, 1, 1)$ справедлива точная оценка

$$\frac{(1-r)^2}{1+r} \leq |f'(z)| \leq \frac{(1+r)^2}{1-r}$$

и радиус выпуклости $r^0 = (\sqrt{17} - 3)/2$.

4. Почти звездообразные и дважды почти звездообразные функции

Взаимосвязь (11) классов $K_C'(a, \gamma, b, \delta)$ и $K_C^*(a, \gamma, b, \delta)$ с учетом соотношения $f \in C \Leftrightarrow F = zf'(z) \in S^*$ позволяет результаты теорем 1 и 2 перенести на класс $K_C^*(a, \gamma, b, \delta)$ дважды почти звездообразных функций.

Теорема 3. Пусть $a_1 = 1 - \frac{1}{a}, b_1 = 1 - \frac{1}{b}$ и $\varepsilon_{\gamma} = \text{sign}(\gamma), \varepsilon_{\delta} = \text{sign}(\delta)$. Тогда для любой функции $F(z)$ класса $K_C^*(a, \gamma, b, \delta)$ при $|z| = r, 0 \leq r < 1$, справедлива точная оценка

$$\begin{aligned} & \left(\frac{1 - \varepsilon_{\gamma} r}{1 + \varepsilon_{\gamma} a_1 r} \right)^{\gamma} \left(\frac{1 - \varepsilon_{\delta} r}{1 + \varepsilon_{\delta} b_1 r} \right)^{\delta} \frac{r}{1+r} \leq |F(z)| \leq \\ & \leq \left(\frac{1 + \varepsilon_{\gamma} r}{1 - \varepsilon_{\gamma} a_1 r} \right)^{\gamma} \left(\frac{1 + \varepsilon_{\delta} r}{1 - \varepsilon_{\delta} b_1 r} \right)^{\delta} \frac{r}{1-r} \end{aligned} \quad (16)$$

и функция $F(z)$ отображает круг $|z| < r^*$ на звездообразную область, где r^* – корень уравнения (14).

При $0 < \gamma, \delta \leq 1$ результат точный. Экстремальная функция имеет вид

$$F_0(z) = \left(\frac{1 + \varepsilon_{\gamma} z}{1 - \varepsilon_{\gamma} a_1 z} \right)^{\gamma} \left(\frac{1 + \varepsilon_{\delta} z}{1 - \varepsilon_{\delta} b_1 z} \right)^{\delta} \frac{z}{1-z}. \quad (17)$$

Примечание 1. Обозначим через $r^*(\alpha)$ радиус звездообразности порядка α класса $K_C^*(a, \gamma, b, \delta)$. Если в теоремах 2 и 3 вместо условия выпуклости и условия звездообразности использовать условие $1 + \text{Re} \left(z \frac{f''(z)}{f'(z)} \right) > \alpha$ выпуклости порядка α (в теореме 2) и, соответ-

ственно, условие $\operatorname{Re} \left(z \frac{F'(z)}{F(z)} \right) > \alpha$ звездообразности порядка α (в теореме 3), то по аналогии с предыдущим найдем радиус $r^*(\alpha)$ как корень уравнения

$$\frac{1}{1+r} - \frac{r}{1-r} \left(|\gamma| \frac{1+a_1}{1+a_1 r} + |\delta| \frac{1+b_1}{1+b_1 r} \right) - \alpha = 0. \quad (18)$$

Поскольку частью экстремальной функции $F_0(z)$ из (17) является выпуклая функция $h_0(z) = \frac{z}{1-z}$, то частным случаем класса $K_C^*(a, \gamma, b, \delta)$ является класс $F(a, \gamma, b, \delta)$ функций $F(z)$, определенный с помощью условия

$$\left| \left(\frac{F(z)}{G(z)} \right)^{\frac{1}{\gamma}} - a \right| < a, \quad \left| \left(\frac{1-z}{z} G(z) \right)^{\frac{1}{\delta}} - b \right| < b.$$

Поэтому в качестве следствия из теоремы 3 с учетом примечания 1 получаем

Следствие 5. Пусть $a_1 = 1 - \frac{1}{a}$, $b_1 = 1 - \frac{1}{b}$, $\varepsilon_\gamma = \operatorname{sign}(\gamma)$, $\varepsilon_\delta = \operatorname{sign}(\delta)$, $a, b > \frac{1}{2}$, $\gamma, \delta \in [-1; 1] \setminus \{0\}$. Тогда в классе $F(a, \gamma, b, \delta)$ при $|z| = r$ справедлива точная оценка (16) и любая функция $F(z)$ из $F(a, \gamma, b, \delta)$ отображает круг $|z| < r^*(\alpha)$ на звездообразную область, где $r^*(\alpha)$ – корень уравнения (18), принадлежащий интервалу $(0; 1)$.

При $0 < \gamma, \delta \leq 1$ результат точный. Экстремальная функция задается по формуле (17).

В последние годы вышел ряд работ (например, [15–20]), в которых получены различные радиусы звездообразности некоторых классов почти звездообразных и дважды почти звездообразных функций.

Частными случаями класса $F(a, \gamma, b, \delta)$ являются классы, рассматриваемые в [15]:

$$\begin{aligned} F_1 &= F(\infty, 1, \infty, 1) = \left\{ F(z) : \operatorname{Re} \frac{F(z)}{G(z)} > 0, \operatorname{Re} \frac{1-z}{z} G(z) > 0 \right\}, \\ F_2 &= F(1, 1, \infty, 1) = \left\{ F(z) : \left| \frac{F(z)}{G(z)} - 1 \right| < 1, \operatorname{Re} \left(\frac{1-z}{z} G(z) \right) > 0 \right\}, \\ F_3 &= F(\infty, 1, \infty, 0) = \left\{ F(z) : \operatorname{Re} \left(\frac{1-z}{z} F(z) \right) > 0 \right\}. \end{aligned}$$

Поэтому при соответствующем выборе параметров a, γ, b, δ из следствия 5 получаются радиусы звездообразности порядка α классов F_1, F_2, F_3 из [15] и в дополнение к результатам из [15] – теоремы искажения классов F_1, F_2, F_3 .

Следствие 6. Теоремы искажения и радиусы звездообразности порядка α классов F_1, F_2, F_3 имеют следующий вид:

$$1) \text{ в классе } F_1 \quad \frac{r(1-r)^2}{(1+r)^3} \leq |F(z)| \leq \frac{r(1+r)^2}{(1-r)^3}, \quad r(F_1, \alpha) = \frac{2(1-\alpha)}{5 + \sqrt{25 - 4\alpha(1-\alpha)}};$$

$$2) \text{ в классе } F_2 \quad \frac{r(1-r)^2}{(1+r)^2} \leq |F(z)| \leq \frac{r(1+r)^2}{(1-r)^2}, \quad r(F_2, \alpha) = \frac{1-\alpha}{2 + \sqrt{4 + (1-\alpha)^2}};$$

$$3) \text{ в классе } F_3 \quad \frac{r(1-r)}{(1+r)^2} \leq |F(z)| \leq \frac{r(1+r)}{(1-r)^2}, \quad r(F_3, \alpha) = \frac{2(1-\alpha)}{3 + \sqrt{9 - 4\alpha(1-\alpha)}}.$$

Заключение

В статье получило дальнейшее развитие исследование классов аналитических нормированных функций $f(z)$ из N , введенных в [5] и [9]. Для этого вводится и исследуется обобщен-

ный класс дважды почти выпуклых функций, который при определенных значениях параметров приводит как к классам из [5], [9] и других статей, а также дает ряд новых интересных подклассов.

Установлена связь введенного класса функций с некоторым подклассом класса дважды почти звездообразных функций, на основе которой исследованы свойства этого класса. В том числе получены обобщения некоторых ранее известных результатов.

Доказанные в статье утверждения являются новыми результатами в теории экстремальных задач для подклассов дважды почти выпуклых и дважды почти звездообразных функций.

ЛИТЕРАТУРА

- 1 Kaplan W. Close-to-convex schlicht functions // Michigan Math. J. – 1952. – Vol. 1 – No. 2 – P. 169–185. <https://doi.org/10.1307/mmj/1028988895>.
- 2 Renyi A. Some remarks on univalent functions // Ann. Univ. Mariae Curie-Sklodowska. – 1959. – Sec. A.3. – № 2. – P. 111–121. <http://sci-gems.math.bas.bg:8080/jspui/bitstream/10525/2878/1/1959-111-121.pdf>.
- 3 Reade M.O. The coefficients of close-to-convex functions // Duke Math. J. – 1956. – Vol. 23. – No. 3. – P. 459–462. <https://doi.org/10.1215/S0012-7094-56-02342-0>
- 4 Майер Ф.Ф., Тастанов М.Г., Утемисова А.А., Ысмағұл Р.С. Точные оценки регулярных функций и радиусы выпуклости и звездообразности некоторых классов звездообразных и почти звездообразных функций // Вестник Казахстанско-Британского технического университета. – 2024. – 21(2). – С. 127–138. <https://doi.org/10.55452/1998-6688-2024-21-2-127-138>.
- 5 Abdel-Gawad H.R., Thomas D.K. A subclass of close-to-convex functions // Publications De L’Institut Mathématique, Nouvelle série tome. – 1991. – Vol. 49. No. 63. – P. 61–66.
- 6 Selvaraj C. A subclass of close-to-convex functions // South East Asian Bull. of Math. – 2004. – Vol. 28 – P. 113–123.
- 7 Mehork B.S., Singh G. A subclass of close-to-convex functions // Int. J. of Math. Analysis. – 2010, Vol. 4 – No. 27 – P. 1319–1327. <http://www.m-hikari.com/ijma/ijma-2010/ijma-25-28-2010/singhIJMA25-28-2010.pdf>.
- 8 Peng Z.G. On a subclass of close to convex functions // Acta Mathematica Scientia, – 2010. – Vol. 30. – No. 5. – P. 1449–1456.
- 9 Stelin S., Selvaraj C. On a generalized class of close-to-convex functions of order α . Intern // J. of Pure and Applied Math. – 2016. – Vol. 109. – No. 5. – P. 141–149.
- 10 Reade M.O. On close-to-convex univalent functions // Michigan Math. J. – 1955. – Vol. 3 – P. 59–62.
- 11 MacGregor T.H. The radius of univalence of certain analytic functions // Proc. Amer. Math. Soc. – 1963. – Vol. 14. – P. 514–520.
- 12 MacGregor T.H. The radius of univalence of certain analytic functions, II // Proc. Am. Math. Soc. – 1963. – Vol. 14. – P. 521–524. <http://dx.doi.org/10.1090/s0002-9939-1963-0148892-5>.
- 13 Kohr G., Graham I. Geometric function theory in one and higher dimensions // New York: Pure and Applied Mathematics (M. Dekker), Inc. 56. – 2003. – 528 p. <https://doi.org/10.1201/9780203911624>.
- 14 Strohäcker E. Beiträge zur theorie der schlichten funktionen // Mathematische Zeitschrift. – 1933. – Vol. 37 – P. 356–380.
- 15 Sebastianc A., Ravichandran V. Radius of starlikeness of certain analytic functions // Math. Slovaca. – 2021. – Vol. 71. – No. 1. – P. 83–104. <https://doi.org/10.48550/arXiv.2001.06999>.
- 16 Khatter K., Lee S.K., Ravichandran V. Radius of starlikeness for classes of analytic functions // arXiv preprint arXiv:2006.11744 – 2020. <https://doi.org/10.48550/arXiv.2006.11744>.
- 17 El-Faqeer A.S.A., Mohd M.H., Ravichandran V., Supramaniam S. Starlikeness of certain analytic functions // arXiv preprint arXiv:2006.11734 – 2020. <https://doi.org/10.48550/arXiv.2006.11734>.
- 18 Sharma M., Jain N.K., Kumar S. Constrained radius estimates of certain analytic functions // arXiv:2305.16210v1 [math.CV] – 2023.

19 Kanaga R., Ravichandran V. Starlikeness for certain close-to-star functions // Hacet. J. Math. Stat. – 2021. – Vol. 50. – No. 2. – P. 414–432. <https://doi.org/10.15672/hujms.702703>.

20 Ali R.M., Jain N.K., Ravichandran V. On the radius constants for classes of analytic functions // arXiv:1207.4529v1 [math.CV]. – 2012. <https://doi.org/10.48550/arXiv.1207.4529>.

REFERENCES

1 Kaplan, W. Close-to-convex schlicht functions. Michigan Math. J., 1 (2), 169–185 (1952) . <https://10.1307/mmj/1028988895>.

2 Renyi, A. Some remarks on univalent functions, Ann. Univ. Mariae Curie-Sklodowska, Sec. A.3, no. 2, pp. 111–121 (1959). <http://sci-gems.math.bas.bg:8080/jspui/bitstream/10525/2878/1/1959-111-121.pdf>.

3 Reade, M.O. The coefficients of close-to-convex functions, Duke Math. J., vol. 23, no. 3, pp. 459–462 (1956). <https://doi.org/10.1215/S0012-7094-56-02342-0>.

4 Maiyer, F.F., Tastanov, M.G., Utemissova, A.A. and Ysmagul, R.S. Exact estimates of regular functions and radii of convexity and starlikeness of some classes of starlike and close-to-starlike functions, Herald of the Kazakh-British technical university, 21(2), 127–138 (2024). <https://doi.org/10.55452/1998-6688-2024-21-2-127-138>.

5 Abdel-Gawad, H.R., Thomas, D.K. A subclass of close-to-convex functions, Publications De L'Institut Mathématique, Nouvelle série tome, 49 (63), 61–66 (1991).

6 Selvaraj, C. A subclass of close-to-convex functions, South East Asian Bull. of Math., 28, 113–123 (2004) .

7 Mehork, B.S., Singh, G. A subclass of close-to-convex functions, Int. J. of Math. Analysis, 4 (27), 1319–1327 (2010) . <http://www.m-hikari.com/ijma/ijma-2010/ijma-25-28-2010/singhIJMA25-28-2010.pdf>.

8 Peng, Z.G. On a subclass of close to convex functions, Acta Mathematica Scientia, 30(5), 1449–1456 (2010).

9 Stelin, S., Selvaraj, C. On a generalized class of close-to-convex functions of order α , Intern. J. of Pure and Applied Math., 109 (5), 141–149 (2016).

10 Reade, M.O. On close-to-convex univalent functions, Michigan Math. J., 3, 59–62 (1955).

11 MacGregor, T.H. The radius of univalence of certain analytic functions, Proc. Amer. Math. Soc., 14, 514–520 (1963).

12 MacGregor, T.H. The radius of univalence of certain analytic functions, II. Proc. Am. Math. Soc., 14, 521–524 (1963). <http://dx.doi.org/10.1090/s0002-9939-1963-0148892-5>.

13 Kohr, G., Graham, I. Geometric function theory in one and higher dimensions, New York: Pure and Applied Mathematics (M. Dekker), Inc. 56, 528 p. (2003). <https://doi.org/10.1201/9780203911624>.

14 Strohhäcker, E. Beiträge zur theorie der schlichten funktionen, Mathematische Zeitschrift, 37, 356–380 (1933).

15 Sebastianc, A., Ravichandran, V. Radius of starlikeness of certain analytic functions, Math. Slovaca, 71 (1), 83–104 (2021). <https://doi.org/10.1515/ms-2017-0454>.

16 Khatter, K., Lee, S. K. and Ravichandran, V. Radius of starlikeness for classes of analytic functions, arXiv preprint arXiv:2006.11744 (2020). <https://doi.org/10.48550/arXiv.2006.11744>.

17 El-Faqeer, A.S.A., Mohd, M.H., Ravichandran, V. and Supramaniam S. Starlikeness of certain analytic functions, arXiv preprint arXiv:2006.11734 (2020). <https://doi.org/10.48550/arXiv.2006.11734>.

18 Sharma, M., Jain, N.K. and Kumar, S. Constrained radius estimates of certain analytic functions, arXiv:2305.16210v1 (2023) .

19 Kanaga, R., Ravichandran, V. Starlikeness for certain close-to-star functions, Hacet. J. Math. Stat., 50 (2), 414–432 (2021). <https://doi.org/10.15672/hujms.702703>.

20 Ali, R.M., Jain, N.K. and Ravichandran, V. On the radius constants for classes of analytic functions, arXiv:1207.4529v1 (2012). <https://doi.org/10.48550/arXiv.1207.4529>

¹Майер Ф.Ф.,

ф.-м.ғ.к., профессор, ORCID ID: 0000-0002-2278-2723,
e-mail: maiyer@mail.ru

¹Тастанов М.Г.,

ф.-м.ғ.к., профессор, ORCID ID: 0000-0003-1926-8958,
e-mail: tastao@mail.ru

^{1*}Утемисова А.А.,

пед.ғ.к., доцент, ORCID ID: 0000-0001-5143-0260,
*e-mail: anar_utemisova@mail.ru

¹Калаков Б.А.,

ф.-м.ғ.к., ORCID ID: 0009-0003-6007-02546
e-mail: kalakov1968@gmail.com

¹Ахмет Байтұрсынұлы атындағы Қостанай өңірлік университеті,
Қостанай қ., Қазақстан

ДӨНЕС ДЕРЛІК ЖӘНЕ ЕКІ ЕСЕ ДӨНЕС ДЕРЛІК ФУНКЦИЯЛАР КЛАСЫНЫҢ КЕЙБІР ҚОСАЛҚЫ КЛАСТАРЫ ТУРАЛЫ

Аңдатпа

Дөнес дерлік функциялардың K класы функциялық $Re \left(\frac{zf'(z)}{g(z)} \right)$ оңдылығымен анықталатыны белгілі, мұндағы $g(z)$ – жұлдыз тәрізді функция. Жұлдыз тәрізді функцияны $g(z)$ дөнес функциямен алмастырсақ, онда K класының белгілі бір қосалқы класы C' алынады. Бұл мақалада біз мәндер жиыны $\frac{zf'(z)}{g(z)}$ белгілі бір жағдайда жарты жазықтықпен сәйкес келетін ерекше түрдегі аймақта болған жағдайға белгілі бір жағдайда жарты жазықтықпен сәйкес келетін типтегі облыста орналасқан жағдайға арналған C' класының жалпылауын енгіземіз. Бұл жалпылау, сонымен қатар, нормаланған қос дерлік дөнес функциялар класының белгілі бір ішкі сыныбына дейін C' класын кеңейтуге байланысты. Арнайы типтегі облыстардың әртүрлілігі мен екі есе дерлік дөнес функцияларға көшу бізге жаңа нәтижелер алуға, сондай-ақ бұрын белгілі нәтижелерді жалпылауға мүмкіндік береді. Мақаланың негізгі зерттеу бағыттары – бұрмалану туралы теоремаларды дәлелдеу, қарастырылған функциялар класының дөңестік радиустарын анықтау және алынған нәтижелердің дәлдігін негіздеу болып табылады. Сонымен қатар, енгізілген функциялар класы мен екі есе дерлік жұлдыз тәрізді функциялардың жаңа класы арасындағы байланыс орнатылды. Бұл класс пен оның ішкі сыныптары үшін жұлдыз тәрізділіктің өсуі мен радиусы туралы, сондай-ақ бұрын белгілі нәтижелерді жалпылауға қатысты жаңа теоремалар алынды.

Тірек сөздер: жұлдыз тәрізді функциялар, дөнес функциялар, бұрмалау теоремалары, дөнес радиустар, өсу теоремалары, жұлдыз тәрізді радиустар.

¹**Maiyer F.F.,**

Cand. Phys.-Math. Sc., Professor, ORCID ID: 0000-0002-2278-2723,

e-mail: maiyer@mail.ru

¹**Tastanov M.G.,**

Cand. Phys.-Math. Sc., Professor, ORCID ID: 0000-0003-1926-8958,

e-mail: tastao@mail.ru

^{1*}**Utemissova A.A.,**

Cand. Ped. Sc., Associate Professor, ORCID ID: 0000-0001-5143-0260,

*e-mail: anar_utemisova@mail.ru

¹**Kalakov B.A.,**

Cand. Phys.-Math. Sc., ORCID ID: 0009-0003-6007-0254

e-mail: kalakov1968@gmail.com

¹Akhmet Baitursynuly Kostanay Regional University,
Kostanay, Kazakhstan**ABOUT SOME SUBCLASSES OF CLASSES OF CLOSE-TO-CONVEX
AND DOUBLY CLOSE-TO-CONVEX FUNCTIONS****Abstract**

It is known that the class K close-to-convex functions is defined by the condition of positivity of the functional $Re \left(\frac{zf'(z)}{g(z)} \right)$ with a starlike function $g(z)$. Replacing the starlike function $g(z)$ with a convex one leads to a known subclass C' of the class K . In this article, we introduce a generalization of the class C' to the case when the set of values $\frac{zf'(z)}{g(z)}$ is contained in a region of a special type, which, in a particular case, can coincide with a half-plane. The generalization is also associated with the extension of the class C' to a certain subclass of the class of normalized doubly close-to-convex functions. The diversity of special cases of a domain of a special type and the transition to doubly close-to-convex functions allows us to obtain both new original results and generalizations of previously known results. The main research of this article is aimed at proving theorems on distortion, finding the radii of convexity of the considered classes of functions and justifying the accuracy of the obtained results. A connection has also been established between the introduced class of functions and a certain new class of doubly close-to-starlike functions, special cases of which have been actively studied in recent years. For this class and its subclasses, new results have also been obtained in the form of theorems on growth and radius of starlikeness and generalization of previously known results.

Keywords: starlike functions, close-to-starlike functions, close-to-convex functions, distortion theorems, radii of convexity, growth theorems, radii of starlikeness.

Дата поступления статьи в редакцию: 23.06.2025

УДК 517.51
МРНТИ 27.39.19

<https://doi.org/10.55452/1998-6688-2025-22-4-295-305>

¹Әбек А.Н.,

PhD, ORCID ID: 0009-0004-7158-3597,

*e-mail: azhar.abekova@gmail.com

²Гогатишвили А.,

PhD, ORCID ID: 0000-0003-3459-0355,

e-mail: gogatish@math.cas.cz

³*Бокаев Н.А.,

доктор физ.-матем. наук, профессор, ORCID ID: 0000-0002-7071-1882,

e-mail: bokayev2011@yandex.kz

⁴Унвер Т.,

PhD, ORCID ID: 0000-0003-0414-8400,

e-mail: tugceunver@kku.edu.tr

^{1,3}Евразийский национальный университет им. Л.Н. Гумилева, г. Астана, Казахстан

²Институт математики Академии наук Чехии, г. Прага, Республика Чехия

⁴Университет Кириккале, г. Кириккале, Турция

ВЕСОВЫЕ НЕРАВЕНСТВА ДЛЯ СУПЕРПОЗИЦИИ ТРЕХ ОПЕРАТОРОВ

Аннотация

Мы изучаем трехвесовое неравенство для суперпозиции операторов Копсона, Харди и Тандори. Целью данной работы является доказательство полной характеристики ограниченности оператора, являющегося комбинацией этих трех операторов в весовых пространствах Лебега от $L_1(v)$ в $L_q(w)$. Основное внимание уделяется определению необходимых и достаточных условий, при которых это неравенство выполняется для всех неотрицательных измеримых функций на положительной вещественной оси. Существенно используется понятие фундаментальной функции борелевской меры относительно возрастающей функции. Поскольку оператор Тандори не является линейным оператором, мы не можем использовать методы двойственности, применявшиеся в более ранних работах. Для решения этой проблемы мы разрабатываем новый, упрощенный метод дискретизации, позволяющий избежать сложностей ранее известных методов. Был получен явный вид наилучшей константы в неравенстве, что демонстрирует точность и оптимальность результатов. Устанавливая необходимые и достаточные условия ограниченности этих составных операторов, мы улучшаем неравенства, ранее установленные в работах Гогатишвили А., Пик Л., Опич Б. [1]. Полученные в статье результаты расширяют и дополняют существующие исследования в области теории весовых неравенств и операторного анализа в функциональных пространствах и предлагают потенциальные приложения в теории приближений, гармоническом анализе и смежных областях.

Ключевые слова: весовое неравенство, супремальный оператор, оператор Копсона, оператор Харди, оператор Тандори, наилучшая константа, суперпозиция операторов, дискретизация.

Введение

В последние годы весовые неравенства для операторов типа Харди привлекли значительное внимание из-за их широкого применения в функциональном анализе и теории операторов. Классические результаты в этой области, такие как исследованные Bennett C. и Sharpley R. [2] заложили основу для дальнейшего прогресса в анализе ограниченности операторов в весовых пространствах. Изучение составных операторов, таких как комбинация операторов Копсона, Харди и Тандори, еще больше расширило эту теорию [3–12]. Более ранние работы, такие как

работы Gogatishvili A., Pick L. [1, 13], были сосредоточены на весовых неравенствах для отдельных операторов, но отсутствовали всесторонние исследования их взаимодействия при объединении. Наше исследование устраняет этот пробел, предоставляя подробную характеристику ограниченности составного оператора, образованного операторами Харди, Копсона и Тандори в весовых пространствах Лебега. Это является естественным продолжением более ранних результатов [14–17], которые исследовали связанные семейства операторов, но не рассматривали всю сложность нелинейных членов в таких неравенствах.

Через $\mathfrak{M}^+(0, \infty)$ обозначим множество всех неотрицательных измеримых функций на $(0, \infty)$.

Пусть $1 \leq q < \infty$ и пусть u, v и w – веса (т.е. локально интегрируемые неотрицательные функции на $(0, \infty)$), φ – строго возрастающая функция на $(0, \infty)$, а $\frac{\varphi}{U}$ – убывающая на $(0, \infty)$, где

$$U(s) = \int_0^s u(t) dt.$$

Наша цель в данной работе – охарактеризовать следующее неравенство

$$\left(\int_0^\infty \left(\sup_{t < s < \infty} \frac{1}{\varphi(s)} \int_0^s \left(\int_\tau^\infty h(y) dy \right) u(\tau) d\tau \right)^q w(t) dt \right)^{\frac{1}{q}} \leq C \int_0^\infty h(s) v(s) ds \quad (1)$$

для всех $h \in \mathfrak{M}^+(0, \infty)$.

Наименьшую константу C , для которой неравенство (1) выполняется, назовем наилучшей константой.

Неравенство (1) представляет собой трехвесовое неравенство, включающее композицию выражений типа Копсона, Харди и Тандори. Внутренний интеграл нелинеен из-за структуры оператора Тандори.

Используя теорему Фубини для неотрицательных функций, имеем

$$\int_0^s u(\tau) \int_\tau^\infty h(t) dt d\tau = \int_0^s U(\tau) h(\tau) d\tau + U(s) \int_s^\infty h(\tau) d\tau.$$

Следовательно, неравенство (1) эквивалентно следующему неравенству

$$\left(\int_0^\infty \left(\sup_{t < s < \infty} \frac{1}{\varphi(s)} \left(\int_0^s U(\tau) h(\tau) d\tau + U(s) \int_s^\infty h(\tau) d\tau \right) \right)^q w(t) dt \right)^{\frac{1}{q}} \leq C \int_0^\infty h(\tau) v(\tau) d\tau. \quad (2)$$

Следующая теорема является основным результатом данной статьи и дает критерий выполнения неравенства (2).

Теорема 1. Пусть $1 \leq q < \infty$ и пусть u, v и w – веса на $(0, \infty)$. φ это U – квазивогнутая функция на $(0, \infty)$. Тогда существует постоянная $C > 0$ такая, что неравенство (2) выполняется для всех $h \in \mathfrak{M}^+(0, \infty)$ тогда и только тогда, когда выполняется следующее условие:

$$C_1 := \sup_{t \in (0, \infty)} \left(\int_0^\infty \frac{w(s) ds}{(\varphi(s) + \varphi(t))^q} \right)^{\frac{1}{q}} \left(\sup_{s \in (0, \infty)} \frac{U(t)U(s)}{(U(s) + U(t))v(s)} \right) < \infty.$$

Более того, наилучшая константа C в (2) удовлетворяет условию $C_1 \approx C$.

Всюду в статье мы всегда обозначаем через c или C положительную константу, которая не зависит от основных параметров, но может меняться от строки к строке. Однако константа с нижним индексом, такая как c_1 , не меняется в разных случаях.

Мы пишем $A \lesssim B$ (или $A \gtrsim B$), если $A \leq C_1 B$ (или $C_2 A \geq B$), для некоторой положительной константы C , не зависящей от соответствующих величин, входящих в выражения A и B , и $A \approx B$, если $A \lesssim B$ и $A \gtrsim B$. Всюду в этих заметках мы используем сокращение $LHS(*)$ (и $RHS(*)$) для левой (правой) части соотношения $(*)$.

Материалы и методы

Определение 1. Пусть a_k – последовательность положительных действительных чисел. Будем говорить, что a_k строго возрастает или строго убывает, и писать $a_k \uparrow\uparrow$ или $a_k \downarrow\downarrow$,

$$\inf_{k \in \mathbb{Z}} \frac{a_{k+1}}{a_k} > 1 \text{ или } \sup_{k \in \mathbb{Z}} \frac{a_{k+1}}{a_k} < 1,$$

соответственно.

Определение 2. Пусть φ – непрерывная строго возрастающая функция на $[0, \infty)$ такая, что $\varphi(0) = 0$ и $\lim_{t \rightarrow \infty} \varphi(t) = \infty$. Тогда будем говорить, что φ допустима.

Пусть φ – допустимая функция. Будем говорить, что функция h является φ -квазивогнутой, если h эквивалентна возрастающей функции на $[0, \infty)$, а отношение $\frac{h}{\varphi}$ эквивалентно убывающей функции на $(0, \infty)$. Будем говорить, что φ -квазивогнутая функция h невырождена, если

$$\lim_{t \rightarrow 0+} h(t) = \lim_{t \rightarrow \infty} \frac{1}{h(t)} = \lim_{t \rightarrow \infty} \frac{h(t)}{\varphi(t)} = \lim_{t \rightarrow 0+} \frac{\varphi(t)}{h(t)} = 0.$$

Семейство невырожденных φ -квазивогнутых функций обозначим через Ω_φ . Будем говорить, что φ квазивогнута, когда $\varphi \in \Omega_\varphi$ с $U(t) = t$.

Определение 3. Предположим, что φ допустима и $h \in \Omega_\varphi$. Говорят, что $\{x_k\}_{k \in \mathbb{Z}}$ – дискретизационная последовательность для h относительно φ , если

- (i) $x_0 = 1$ и $\varphi(x_k) \uparrow\uparrow$;
- (ii) $h(x_k) \uparrow\uparrow$ и $\frac{h(x_k)}{\varphi(x_k)} \downarrow\downarrow$;
- (iii) существует разложение $\mathbb{Z} = \mathbb{Z}_1 \cup \mathbb{Z}_2$ такое, что $\mathbb{Z}_1 \cap \mathbb{Z}_2 = \emptyset$ и для любого $t \in [x_k, x_{k+1}]$

$$\begin{aligned} h(x_k) &\approx h(t) \text{ if } k \in \mathbb{Z}_1 \\ \frac{h(x_k)}{\varphi(x_k)} &\approx \frac{h(t)}{\varphi(t)} \text{ if } k \in \mathbb{Z}_2. \end{aligned}$$

Определение 4. Пусть φ – допустимая функция, а ν – неотрицательная борелевская мера на $[0, \infty)$. Будем говорить, что функция h , определяемая формулой

$$h(t) = \varphi(t) \int_{[0, \infty)} \frac{d\nu(s)}{\varphi(s) + \varphi(t)}, t \in (0, \infty)$$

является фундаментальной функцией меры ν относительно φ . Будем также говорить, что ν является мерой представления h относительно φ .

Будем говорить, что ν невырождена, если для каждого $t \in (0, \infty)$ выполняются следующие условия:

$$\int_{[0, \infty)} \frac{d\nu(s)}{\varphi(s) + \varphi(t)} < \infty, t \in (0, \infty) \quad \text{и} \quad \int_{[0, 1]} \frac{d\nu(s)}{\varphi(s)} = \int_{[1, \infty)} d\nu(s) = \infty.$$

Следующая лемма: устойчивость супремума и сумм по дискретизациям имеет решающее значение для согласованных эквивалентностей нормы.

Лемма 1. [13, Лемма 3.7] Пусть $q \in (0, \infty)$. Предположим, что φ – допустимая функция, ν – невырожденная положительная борелевская мера на $[0, \infty)$, h – фундаментальная функция ν относительно U^q , и $f \in \mathfrak{M}^+(0, \infty)$. Если $\{x_k\}$ – дискретизационная последовательность для h относительно φ^q , то

$$\int_{[0, \infty)} \left(\operatorname{esssup}_{y \in (0, \infty)} \frac{|f(y)|}{\varphi(x) + \varphi(y)} \right)^q d\nu(x) \approx \sum_{k \in \mathbb{Z}} \left(\operatorname{esssup}_{y \in (0, \infty)} \frac{|f(y)|}{\varphi(x_k) + \varphi(y)} \right)^q \varphi(x_k) \\ \approx \sum_{k \in \mathbb{Z}} \operatorname{esssup}_{x_k \leq y < x_{k+1}} |f(y)|^q \varphi^{-q}(y) h(y).$$

Лемма 2. [13, лемма 3.9] Пусть u – допустимая функция и $\varphi \in \Omega_U$. Пусть $\{x_k\}$ – дискретизационная последовательность для φ относительно U . Тогда

$$\sup_{x \in (0, \infty)} \varphi(x) \sup_{0 < y < \infty} \frac{|f(y)|}{U(x) + U(y)} \approx \sup_{k \in \mathbb{Z}} \varphi(x_k) \sup_{0 < y < \infty} \frac{|f(y)|}{U(x_k) + U(y)} \approx \\ \approx \sup_{k \in \mathbb{Z}} \varphi(x_k) U(x_k)^{-1} \sup_{x_{k-1} < y < x_k} |f(y)| + \sup_{k \in \mathbb{Z}} \varphi(x_k) \sup_{x_k < y < x_{k+1}} |f(y)| U(y)^{-1} \approx \\ \approx \sup_{k \in \mathbb{Z}_1} \varphi(x_k) \sup_{x_k \leq y < x_{k+1}} |f(y)| U(y)^{-1} + \sup_{k \in \mathbb{Z}_2} \varphi(x_k) U(x_k)^{-1} \sup_{x_k < y < x_{k+1}} |f(y)| \\ \approx \sup_{k \in \mathbb{Z}} \sup_{x_k \leq y < x_{k+1}} |f(y)| U(y)^{-1} \varphi(y).$$

Лемма 3. Пусть $\{x_k\}_{k \in \mathbb{Z}}, \{\tau_k\}_{k \in \mathbb{Z}}$ – произвольные последовательности неотрицательных чисел, $\tau_k \downarrow$, $h \in \mathfrak{M}^+(0, \infty)$. Тогда для любого $q > 0$

$$\sum_{k \in \mathbb{Z}} \left(\int_0^{x_k} h \right)^q \tau_k \approx \sum_{k \in \mathbb{Z}} \left(\int_{x_{k-1}}^{x_k} h \right)^q \tau_k, \\ \sup_{k \in \mathbb{Z}} \left(\int_0^{x_k} h \right)^q \tau_k \approx \sup_{k \in \mathbb{Z}} \left(\int_{x_{k-1}}^{x_k} h \right)^q \tau_k, \\ \sum_{k \in \mathbb{Z}} \left(\int_{x_k}^{\infty} h \right)^q \tau_k^{-1} \approx \sum_{k \in \mathbb{Z}} \left(\int_{x_k}^{x_{k+1}} h \right)^q \tau_k^{-1}, \\ \sup_{k \in \mathbb{Z}} \left(\int_{x_k}^{\infty} h \right)^q \tau_k^{-1} \approx \sup_{k \in \mathbb{Z}} \left(\int_{x_k}^{x_{k+1}} h \right)^q \tau_k^{-1}.$$

Лемма 4. Пусть $\{x_k\}_{k \in \mathbb{Z}}$ и $\{\tau_k\}_{k \in \mathbb{Z}}$ – произвольные последовательности неотрицательных чисел. $\tau_k \downarrow$ и $h \in \mathfrak{M}^+(0, \infty)$. Тогда

$$\sup_{k \in \mathbb{Z}} \left(\operatorname{esssup}_{0 \leq t \leq x_k} h(t) \right) \tau_k \approx \sup_{k \in \mathbb{Z}} \left(\operatorname{esssup}_{x_{k-1} \leq t \leq x_k} h(t) \right) \tau_k, \\ \sup_{k \in \mathbb{Z}} \left(\operatorname{esssup}_{x_k \leq t \leq \infty} h(t) \right) \tau_k^{-1} \approx \sup_{k \in \mathbb{Z}} \left(\operatorname{esssup}_{x_k \leq t \leq x_{k+1}} h(t) \right) \tau_k^{-1}.$$

Результаты и обсуждение

Мы начинаем этот раздел с замечания, что неравенство (2) эквивалентно дискретному неравенству, и представляем его характеристику в дискретной форме, что само по себе заслуживает внимания.

Пусть $q \in (0, \infty)$ и

$$G(t) = \left(\int_0^t w(s) ds + \varphi(t)^q \int_t^\infty \varphi^{-q}(s) w(s) ds \right)^{\frac{1}{q}}. \quad (3)$$

Лемма 5. Пусть $q \in (0, \infty)$, и пусть $h \in \mathfrak{M}^+(0, \infty)$, w – вес на $(0, \infty)$. U и φ допустимые функции на $(0, \infty)$ и $\frac{\varphi}{U}$ убывает. G определена в (3) и пусть $\{x_k\}$ дискретизационная последовательность G относительно φ . Тогда из левой части (2) следует, что

$$\begin{aligned} & \int_0^\infty \left(\sup_{t < s < \infty} \frac{1}{\varphi(s)} \left(\int_0^s U(\tau) h(\tau) d\tau + U(s) \int_s^\infty h(\tau) d\tau \right) \right)^q w(t) dt \approx \\ & \approx \sum_{k \in \mathbb{Z}} G(x_k) \left[\sup_{x_k < t < \infty} \frac{1}{\varphi(t)} \left(\int_0^t U(\tau) h(\tau) d\tau + U(t) \int_t^\infty h(\tau) d\tau \right) \right]^q \\ & \approx \sum_{k \in \mathbb{Z}} \left[\sup_{x_k < t < x_{k+1}} \frac{G(t)}{\varphi(t)} \left(\int_{x_k}^t U(\tau) h(\tau) d\tau + U(t) \int_t^{x_{k+1}} h(\tau) d\tau \right) \right]^q. \quad (4) \end{aligned}$$

Доказательство. Обозначим

$$\begin{aligned} H(t) &:= \varphi(t) \sup_{t < s < \infty} \frac{1}{\varphi(s)} \left(\int_0^s U(\tau) h(\tau) d\tau + U(s) \int_s^\infty h(\tau) d\tau \right) \\ &= \sup_{t < s < \infty} \frac{\varphi(t)}{\varphi(s)} \left(\int_0^s U(\tau) h(\tau) d\tau + U(s) \int_s^\infty h(\tau) d\tau \right). \end{aligned}$$

Мы имеем

$$H(t) = \sup_{0 < s < \infty} \min \left(1, \frac{\varphi(t)}{\varphi(s)} \right) \left[\int_0^s U(\tau) h(\tau) d\tau + U(s) \int_s^\infty h(\tau) d\tau \right].$$

Следовательно, $H \uparrow$ и $\frac{H}{\varphi} \downarrow$, то есть $H \in \Omega_\varphi$. Используя элементарные неравенства

$$\frac{\varphi(t)}{\varphi(s) + \varphi(t)} < \min \left(1, \frac{\varphi(t)}{\varphi(s)} \right) < \frac{2\varphi(t)}{\varphi(s) + \varphi(t)},$$

получаем

$$H(t) \approx \sup_{0 < s < \infty} \frac{\varphi(t)}{\varphi(s) + \varphi(t)} \left[\int_0^s U(\tau) h(\tau) d\tau + U(s) \int_s^\infty h(\tau) d\tau \right] = \sup_{0 < s < \infty} \frac{\varphi(t)}{\varphi(s) + \varphi(t)} f(s).$$

Используя Лемму 1 для функции $f(s) = \int_0^s U(\tau) h(\tau) d\tau + U(s) \int_s^\infty h(\tau) d\tau$, имеем:

$$\begin{aligned} \int_0^\infty H(t)^q \varphi^q(t) w(t) dt &\approx \int_0^\infty \left[\sup_{0 < s < \infty} \frac{\int_0^s U(\tau) h(\tau) d\tau + U(s) \int_s^\infty h(\tau) d\tau}{\varphi(s) + \varphi(t)} \right]^q \varphi^q(t) w(t) dt \\ &\approx \sum_{k \in \mathbb{Z}} G(x_k) \left[\sup_{x_k < t < \infty} \frac{1}{\varphi(t)} \left(\int_0^t U(\tau) h(\tau) d\tau + U(t) \int_t^\infty h(\tau) d\tau \right) \right]^q \\ &\approx \sum_{k \in \mathbb{Z}} \left[\sup_{x_k < t < x_{k+1}} \left(\int_0^t U(\tau) h(\tau) d\tau + U(t) \int_t^\infty h(\tau) d\tau \right) \frac{G(t)}{\varphi(t)} \right]^q \\ &\approx \sum_{k \in \mathbb{Z}} \left[\sup_{x_k < t < x_{k+1}} \left(\int_{x_k}^t U(\tau) h(\tau) d\tau + U(t) \int_t^{x_{k+1}} h(\tau) d\tau \right) \frac{G(t)}{\varphi(t)} \right]^q \\ &\quad + \sum_{k \in \mathbb{Z}} \left[\sup_{x_k < t < x_{k+1}} \frac{G(t)}{\varphi(t)} \int_0^{x_k} U(\tau) h(\tau) d\tau \right]^q \end{aligned}$$

$$+ \sum_{k \in \mathbb{Z}} \left[\sup_{x_k < t < x_{k+1}} \frac{G(t)U(t)}{\varphi(t)} \int_{x_{k+1}}^{\infty} h(\tau) d\tau \right]^q =: I_1 + I_2 + I_3.$$

Поскольку $\frac{G(x_k)}{\varphi(x_k)} \downarrow\downarrow$, используя Лемму 3, получаем

$$\begin{aligned} I_2 &\approx \sum_{k \in \mathbb{Z}} \left[\frac{G(x_k)}{\varphi(x_k)} \int_0^{x_k} U(\tau) h(\tau) d\tau \right]^q \approx \sum_{k \in \mathbb{Z}} \left[\frac{G(x_k)}{\varphi(x_k)} \int_{x_{k-1}}^{x_k} U(\tau) h(\tau) d\tau \right]^q \leq \\ &\leq \sum_{k \in \mathbb{Z}} \left[\sup_{x_{k-1} < t < x_k} \frac{G(t)}{\varphi(t)} \int_{x_{k-1}}^t U(\tau) h(\tau) d\tau \right]^q \leq I_1. \end{aligned}$$

Так как $\frac{U(x_{k+1})}{\varphi(x_{k+1})} \uparrow$ и $G(x_{k+1}) \uparrow\uparrow$, то $\frac{G(x_{k+1})U(x_{k+1})}{\varphi(x_{k+1})} \uparrow\uparrow$, и используя Лемму 3, получаем

$$\begin{aligned} I_3 &= \sum_{k \in \mathbb{Z}} \left[\frac{G(x_{k+1})U(x_{k+1})}{\varphi(x_{k+1})} \int_{x_{k+1}}^{\infty} h(\tau) d\tau \right]^q \approx \sum_{k \in \mathbb{Z}} \left[\frac{G(x_{k+1})U(x_{k+1})}{\varphi(x_{k+1})} \int_{x_{k+1}}^{x_{k+2}} h(\tau) d\tau \right]^q \\ &\leq \sum_{k \in \mathbb{Z}} \left[\sup_{x_{k+1} < t < x_{k+2}} \frac{G(t)U(t)}{\varphi(t)} \int_t^{x_{k+2}} h(\tau) d\tau \right]^q \leq I_1. \end{aligned}$$

Таким образом, все части управляются I_1 , и мы получаем:

$$I_1 + I_2 + I_3 \approx I_1$$

что доказывает эквивалентность (8). Лемма 8 доказана.

Предложение 1. Пусть $1 \leq q < \infty$, и пусть u, v и w – веса на $(0, \infty)$, U и φ допустимые функции на $(0, \infty)$ и $\frac{\varphi}{U}$ убывает. G определена в (3), и пусть $\{x_k\}$ дискретизационная последовательность G относительно φ . Тогда неравенство (2) выполняется для всех $h \in \mathfrak{M}^+(0, \infty)$, тогда и только тогда, когда существует положительная константа $\mathcal{C}$ такая, что

$$\begin{aligned} \left[\sum_{k \in \mathbb{Z}} \sup_{x_k < t < x_{k+1}} \frac{G(t)}{\varphi(t)} \left(\int_{x_k}^t U(\tau) h(\tau) d\tau + U(t) \int_t^{x_{k+1}} h(\tau) d\tau \right)^q \right]^{\frac{1}{q}} &\leq \\ &\leq \mathcal{C} \sum_{k \in \mathbb{Z}} \int_{x_k}^{x_{k+1}} h(\tau) v(\tau) d\tau \end{aligned} \quad (5)$$

выполняется для всех $h \in \mathfrak{M}^+(0, \infty)$. Более того, наилучшие константы в неравенствах (2) и (5) удовлетворяют условию $\mathcal{C} \approx \mathcal{C}$.

Доказательство. Пусть $\{x_k\}_{k \in \mathbb{N}}$ дискретизационная последовательность функции G . Применяя (4), получаем, что

$$\text{LHS}(2) \approx \left(\sum_{k \in \mathbb{Z}} \left[\sup_{x_k < t < x_{k+1}} \frac{G(t)}{\varphi(t)} \left(\int_{x_k}^t U(\tau) h(\tau) d\tau + U(t) \int_t^{x_{k+1}} h(\tau) d\tau \right)^q \right]^{\frac{1}{q}} \right),$$

и

$$\text{RHS}(2) = \int_0^{\infty} h(\tau) v(\tau) d\tau = \sum_{k \in \mathbb{Z}} \int_{x_k}^{x_{k+1}} h(\tau) v(\tau) d\tau.$$

Тогда ясно, что существует положительная константа $\mathcal{C}$, такая, что неравенство (2) выполняется для всех $h \in \mathfrak{M}^+(0, \infty)$, тогда и только тогда, когда существует положительная константа $\mathcal{C}$, такая, что (5) выполняется для всех $h \in \mathfrak{M}^+(0, \infty)$. Более того, $\mathcal{C} \approx \mathcal{C}$. Предложение 1 доказано.

Для исследования неравенства (5) нам понадобятся следующие обозначения: сначала обозначим через $M(x_k, x_{k+1})$ наилучшую константу весовых итерированных неравенств Копсона и Харди, то есть

$$M(x_k, x_{k+1}) = \sup_{h \geq 0} \frac{\sup_{x_k < t < x_{k+1}} \frac{G(t)}{\varphi(t)} \left(\int_{x_k}^t U(\tau) h(\tau) d\tau + U(t) \int_t^{x_{k+1}} h(\tau) d\tau \right)}{\int_{x_k}^{x_{k+1}} h(\tau) v(\tau) d\tau}. \quad (6)$$

Используя характеристики весовых итерационных неравенств Копсона и Харди (см., [17, теорема 5.3.1]), имеем

$$M(x_k, x_{k+1}) \approx \sup_{x_k < t < x_{k+1}} \frac{G(t)}{\varphi(t)} \left(\sup_{x_k < s < t} \frac{u(s)}{v(s)} + U(t) \sup_{t < s < x_{k+1}} \frac{1}{v(s)} \right). \quad (7)$$

Предложение 2. Пусть $1 \leq q < \infty$, и пусть u, v и w – веса на $(0, \infty)$, U и φ допустимые функции на $(0, \infty)$ и $\frac{\varphi}{U}$ убывает. G определена в (3), а $\{x_k\}$ – дискретизационная последовательность G относительно φ . $M(x_k, x_{k+1})$ определена в (6), Тогда неравенство (5) выполняется для всех $h \in \mathfrak{M}^+(0, \infty)$, тогда и только тогда, если существуют положительные константы $\mathcal{C}'$, такие, что

$$\left(\sum_{k \in \mathbb{Z}} a_k^q M(x_k, x_{k+1})^q \right)^{\frac{1}{q}} \leq \mathcal{C}' \sum_{k \in \mathbb{Z}} a_k, \quad (8)$$

выполняется для всех последовательностей неотрицательных чисел $\{a_k\}_{k \in \mathbb{Z}}$. Более того, наилучшие константы в неравенствах (5) и (8) удовлетворяют условию $\mathcal{C} \approx \mathcal{C}'$.

Доказательство. Предположим, что (5) выполняется для всех $h \in \mathfrak{M}^+(0, \infty)$. По определению $M(x_k, x_{k+1})$ существует функция $h_k \geq 0$, такая, что

$$\text{supp } h_k \subset (x_k, x_{k+1}), \quad \int_{x_k}^{x_{k+1}} h_k(\tau) v(\tau) d\tau = 1$$

и

$$\sup_{x_k < t < x_{k+1}} \frac{G(t)}{\varphi(t)} \left(\int_{x_k}^t U(\tau) h_k(\tau) d\tau + U(t) \int_t^{x_{k+1}} h_k(\tau) d\tau \right) \geq \frac{1}{2} M(x_k, x_{k+1}).$$

Если подставим функцию

$$h(s) = \sum_{k \in \mathbb{Z}} a_k h_k(s)$$

в (5), где $\{a_k\}_{k \in \mathbb{Z}}$ любая последовательность неотрицательных чисел, то из (5) следует (8). Более того, $\mathcal{C}' \lesssim \mathcal{C}$. Обратно, сначала заметим, что для каждого $h \in \mathfrak{M}^+(0, \infty)$

$$M(x_k, x_{k+1}) \geq \frac{\sup_{x_k < t < x_{k+1}} \frac{G(t)}{\varphi(t)} \left(\int_{x_k}^t U(\tau) h(\tau) d\tau + U(t) \int_t^{x_{k+1}} h(\tau) d\tau \right)}{\int_{x_k}^{x_{k+1}} h(\tau) v(\tau) d\tau}.$$

Затем (5) следует при подставлении

$$a_k = \int_{x_k}^{x_{k+1}} h(\tau) v(\tau) d\tau$$

в (8) и $\mathcal{C} \leq \mathcal{C}'$. Далее, мы имеем $\mathcal{C} \approx \mathcal{C}'$. Предложение 2 доказано.

Лемма 6. [13, предл. 4.1] Пусть $\{v_k\}_{k \in \mathbb{Z}}$ последовательность положительных действительных чисел и $1 \leq q < \infty$. Предположим, что неравенство

$$\left(\sum_{k \in \mathbb{Z}} a_k^q v_k^q \right)^{1/q} \lesssim B \sum_{k \in \mathbb{Z}} a_k \quad (9)$$

выполняется для любой последовательности положительных действительных чисел $\{a_k\}$ тогда и только тогда, когда

$$B_1 := \sup_{k \in \mathbb{Z}} v_k < \infty.$$

Более того, наилучшая константа в неравенстве (9) удовлетворяет условию $B \approx B_1$.

Следующая теорема 2 дискретизирует непрерывную задачу в последовательность локальных весовых неравенств на интервалах $[x_k, x_{k+1}]$.

Теорема 2. Пусть $1 \leq q < \infty$ и пусть u, v и w – веса на $(0, \infty)$, φ – U -квазивогнутая функция на $(0, \infty)$. Тогда существует константа $C > 0$ такая, что неравенство (2) выполняется для всех $h \in \mathfrak{M}^+(0, \infty)$, тогда и только тогда, когда выполняется следующее условие:

$$A_1 := \sup_{k \in \mathbb{Z}} \sup_{x_k \leq t \leq x_{k+1}} \frac{G(t)}{\varphi(t)} \left(\operatorname{esssup}_{s \in (x_k, t)} \frac{U(s)}{v(s)} + U(t) \operatorname{esssup}_{s \in (t, x_{k+1})} \frac{1}{v(s)} \right) < \infty.$$

Более того, наилучшая константа в неравенстве (2) удовлетворяет условию $C \approx A_1$.

Доказательство. Используя Предложения 1 и 2, наилучшая константа в (2) удовлетворяет условию $C \approx \mathcal{C}'$, где $\mathcal{C}'$ – наилучшая константа в неравенстве (8). Следовательно, применяя лемму 6 и используя (7), получаем результат.

Доказательство теоремы 1.

Используя теорему 2, достаточно показать, что

$$A_1 \approx C_1.$$

Отметим, что для функций U и v из (7) имеем

$$\operatorname{esssup}_{s \in (0, t)} \frac{U(s)}{v(s)} + U(t) \operatorname{esssup}_{s \in (t, \infty)} \frac{1}{v(s)} \approx \operatorname{esssup}_{s \in (0, \infty)} \frac{U(t)U(s)}{(U(s) + U(t))v(s)}. \quad (10)$$

Действительно, используем следующие простые оценки

$$\frac{U(t)U(s)}{(U(s) + U(t))v(s)} \leq \min(U(s), U(t)) \leq \frac{2U(t)U(s)}{(U(s) + U(t))v(s)}.$$

Следовательно,

$$\begin{aligned} \operatorname{esssup}_{s \in (0, t)} \frac{U(s)}{v(s)} + U(t) \operatorname{esssup}_{s \in (t, \infty)} \frac{1}{v(s)} &= \operatorname{esssup}_{s \in (0, \infty)} \min(U(s), U(t)) \frac{1}{v(s)} \\ &\approx \operatorname{esssup}_{s \in (0, \infty)} \frac{U(t)U(s)}{U(s) + U(t)} \frac{1}{v(s)}. \end{aligned}$$

По лемме 2, лемме 4 и (10) имеем

$$\begin{aligned} A_1 &:= \sup_{k \in \mathbb{Z}} \sup_{x_k \leq t \leq x_{k+1}} \frac{G(t)}{\varphi(t)} \left(\operatorname{esssup}_{s \in (x_k, t)} \frac{U(s)}{v(s)} + U(t) \operatorname{esssup}_{s \in (t, x_{k+1})} \frac{1}{v(s)} \right) \\ &\approx \sup_{k \in \mathbb{Z}} \sup_{x_k \leq t \leq x_{k+1}} \frac{G(t)}{\varphi(t)} \left(\operatorname{esssup}_{s \in (0, t)} \frac{U(s)}{v(s)} + U(t) \operatorname{esssup}_{s \in (t, \infty)} \frac{1}{v(s)} \right) \\ &= \sup_{t \in (0, \infty)} \frac{G(t)}{\varphi(t)} \left(\operatorname{esssup}_{s \in (0, t)} \frac{U(s)}{v(s)} + U(t) \operatorname{esssup}_{s \in (t, \infty)} \frac{1}{v(s)} \right) \\ &\approx \sup_{t \in (0, \infty)} \left(\int_0^\infty \frac{w(s) ds}{(\varphi(s) + \varphi(t))^q} \right)^{\frac{1}{q}} \left(\operatorname{esssup}_{s \in (0, \infty)} \frac{U(t)U(s)}{(U(s) + U(t))v(s)} \right) = C_1. \end{aligned}$$

Теорема 1 доказана.

Заключение

В данной работе рассмотрено весовое неравенство для оператора, состоящего из суперпозиции операторов Харди, Копсона и Тандори на классе неотрицательных измеримых функций. Получены необходимые и достаточные условия выполнения (справедливости) весового неравенства для суперпозиции рассматриваемых операторов. Получены критерии справедливости такого неравенства при условии $1 \leq q < \infty$. Для доказательства основной теоремы использовались методы дискретизации: сведение интегральных неравенств к неравенствам для суммы, критерии дискретизации для последовательностей и дискретное эквивалентное условие для неравенства (2). Установлен точный вид наилучшей константы в соответствующем неравенстве, что подтверждает точность и оптимальность полученных результатов. При доказательстве основных утверждений существенно использован метод дискретизации, ранее рассмотренный в работе Gogatishvili A., Pick L. [1, 13]. Более того, наш подход использует методы дискретизации, технику, введенную в более ранних работах [1, 13], которая позволяет нам эффективно обрабатывать нелинейности в операторе Тандори.

Полученные результаты вносят существенный вклад в развитие теории весовых неравенств и операторного анализа в функциональных пространствах, а также демонстрируют потенциал для дальнейшего применения в теории операторов в функциональных пространствах, гармоническом анализе и смежных областях математического анализа.

Информация о финансировании

Исследование А.Н. Абек и А.Гогатишвили выполнено при финансовой поддержке гранта Министерства науки и высшего образования Республики Казахстан (проект № AP22686420).

Исследование Н.А. Бокаев, А.Н. Абек, А.Гогатишвили выполнено при финансовой поддержке гранта Министерства науки и высшего образования Республики Казахстан (проект № AP26196065).

Исследование А. Гогатишвили частично поддержано грантом проекта 23-04720S Чешского научного фонда (GA ČR), Институт математики CAS поддержан RVO:67985840, Национальным научным фондом имени Шота Руставели (SRNSF), номер гранта: FR22-17770.

ЛИТЕРАТУРА

- 1 Gogatishvili, A., Pick, L., Opic, B. Weighted inequalities for Hardy-type operators involving suprema. *Collectanea Mathematica*, 57 (3), 227–255 (2006). <https://raco.cat/index.php/CollectaneaMathematica/article/view/56609/67919>.
- 2 Bennett, C., Sharpley, R. *Interpolation of Operators*. Pure and Applied Mathematics 129, Academic Press, Boston, MA, 1988, p. 469.

- 3 Křepela, M. Integral conditions for Hardy-type operators involving suprema. *Collectanea Mathematica*, 68, 21–50 (2017). <https://doi.org/10.1007/s13348-016-0170-6>.
- 4 Gogatishvili, A., Mustafayev, R.C. Iterated Hardy-type inequalities involving suprema. *Mathematical Inequalities & Applications*, 20, 901–927 (2017). <https://doi.org/10.7153/mia-2017-20-57>.
- 5 Astashkin, S.V., Maligranda, L. Structure of Cesaro function spaces: A survey. *Banach Center Publications*, 102, 13–40 (2014). <https://doi.org/10.4064/bc102-0-1>.
- 6 Kolyada, V.I. On Cesaro and Copson norms of nonnegative sequences. *Ukrainian Mathematical Journal*, 71 (2), 248–258 (2019). <https://doi.org/10.1007/s11253-019-01642-7>.
- 7 Stepanov, V.D., Shambilova, G.É. On weighted iterated Hardy-type operators. *Anal. Math.*, 44 (2), 273–283 (2018). <https://doi.org/10.1007/s10476-018-0211-3>.
- 8 Stepanov, V.D., Shambilova, G.É. On iterated and bilinear integral Hardy-type operators. *Math. Inequal. Appl.*, 22 (4), 1505–1533 (2019). <https://doi.org/10.7153/mia-2019-22-105>.
- 9 Stepanov, V.D., Shambilova, G.É. On the Iterated Integral Operators on the Cone of Monotone Functions. *Sib Math J.*, 66, 345–363 (2025). <https://doi.org/10.1134/S0037446625020119>.
- 10 Křepela, M. Integral conditions for Hardy-type operators involving suprema. *Collectanea Mathematica*, 68, 21–50 (2017). <https://doi.org/10.1007/s13348-016-0170-6>.
- 11 Mustafayev, R.Ch., Bilgili, N. Generalized fractional maximal functions in Lorentz spaces A. *Journal of Mathematical Inequalities*, 12 (3), 827–851 (2018).
- 12 Bakhtigareeva, E.G., Goldman, M.L. On the relationship between embeddings and coverings of cones of functions, *Mat. Sb.*, 216 (3), 26–48 (2025). <https://doi.org/10.4213/sm10199>.
- 13 Gogatishvili, A., Pick, L. Discretization and anti-discretization of rearrangementinvariant norms. *Publicacions Matemàtiques*, 47 (2), 311–358 (2003).
- 14 Bokayev, N.A., Gogatishvili, A., Abek, A.N. Cones of monotone functions generated by generalized fractional maximal function. *TWMS Journal of Pure and Applied Mathematics*, 15 (1), 127–141 (2024). <https://doi.org/10.30546/22191259.15.1.2024.2487>.
- 15 Unver, T. Embeddings between weighted Cesaro function spaces. *Mathematical Inequalities and Applications*, 23 (3), 925–942 (2020). <http://dx.doi.org/10.7153/mia-2020-23-72>.
- 16 Unver, T. Embeddings between weighted Tandori and Cesaro function spaces. *Commun. Faculty of Sciences Univ. of Ankara Series A1 Mathematics and Statistics*, 70 (2), 837–848 (2021).
- 17 Evans, W.D., Gogatishvili, A., Opic, B. *Weighted inequalities involving $\mathbf{P}$ -quasiconcave operators*. World Scientific Publishing Co.Pte.Ltd., Hackensack, NJ. 2018. ISBN:978-981-3239-62-3.

^{1*}Әбек А.Н.,

PhD, ORCID ID: 0009-0004-7158-3597,

*e-mail: azhar.abekova@gmail.com

²Гогатишвили А.,

PhD, ORCID ID: 0000-0003-3459-0355,

e-mail: gogatish@math.cas.cz

³Бокаев Н.А.,

ф.-м.ғ.д., профессор, ORCID ID: 0000-0002-7071-1882,

e-mail: bokayev2011@yandex.kz

⁴Унвер Т.,

PhD, ORCID ID: 0000-0003-0414-8400,

e-mail: tugceunver@kku.edu.tr

^{1,3}Л.Н. Гумилев атындағы Еуразия ұлттық университеті, Астана қ., Қазақстан

²Чехия ғылым Академиясының математика институты, Прага қ., Чехия

⁴Кириккале университеті, Кириккале қ., Түркия

ҮШ ОПЕРАТОРДЫҢ СУПЕРПОЗИЦИЯСЫ ҮШІН САЛМАҚТЫ ТЕҢСІЗДІКТЕР

Аңдатпа

Копсон, Харди және Тандори операторларының суперпозициясы үшін үш салмақтық теңсіздікті зерттейміз. Бұл жұмыстың мақсаты – $L_1(v)$ -ден $L_q(w)$ дейінгі салмақты Лебег кеңістіктеріндегі осы үш

оператордың комбинациясы болатын оператордың шенелгендігінің толық сипаттамасын дәлелдеу. Негізгі назар оң нақты осьтегі барлық теріс емес өлшемді функциялар үшін осы теңсіздіктің орындалуының қажетті және жеткілікті шарттарын анықтауға бағытталған. Өспелі функцияға қатысты Борел өлшемінің іргелі функциясының түсінігі айтарлықтай қолданылады. Тандори операторы сызықтық оператор болмағандықтан, біз бұрынғы жұмыстарда қолданылған екі жақтылық әдістерін пайдалана алмаймыз. Бұл мәселені шешу үшін біз бұрын белгілі әдістердің күрделілігін болдырмайтын жаңа жеңілдетілген дискретизация әдісін әзірлейміз. Нәтижелердің дәлдігі мен оңтайлылығын көрсететін теңсіздіктегі ең жақсы тұрақтының айқын түрі алынды. Осы құрама операторлардың шенелгендігі үшін қажетті және жеткілікті шарттарды орнату арқылы біз Гогатишвили А., Пик Л., Опич Б. [1] еңбектерінде бұрын белгіленген теңсіздіктерді жақсартамыз. Жұмыста алынған нәтижелер функционалдық кеңістіктердегі салмақты теңсіздіктер және операторлық талдау саласындағы бар зерттеулерді кеңейтеді және толықтырады және жуықтау теориясында, гармоникалық талдауда және қатысты салаларда әлеуетті қолданбаларды ұсынады.

Тірек сөздер: салмақты теңсіздік, супремалды оператор, Копсон операторы, Харди операторы, Тандори операторы, ең жақсы тұрақты, операторлардың суперпозициясы, дискретизация.

^{1*}**Abek A.N.,**

PhD, ORCID ID: 0009-0004-7158-3597,

*e-mail: azhar.abekova@gmail.cm

²**Gogatishvili A.,**

PhD, ORCID ID: 0000-0003-3459-0355,

e-mail: gogatish@math.cas.cz

³**Bokayev N.A.,**

Dr. Phys.-Math. Sc., Professor, ORCID ID: 0000-0002-7071-1882,

e-mail: bokayev2011@yandex.kz

⁴**Unver T.,**

PhD, ORCID ID: 0000-0003-0414-8400,

e-mail: tugceunver@kku.edu.tr

^{1,3}L.N. Gumilyov Eurasian national university, Astana, Kazakhstan

²Institute of Mathematics of the Czech Academy of Sciences, Prague, Czech Republic

⁴Kirikkale University, Kirikkale, Turkey

WEIGHTED INEQUALITIES FOR A SUPERPOSITION OF THE THREE OPERATORS

Abstract

We study a three-weight inequality for a superposition of the Copson, Hardy, and Tandori operators. The goal of this paper is to prove a complete characterization of the boundedness of the operator that is a combination of these three operators in weighted Lebesgue spaces from $L_1(v)$ to $L_q(w)$. The main focus is on determining necessary and sufficient conditions under which this inequality holds for all non-negative measurable functions on the positive real axis. The notion of a fundamental function of a Borel measure with respect to an increasing function is used substantially. Since the Tandori operator is not a linear operator, we cannot use the duality methods used in earlier works. To solve this problem, we develop a new, simplified discretization method that avoids the complexities of previously known methods. An explicit form of the best constant in the inequality is obtained, demonstrating the accuracy and optimality of the results. By establishing necessary and sufficient conditions for the boundedness of these composite operators, we improve the inequalities previously established in the works of Gogatishvili A., Pick L., Opic B. [1]. The results obtained in the paper extend and complement existing research in the field of weighted inequalities and operator analysis in function spaces and offer potential applications in approximation theory, harmonic analysis and related areas.

Keywords: weighted inequality, supremal operator, Copson operator, Hardy operator, Tandori operator, best constant, superposition of operators, discretization.

Дата поступления статьи в редакцию: 18.08.2025

UDC 510.67
IRSTI 27.03.66<https://doi.org/10.55452/1998-6688-2025-22-4-306-312>¹Baizhanov B.,

Dr. Phys.-Math. Sc., Professor, ORCID ID: 0000-0002-3743-7404,

e-mail: baizhanov@math.kz

^{1*}Sargulova F.,

PhD student, ORCID ID: 0000-0001-5457-5078,

e-mail: fsargulova@gmail.com

¹Institute of Mathematics and Mathematical Modeling, Almaty, Kazakhstan

UNIFORMLY EXTERNALLY DEFINABLE EXPANSION

Abstract

In this article, we study the expansion of a structure by adding a new predicate that is not definable by any formula in the original language. To consider an externally definable expansion, we define the extension of a model in both essential and non-essential case. Such expansions can lead to significant changes in the properties of the resulting structure. We focus on the case of externally definable expansions, where the new relation is given by the intersection of a formula defined in an elementary extension with the original structure. The concept of a uniformly externally definable expansion was first introduced by Macpherson, Marker, and Steinhorn in the context of expansions by cuts in submodels of o-minimal structures over the real numbers. Subsequently, Baizhanov demonstrated that expanding a model of a weakly o-minimal theory by a family of convex sets preserves both weak o-minimality and uniform external definability. We establish conditions for external expansions under which the key properties of the original structure are preserved.

Keywords: expansion, externally definable expansion, convex-to-right (left) 2-formula, quasineighborhood and neighborhood.

Introduction

External definability. Let $\mathfrak{M} = \langle M, \Sigma \rangle$ be a structure of a signature Σ . We consider an extension $\Sigma^+ = \Sigma \cup \{U^n\}$ of the signature Σ . We say that $\mathfrak{M}^+ = \langle M, \Sigma^+ \rangle$ is an expansion of $\mathfrak{M}$, if $U \notin \Sigma$. An expansion is essential if there is a formula $\psi^+(\bar{x}, \bar{y})$ of the signature Σ^+ such that $\psi^+(\mathfrak{M}^+, \bar{a}) \neq \emptyset(\mathfrak{M}, \bar{b})$ for any definable set $\emptyset(\mathfrak{M}, \bar{b})$ of $\mathfrak{M}$. In the opposite case we say that the expansion is non-essential. In particular, any constant expansion is non-essential. Let $\mathfrak{M}$ be an elementary substructure of $\mathfrak{N}$ $\mathfrak{M} < \mathfrak{N}$. We say that U is externally definable, if there exists a formula $\psi(\bar{x}, \bar{a})$ of the signature Σ with $\bar{a} \in N$ and $\bar{a} \notin M$ such that

$$U(\mathfrak{M}^+) = \psi(\mathfrak{N}, \bar{a}) \cap M^n.$$

We say that an expansion is externally definable, if any definable subset in the expanded language is externally definable.

Let $A \subseteq M$. A complete type $p(\bar{x}) \in S(A)$ is said to be definable, if for any formula $\psi(\bar{x}, \bar{y})$ there exists $\emptyset_\psi(\bar{y}, \bar{a})$, where $\bar{a} \in A$, such that for any $\bar{b} \in A$

$$[\psi(\bar{x}, \bar{b}) \in p \Leftrightarrow \mathfrak{M} \models \emptyset_\psi(\bar{b}, \bar{a})].$$

S. Shelah proved that a complete theory T is stable if and only if for any model $\mathfrak{M} = \langle M; \Sigma \rangle \models T$ and for any set $A \subseteq M$ any type over A is definable. Thus, it is well-known that if a theory T is stable, any externally definable expansion is non-essential. Indeed, we consider an externally definable expansion of a model $\mathfrak{m}$ of a stable theory such that for $\psi(\bar{x}, \bar{a})$, $\bar{a} \in N \setminus M$, $\psi(\mathfrak{N}, \bar{a}) \cap M = U(\mathfrak{M}^+)$,

here $\mathfrak{M}$ is an elementary substructure of $\mathfrak{N}$. Let $\bar{b} \in M$ be such that $\mathfrak{M}^+ \models U(\bar{b})$. Then $\mathfrak{M} \models \psi(\bar{b}, \bar{a})$, it means that $\psi(\bar{b}, \bar{y}) \in tp(\bar{a}|M)$. Since T is stable then $tp(\bar{a}|M)$ is definable, and consequently there exists an M -definable formula $\Theta_\psi(\bar{y}, \bar{a})$, $\bar{a} \in M$ such that

$$[\psi(\bar{b}, \bar{y}) \in tp(\bar{a}|M) \Leftrightarrow \mathfrak{M} \models \Theta(\bar{b}, \bar{a})].$$

Thus,

$$U(\mathfrak{M}^+) = \Theta_\psi(\mathfrak{M}, \bar{a}).$$

In each formula of Σ^+ we replace $U(\bar{y})$ with $\Theta_\psi(\bar{y}, \bar{a})$. So, any parametrically definable set in $\mathfrak{M}^+$ is equal to a parametrically definable set $\mathfrak{M}$. This means that this expansion is non-essential. Thus, any externally definable expansion of a model of a stable theory is non-essential expansion. So, we consider externally definable expansions for unstable theories.

Materials and methods

Bruno Poizat introduced the notion of a beautiful pair. We say that a pair $(\mathfrak{N}, \mathfrak{M})$ of models of T is beautiful, if $\mathfrak{M}$ is $|T|^+$ -saturated and for each $\bar{a} \in N$ each type over $M \cup \{\bar{a}\}$ is realized in $\mathfrak{N}$. Let $\bar{a} \in N \setminus M$ and $p := tp(\bar{a}|M)$. Then for each formula $\psi(\bar{x}, \bar{y})$ we define the predicate $R_{(\psi, p)}(\bar{y})$ on the set M as follows:

$$\mathfrak{M}^+ \models R_{(\psi, p)}(\bar{a}) \Leftrightarrow \psi(\bar{x}, \bar{a}) \in tp(\bar{a}|M) \Leftrightarrow \mathfrak{N} \models \psi(\bar{a}, \bar{a}).$$

Here, we denote the expanded structure by $\mathfrak{M}^+ = \langle M; \Sigma^+ \rangle$, where

$$\Sigma^+ := \{R_{(\psi, p)}(\bar{y}) \mid p \in S(M), \psi \in \Sigma\}.$$

We say that $\mathfrak{M}^+$ is externally definable if each new predicate is externally definable. It follows from definition that if a pair of models (M, N) is a conservative pair (the type of any tuple of elements from N over M is definable), then the structure $\mathfrak{M}^+$ is the structure obtained from $\mathfrak{M}$ by Scolemization of $\mathfrak{M}$. In the paper, we consider the cases where $\mathfrak{M}^+$ is constructed from one 1-type for an unstable theory.

Let $\mathfrak{M}$ be a model of a complete theory T of a signature Σ . We say that $\mathfrak{M}_p^+$ is the expansion of $\mathfrak{M}$ by the type $p \in S_1(M)$, if $\mathfrak{M}_p^+ := \langle M; \Sigma_p^+ \rangle$, where $\Sigma_p^+ := \{R_{(\psi, p)}(\bar{y}) \mid \psi \in \Sigma\}$.

An externally definable expansion $\mathfrak{M}^+ = \langle M; \Sigma \cup \{U^1\} \rangle$ is a uniformly externally definable expansion, if for some $\mathfrak{N} \succ \mathfrak{M}$, for any $\varphi(\bar{x})$ of Σ^+ , there exists $K_\varphi(\bar{x}, \bar{a})$ of Σ with $\bar{a} \in N$, such that for each $\bar{a} \in M$ the following holds:

$$\mathfrak{M}^+ \models \varphi(\bar{a}) \Leftrightarrow \mathfrak{N} \models K_\varphi(\bar{a}, \bar{a}).$$

Originally, the notion of a uniformly externally definable expansion was considered in the paper by D. Macpherson, D. Marker, and C. Steinhorn for a cut in a submodel of an o-minimal model with the universe consisting of all real numbers [2]. B. Baizhanov proved that an expansion of a model of a weakly o-minimal theory by a family of convex sets is weakly o-minimal and uniformly externally definable [3]. S. Shelah in proved that the expansion of a model of a dependent theory (NIP) by all externally definable relations is uniformly externally definable and NIP [4].

Results and discussion

A set $A \subset M$ of linearly ordered structure $\mathfrak{M}$ is said to be convex, if for any $a, b \in A$ such that $a < b$, it follows for any $c \in M$ ($a \leq c \leq b \rightarrow c \in A$). A 2-formula $\psi(x, y)$ is said to be convex-to-right if for any $\alpha \in (\exists x\psi(x, y))(\mathfrak{M})$, the set $\psi(\mathfrak{M}, \alpha)$ is a convex set and $\alpha = \inf\psi(\mathfrak{M}, \alpha)$. Similarly, a 2-formula $\psi(x, y)$ is said to be convex to the left if for any $\alpha \in (\exists x\psi(x, y))(\mathfrak{M})$, the set $\psi(\mathfrak{M}, \alpha)$ is a convex set and $\alpha = \sup\psi(\mathfrak{M}, \alpha)$.

Definition 1. Let $\mathfrak{M} = \langle M; <, \dots \rangle$ be a linearly ordered structure, $A \subseteq M$, and let $\varphi(x, \bar{y})$ be a definable formula [5].

1) The convex closure of a formula $\varphi(x, \bar{a})$ is the following formula:

$$\varphi^c(x, \bar{a}) := \exists y_1 \exists y_2 (\varphi(y_1, \bar{a}) \wedge \varphi(y_2, \bar{a}) \wedge (y_1 \leq x \leq y_2)).$$

2) The convex closure of a type $p(x) \in S_1(A)$ is the following type:

$$p^c(x) := \{\varphi^c(x, \bar{a}) \mid \varphi(x, \bar{a}) \in p\}.$$

Main Theorem. Let $\mathfrak{M} = \langle M; <, \dots \rangle$ be a linearly ordered structure. Let A and B be subsets of M . We write $A < B$ if $a < b$ for any $a \in A$ and $b \in B$. We write $A < b$ if $A < \{b\}$ and $b < A$ if $\{b\} < A$. A partition of M to C and D is said to be a cut, if $C < D$. A cut (C, D) is said to be irrational, if $\theta(\mathfrak{M}) \neq C$ and $\theta(\mathfrak{M}) \neq D$ for any M -definable 1-formula $\theta(x)$.

Theorem 1. Let $\mathfrak{M} = \langle M; \Sigma \rangle$ be a structure, where $\Sigma = \{=, <, \dots\}$ and $<$ is interpreted as a linear order. Let (C, D) be an irrational cut in $\mathfrak{M}$ and let for some $\mathfrak{N} > \mathfrak{M}$ there exists $\alpha \in N \setminus M$ such that $C < \alpha < D$. Assume also that for any formula $H(x, y)$, such that $H(\mathfrak{N}, \alpha)$ is convex-to-right or to the left and contains an element different from α , it holds that $H(\mathfrak{N}, \alpha) \cap M \neq \emptyset$. Then $\mathfrak{M}^+ = \langle M; \Sigma \cup \{U^1\} \rangle$ is a uniformly externally definable expansion, where $U(\mathfrak{M}^+) = C$.

Proof. For each formula $L(x)$ we define

$$E_L(x, y) := L(x) \wedge L(y) \wedge \forall z ((x \leq z \leq y \vee y \leq z \leq x) \rightarrow L(z)).$$

We note that E_L is an equivalence relation on $L(\mathfrak{M})$ such that each E_L -class is a convex set.

Lemma 3. Let $p \in S_1(M)$ be such that $\theta(N, \alpha) \cap M \neq \emptyset$ for any $\alpha \in p^c(N)$ and for any convex to the left M -formula $\theta(x, y)$ that is not equivalent to $x = y$. Let $L(x)$ be an M -formula and $\beta \in L(N)$.

If $E_L(N, \beta) \cap p(N) \neq \emptyset$, then $p^c(N) \subset E_L(N, \beta)$.

Proof. By definition of E_L , it holds that $L(N) = \bigcup_{\beta \in L(N)} E_L(N, \beta)$. By definition of a type, we obtain that $L(x) \in p \Leftrightarrow$ there is $\beta \in L(N) \cap p(N)$. Then $E_L(x, \beta) \wedge x < \beta$ is a formula that is convex to the left. By the hypothesis of this lemma, $E_L(N, \beta) \cap M \neq \emptyset$.

Let $b \in E_L(N, \beta) \cap M$ and $E_L(N, \beta) = E_L(N, b)$. We consider the right side of $E_L(x, \beta)$, that is, $E_L(x, \beta) \wedge x > \beta$, and conclude that there is $c \in M$ such that $E(N, c) = E(N, \beta)$, that is why $p^c(N) < c$. Then $b < \beta < c$ and $\mathfrak{M} \models E_L(b, c)$. The last implies the following inclusions: $p(N) \subseteq p^c(N) \subset E_L(N, \beta)$. $\square$

We prove that for each formula $\varphi(x_1, x_2, \dots, x_n)$ of $\Sigma^+ = \Sigma \cup \{U\}$ there exists a formula $K_\varphi(x_1, \dots, x_n, \bar{a})$ and $\bar{a} \in N \setminus M$ such that $\varphi(\mathfrak{M}^+) = K_\varphi(\mathfrak{N}^n, \bar{a}) \cap M^r$.

We prove it by induction in the complexity of the construction of φ . First, we assume that $\varphi(x_1 \dots x_n)$ has the following form:

$$\begin{aligned} \varphi(x_1, x_2, \dots, x_n) = & \left(H_1(x_1 \dots x_n) \wedge \bigwedge_{i=1}^n \epsilon_i^1 U(x_i) \right) \vee \left(H_2(x_1 \dots x_n) \wedge \bigwedge_{i=1}^n \epsilon_i^2 U(x_i) \right) \vee \\ & \vee \dots \vee \left(H_s(x_1 \dots x_n) \wedge \bigwedge_{i=1}^n \epsilon_i^s U(x_i) \right), \end{aligned}$$

where $H_1(x_1, \dots, x_n), H_2(x_1, \dots, x_n), \dots, H_s(x_1, \dots, x_n)$ are formulas of Σ , and

$$\begin{cases} \epsilon_i^j U(x) = U(x), & \text{if } \epsilon_i^j = 1, \text{ and} \\ \epsilon_i^j U(x) = \neg U(x), & \text{if } \epsilon_i^j = 0. \end{cases}$$

We put $\epsilon^j = \langle \epsilon_1^j, \epsilon_2^j, \dots, \epsilon_n^j \rangle$.

Let $\bar{a} \in M^n$. Assume that $\mathfrak{M}^+ \models \varphi(\bar{a})$. Then

$$\mathfrak{M}^+ \models \left(H_1(\bar{a}) \wedge \bigwedge_{i=1}^n \epsilon_i^1 U(a_i) \right) \vee \left(H_2(\bar{a}) \wedge \bigwedge_{i=1}^n \epsilon_i^2 U(a_i) \right) \vee \dots \vee \left(H_s(\bar{a}) \wedge \bigwedge_{i=1}^n \epsilon_i^s U(a_i) \right).$$

Suppose, that $\mathfrak{M}^+ \models H_1(\bar{a}) \wedge \bigwedge_{i=1}^n \epsilon_i^1 U(a_i)$. Then $\mathfrak{M} \models H_1(\bar{a})$, and then $\mathfrak{M}^+ \models {}^1 U(a_i)$ if and only if $\mathfrak{N} \models a_i < \alpha$; as well as $\mathfrak{M}^+ \models {}^0 U(a_i)$ if and only if $\mathfrak{N} \models \alpha < a_i$.

We introduce the following notation. We write

$$\begin{cases} \mathfrak{N} \models (a_i < \alpha)^{\epsilon_i} \Leftrightarrow \mathfrak{N} \models a_i < \alpha, & \text{if } \epsilon_i^1 = 1, \text{ and} \\ \mathfrak{N} \models (a_i < \alpha)^{\epsilon_i} \Leftrightarrow \mathfrak{N} \models \alpha < a_i, & \text{if } \epsilon_i^1 = 0. \end{cases}$$

Then

$$\mathfrak{N} \models H_1(\bar{a}) \wedge \bigwedge_{i=1}^n (a_i < \alpha)^{\epsilon_i^1}$$

implies that

$$\begin{aligned} \mathfrak{N} \models & \left(H_1(\bar{a}) \wedge \bigwedge_{i=1}^n (a_i < \alpha)^{\epsilon_i^1} \right) \vee \left(H_2(\bar{a}) \wedge \bigwedge_{i=1}^n (a_i < \alpha)^{\epsilon_i^2} \right) \vee \dots \vee \\ & \vee \left(H_s(\bar{a}) \wedge \bigwedge_{i=1}^n (a_i < \alpha)^{\epsilon_i^s} \right). \end{aligned}$$

We also denote

$$K_\varphi(x_1, \dots, x_n, \alpha) := \bigvee_{j=1}^s \left(H_j(\bar{a}) \wedge \bigwedge_{i=1}^s (a_i < \alpha)^{\epsilon_i^j} \right).$$

Then $\mathfrak{M}^+ \models \varphi(\bar{a})$ if and only if $\mathfrak{N} \models K_\varphi(\bar{a}, \alpha)$.

Suppose that for a formula $\psi(y, \bar{x})$ of Σ^+ there exists a formula $K_\psi(y, x, \alpha)$ of Σ such that for any $c, \bar{a} \in M$ we have that

$$\mathfrak{M}^+ \models \psi(c, \bar{a}) \Leftrightarrow \mathfrak{N} \models K_\psi(c, \bar{a}, \alpha).$$

Then we construct $K_\varphi(\bar{x}, \alpha)$ for the formula $\varphi(\bar{x}) := \exists y \psi(y, \bar{x})$.

We claim that

$$K_\varphi(\bar{x}, \alpha) := \exists z \exists y (K_\psi(y, \bar{x}, \alpha) \wedge \forall v (z \leq v < \alpha \rightarrow K_\psi(y, \bar{x}, v)))$$

satisfies the requirement.

($\Rightarrow$) Let $\bar{a} \in M$ be such that $\mathfrak{N} \models K_\varphi(\bar{a}, \alpha)$, this means

$$\mathfrak{N} \models \exists z \exists y (K_\psi(y, \bar{a}, \alpha) \wedge \forall v (z \leq v < \alpha \rightarrow K_\psi(y, \bar{a}, v))).$$

By Lemma 3 we consider the formula

$$F(z, \bar{a}, \alpha) := \exists y(K_\psi(y, \bar{a}, \alpha) \wedge \forall v(z \leq v < \alpha \rightarrow K_\psi(y, \bar{a}, v))).$$

Note that $F(z, \bar{a}, \alpha)$ is convex to the left. By the hypothesis of the theorem there exists $b \in M$ such that

$$\mathfrak{N} \models F(b, \bar{a}, \alpha) \Leftrightarrow \mathfrak{N} \models \exists y(K_\psi(y, \bar{a}, \alpha) \wedge \forall v(b \leq v < \alpha \rightarrow K_\psi(y, \bar{a}, v));$$

that is why $F(b, \bar{a}, t) \in tp(\alpha/M)$. By Lemma 3 there exists $c \in M$ such that

$$\mathfrak{N} \models F(b, \bar{a}, c) \wedge b < \alpha < c,$$

this means that

$$\mathfrak{N} \models \exists y(K_\psi(y, \bar{a}, c) \wedge \forall v(b \leq v < c \rightarrow K_\psi(y, \bar{a}, v))).$$

Since $\mathfrak{N} < \mathfrak{M}$, so

$$\mathfrak{M} \models \exists y(K_\psi(y, \bar{a}, c) \wedge \forall v(b \leq v < c \rightarrow K_\psi(y, \bar{a}, v))).$$

Let $m \in M$ be such that

$$\mathfrak{M} \models K_\psi(m, \bar{a}, c) \wedge \forall v(b \leq v < c \rightarrow K_\psi(m, \bar{a}, v)).$$

Then $\mathfrak{N} \models \forall v(b \leq v < c \rightarrow K_\psi(m, \bar{a}, v))$. Since $b < \alpha < c$, we obtain $\mathfrak{N} \models K_\psi(m, \bar{a}, \alpha)$. By the induction assumption $\mathfrak{M}^+ \models \psi(m, \bar{a})$. So, $\mathfrak{M}^+ \models \exists y\psi(y, \bar{a})$ and $\mathfrak{M}^+ \models \varphi(\bar{a})$.

($\Leftarrow$) Let $\mathfrak{M}^+ \models \varphi(\bar{a})$. Recall that $\varphi(\bar{x}) = \exists y(\psi(y, \bar{x}))$ then $\mathfrak{M}^+ \models \exists y(\psi(y, \bar{a}))$ and, consequently $\mathfrak{M}^+ \models \psi(m, \bar{a})$ for some $m \in M$. By the induction hypothesis,

$$\mathfrak{M}^+ \models \psi(m, \bar{a}) \Leftrightarrow \mathfrak{N} \models K_\psi(m, \bar{a}, \alpha).$$

We consider p^c , which is the convex hull of $p = tp(\alpha/M)$. By Lemma and since $K_\psi(m, \bar{a}, t)$ is convex, there exists $b, c \in M$ such that $b < p^c(N) < c$. So, we have

$$\mathfrak{N} \models (K_\psi(m, \bar{a}, \alpha) \wedge \forall v(b \leq v < c \rightarrow K_\psi(m, \bar{a}, v)))$$

and, consequently

$$\mathfrak{N} \models \exists y(K_\psi(y, \bar{a}, \alpha) \wedge \forall v(b \leq v < c \rightarrow K_\psi(y, \bar{a}, v))).$$

Since $b < \alpha < c$, we obtain

$$\mathfrak{N} \models \exists y(K_\psi(y, \bar{a}, \alpha) \wedge \forall v(b \leq v < \alpha \rightarrow K_\psi(y, \bar{a}, v))),$$

and

$$\mathfrak{N} \models \exists z \exists y(K_\psi(y, \bar{a}, \alpha) \wedge \forall v(z < v < \alpha \rightarrow K_\psi(y, \bar{a}, v))).$$

Then $\mathfrak{N} \models K_\varphi(\bar{a}, \alpha)$. $\square$

Conclusion

In conclusion, our investigation of externally definable expansions highlights the delicate balance between enriching a structure with new predicates and preserving its fundamental model-theoretic properties. By distinguishing between essential and non-essential extensions, we clarify when such expansions maintain the behavior of the original model. Building on earlier work in o-minimal and weakly o-minimal settings, we provide general conditions under which key properties — such as uniform external definability are retained. These results contribute to a deeper understanding of the model-theoretic stability of structures under externally definable expansions.

Funding

This research is funded by the Science Committee of the Ministry of Science and Higher Education of the Republic of Kazakhstan (Grant No. AP 19677434).

REFERENCES

- 1 Shelah S. Classification theory and the number of non-isomorphic models, Studies in Logic and the Foundations of Mathematics, 1978, 2nd edition 1990, Elsevier, ISBN 978-0-444-70260-9.
- 2 Macpherson, D., Marker, D., Steinhorn, Ch. Weakly o-minimal structures and real closed fields, Translations of the American Mathematical Society, 352, 5435–5483 (2000).
- 3 Baizhanov, B. Expansion of a model of a weakly o-minimal theory by a family of unary predicates, The Journal of Symbolic Logic, 66, 1382–1414 (2001).
- 4 Shelah, S. Dependent first order theories. Continued, Israel Journal of Mathematics 173, 1 (2009). <https://doi.org/10.1007/s11856-009-0082-1>.
- 5 Baizhanov, B.S., Verbovskii, V.V. O-Stable Theories. Algebra and Logic, 50, 211–225 (2011).

¹Байжанов Б.,

ф.-м.ғ.д., профессор, ORCID ID: 0000-0002-3743-7404,

e-mail: baizhanov@math.kz

^{1*}Саргулова Ф.,

докторант, ORCID ID: 0000-0001-5457-5078,

*e-mail: fsargulova@gmail.com

¹Математика және математикалық модельдеу институты, Алматы қ., Қазақстан

БІРКЕЛКІ СЫРТТАЙ АНЫҚТАЛАТЫН БАЙЫТУ

Аңдатпа

Бұл мақалада құрылымдағы тілді байыту мәселелері қарастырылады. Дәл осы кезде құрылымның ешбір формуласымен сәйкес келмейтін жаңа предикат қосылады. Сырттай анықталатын байытуларды қарастыру үшін модельдің маңызды және маңызды емес байытылу ұғымдары енгізіледі. Бұл жағдайда жаңадан алынған құрылымның қасиеттері айтарлықтай өзгеруі мүмкін. Жаңа қатынас элементар байытылымда анықталған формуланың бастапқы құрылыммен қиылысуы болған кезде сырттай анықталатын байыту жағдайы қарастырылады. Бірқалыпты сырттай анықталатын байыту ұғымын алғаш рет Макферсон, Маркер және Стайнхорн нақты сандар жиынындағы о-минимал құрылымдардың ішкі модельдеріндегі қиылысу арқылы байыту контекстінде енгізген. Кейінірек Байжанов әлсіз о-минималды теория моделін дөңес жиындар топтастығы арқылы байыту бұл модельдің әрі әлсіз о-минималдылығын, әрі бірқалыпты сырттай анықталатындығын сақтайтынын дәлелдеді. Сыртқы байыту шарттары анықталып, онда бастапқы құрылымның негізгі қасиеттері сақталады.

Тірек сөздер: байыту, сырттай анықталатын байыту, дөңес оңға (солға) 2-формула, көршілес және квази-көршілес.

¹Байжанов Б.,

д.ф.-м.н., профессор, ORCID ID: 0000-0002-3743-7404

e-mail: baizhanov@math.kz

^{1*}Саргулова Ф.,

докторант, ORCID ID: 0000-0001-5457-5078

*e-mail: fsargulova@gmail.com

¹Институт математики и математического моделирования, г. Алматы, Казахстан

РАВНОМЕРНО ВНЕШНЕ ОПРЕДЕЛИМОЕ ОБОГАЩЕНИЕ

Аннотация

В этой статье мы рассматриваем вопросы обогащения языка в структуре, когда добавляется новый предикат, не совпадающий ни с одной формулой структуры. Для рассмотрения внешне определимых расшире-

ний мы вводим понятие расширения модели в существенном и несущественном случаях. При этом могут существенно меняться свойства вновь полученной структуры. Рассмотрен случай внешне определимого обогащения, когда новое отношение является пересечением формулы, определенной в элементарном расширении, с исходной структурой. Понятие равномерно внешне определимого расширения было впервые введено Макферсоном, Маркером и Стайнхорном в контексте расширений с помощью сечений в подмоделях о-минимальных структур над множеством вещественных чисел. Впоследствии Байжанов показал, что расширение модели слабо о-минимальной теории семейством выпуклых множеств сохраняет как слабую о-минимальность, так и равномерную внешнюю определимость. Получены условия на внешнее обогащение, при котором сохраняются основные свойства исходной структуры.

Ключевые слова: обогащение, внешне определимое обогащение, выпуклая вправо (влево) 2-формула, окрестность и квазиокрестность.

Article submission date: 17.10.2025

УДК 517.5
МРНТИ 27.39.29

<https://doi.org/10.55452/1998-6688-2025-22-4-313-323>

^{1*}**Басаров С.Ж.**,
магистр, м.н.с., ORCID ID: 0000-0001-5422-2424,
*e-mail: bassarov.serzhan98@gmail.com
²**Нурсултанов Е.Д.**,
д.ф.-м.н., профессор, ORCID ID: 0000-0003-3879-2261,
e-mail: er-nurs@yandex.kz

¹ТОО «Geometry», г. Астана, Казахстан
²Казахстанский филиал Московского государственного университета
им. М.В. Ломоносова, г. Астана, Казахстан

ВОССТАНАВЛИВАЮЩАЯ ФУНКЦИЯ ДЛЯ ПРОСТРАНСТВ СОБОЛЕВА С ДОМИНИРУЮЩЕЙ СМЕШАННОЙ ПРОИЗВОДНОЙ

Аннотация

Линейная аппроксимация – это приближение функции из некоторого класса элементами фиксированного конечного подпространства этого же класса. Например, для одномерных периодических функций такими элементами служат тригонометрические полиномы. В интересующем нас многомерном случае для функций, периодических по каждой переменной, в качестве такого подпространства выступает множество тригонометрических полиномов со спектром из ступенчатого гиперболического креста. При этом возникает вопрос выбора коэффициентов для этих полиномов. В данной статье представлен аппарат, восстанавливающий функции из пространств Соболева с доминирующей смешанной производной по заданным точкам и получены оценки погрешности восстановления. Метод основан на построении восстанавливающей функции в виде полинома со спектром из ступенчатого гиперболического креста, коэффициенты которого вычисляются по заданным точкам. Погрешность приближения имеет порядок ортопоперечника, что является оптимальным результатом для таких полиномов. Предложенный метод точен для полиномов со спектром из ступенчатого гиперболического креста. Также получен функционал, восстанавливающий коэффициенты Фурье для функций из указанных пространств. Благодаря явному выражению восстанавливающей функции полученная формула может быть использована для решения прикладных задач.

Ключевые слова: много переменных, теория функций, пространство Соболева, смешанная производная, гиперболический крест.

Введение

Пусть $n \in \mathbb{N}, 1 \leq q < \infty$. Пространство Лебега $L_q = L_q[0,1]^n$ будем называть множество измеримых функций $f(x_1, \dots, x_n)$, для которых

$$\|f\|_{L_q} = \left(\int_0^1 \dots \int_0^1 |f(x_1, \dots, x_n)|^q dx_1 \dots dx_n \right)^{1/q} < \infty.$$

В случае $q = \infty$:

$$\|f\|_{L_\infty} = \operatorname{ess\,sup}_{x \in [0,1]^n} |f(x_1, \dots, x_n)| < \infty.$$

Определение 1. Пусть $\alpha > 0, 1 \leq q \leq \infty$. Пространство Соболева $W_q^\alpha = W_q^\alpha[0,1]^n$ с доминирующей смешанной производной будем называть множеством измеримых функций f таких, что ряд Фурье функции $f^\alpha \in L_n$ (производная по Вейлю) совпадает с рядом $\sum_{k \in \mathbb{Z}^n} \bar{k}^\alpha \hat{f}(k) e^{2\pi i(k,x)}$, где $\bar{k}^\alpha = \prod_{j=1}^n \bar{k}_j^\alpha$, $\bar{k}_j^\alpha = \max(|k_j^\alpha|, 1)$, $\hat{f}(k)$ – коэффициент Фурье функции f , $(k, x) = \sum_{j=1}^n k_j x_j$. К тому же норма определяется следующим образом:

$$\|f\|_{W_q^\alpha} = \|f^\alpha\|_{L_q}.$$

Определение 2. [1] Пусть $m \in \mathbb{N}$ и $p > 1$ – простое. Множество

$$G_m = \bigcup_{\substack{v_1 + \dots + v_n \leq m, \\ v_j = 0, \dots, m}} \rho(v_1, \dots, v_n),$$

$$\rho(v_1, \dots, v_n) = \{(\mu_1, \dots, \mu_n) \in \mathbb{Z}^n : [p^{v_j-1}] \leq |\mu_j| < p^{v_j}, j = 1, \dots, n\}$$

будем называть ступенчатый гиперболический крест.

Определение 3. Пусть $m \in \mathbb{N}$ и $p > 1$ – простое. Функцию

$$S_{G_m}(x; f) = \sum_{\mu \in G_m} \hat{f}(\mu) e^{2\pi i (\mu, x)}$$

будем называть полиномом со спектром из ступенчатого гиперболического креста.

Примечание. Особенность этих полиномов в том, что их погрешности достигают по порядку ортопоперечника для пространств с доминирующей смешанной производной (см. теорему А. ниже) – в частности, для пространств Соболева. Полиномов такого вида можно рассматривать как аналог одномерной частной суммы ряда Фурье, но только для многомерных.

Определение 4. [2] Пусть H – нормированный класс функций, $\{u_j\}_{j=1}^M$ – ортонормальная система и $\langle \cdot, \cdot \rangle$ – скалярное произведение. Определим ортопоперечник как величину

$$\varphi_M(H)_q = \inf_{u_1, \dots, u_M} \sup_{f \in H} \left\| f - \sum_{j=1}^M \langle f, u_j \rangle u_j \right\|_{L_q}.$$

Теорема А. [2-7] Пусть $p > 1$ – простое, $1 < q \leq \tau < \infty$, $\alpha > 1/q - 1/\tau$. Тогда

$$\varphi_M(W_q^\alpha)_\tau \asymp E_{G_m}(W_q^\alpha)_\tau \asymp \frac{1}{p^{(\alpha - \frac{1}{q} + \frac{1}{\tau})m}} \asymp \left(\frac{\log^{n-1} M}{M} \right)^{\alpha - 1/q + 1/\tau},$$

где $M \asymp m^{n-1} p^m$, $E_{G_m}(W_q^\alpha)_\tau = \inf_{t \in T(G_m)} \sup_{\|f\|_{W_q^\alpha}=1} \|f - t\|_{L_\tau}$, $T(G_m)$ – множество полиномов со спектром из ступенчатого гиперболического креста.

Цель данной статьи – получить погрешность восстанавливающей функции для пространств Соболева W_q^α . А именно получить оценку

$$\sup_{\|f\|_{W_q^\alpha}=1} \|f - F_m(\cdot; f)\|_{L_\tau} \asymp g(m),$$

где $F_m(x; f)$ – восстанавливающая функция, $g(m) \rightarrow 0$ при $m \rightarrow \infty$, константы, входящие в неравенства, не зависят от f, m .

Задача о восстановлении функции имеет богатую историю. С прошлого столетия было опубликовано много работ по этому вопросу, которые сделали значительный прогресс в области теории аппроксимации. В частности, для задач о восстановлении функции для пространств с доминирующей смешанной производной алгоритм Смоляка достигает по порядку ортопоперечника [8–16, а также ссылки, приведенные в них], что является оптимальным для параметров $1 < q < \tau < \infty$. В работе [17] была приведена восстанавливающая функция для пространств Соболева, используя техники гармонического анализа (для полного списка результатов и истории развития по теме восстановления функции см. обзор [18]).

Материалы и методы

Приведем несколько вспомогательных лемм.

Лемма 1. [19] Пусть $n \in \mathbb{N}, d \in \mathbb{N}^n, B = B_1 \times \dots \times B_n \in \mathbb{Z}^n$ – сегмент, $d_j > |B_j|, j = 1, \dots, n$,

$$I = \bigcup_{r \in \mathbb{Z}^n} (B_1 + r_1 d_1) \times \dots \times (B_n + r_n d_n).$$

Если $f \in L_1[0,1]^n$ и $f(x) \sim \sum_{r \in \mathbb{Z}^n} \hat{f}(r) e^{2\pi i(r,x)}$, тогда для функции

$$\frac{1}{d_1 d_2 \dots d_n} \sum_{s_1=0}^{d_1-1} \dots \sum_{s_n=0}^{d_n-1} f\left(x_1 + \frac{s_1}{d_1}, \dots, x_n + \frac{s_n}{d_n}\right) D_B\left(\frac{s_1}{d_1}, \dots, \frac{s_n}{d_n}\right),$$

где $D_B(x) = \sum_{k \in B} e^{-2\pi i(k,x)}$ – ядро Дирихле по сегменту B ее рядом Фурье является

$$\sum_{r \in I} \hat{f}(r) e^{2\pi i(r,x)}.$$

Лемма 2. [20, Лемма 4.3] Пусть $p > 1$ —простое, $b = \{b_k\}_{k=-\infty}^{+\infty} \in l_1$. Тогда

$$\sum_{m \in \mathbb{Z} \setminus \{0\}} b_m = \sum_{k=1}^{\infty} \sum_{j=1}^{p-1} \sum_{r \in \mathbb{Z}} b_{p^{k-1}(pr+j)}.$$

Следствие 1. [20, Лемма 4.4] Пусть $p > 1$ —простое, $b = \{b_k\}_{k=-\infty}^{+\infty} \in l_1$. Тогда

$$\sum_{k=0}^m \frac{(-1)^{\text{sign } k}}{(p-1)^{1-\text{sign } k}} \sum_{j=1}^{p-1} \sum_{r \in \mathbb{Z}} b_{p^{k-1}(pr+j)} = \sum_{r \in \mathbb{Z}} b_{p^m r}.$$

Результаты и обсуждения

Рассмотрим следующий функционал

$$\begin{aligned} T_{m,v}(f) &= T_{m,v}(f; p) \\ &= \frac{1}{p^m} \sum_{\substack{k_1+v_1+\dots+k_n+v_n=m, \\ k_j=0,\dots,m}} \sum_{r_1=0}^{p^{k_1+v_1}-1} \dots \sum_{r_n=0}^{p^{k_n+v_n}-1} \prod_{j=1}^{n-1} \frac{\sum_{l=1}^{p-1} e^{2\pi i\left(\frac{2lr_j}{p} + \text{sign } k_j\right)}}{(p-1)^{1-\text{sign } k_j}} f\left(\frac{r_1}{p^{k_1+v_1}}, \dots, \frac{r_n}{p^{k_n+v_n}}\right), \end{aligned}$$

где $p > 1$ —простое, $m \in \mathbb{N}, v = (v_1, \dots, v_n) \in \mathbb{Z}_+^n = \{0, 1, \dots\}^n, v_1 + \dots + v_n \leq m$.

Лемма 3. Пусть $n \in \mathbb{N}, f \in L_1[0,1]^n$ и ряд Фурье функции f абсолютно сходится. Тогда

$$\begin{aligned} T_{m,v}(f) &= \sum_{\substack{k_1+v_1+\dots+k_n+v_n=m, \\ k_j=0,\dots,m}} \prod_{j=1}^{n-1} \frac{(-1)^{\text{sign } k_j}}{(p-1)^{1-\text{sign } k_j}} \sum_{\beta_1=1}^{p-1} \dots \sum_{\beta_{n-1}=1}^{p-1} \sum_{r \in \mathbb{Z}^n} \hat{f}(p^{k_1-1}(pr_1 \\ &\quad + \text{sign } k_1, \dots, p^{k_{n-1}-1}(pr_{n-1} + \text{sign } k_{n-1}), p^{k_n} r_n). \end{aligned}$$

Доказательство. Утверждение доказывается аналогично [20, лемма 4.2].

Лемма 4. Пусть $n \in \mathbb{N}, f \in L_1[0, 1]^n$ и ряд Фурье функции f абсолютно сходится. Тогда

$$T_{m,v}(f) = \hat{f}(0, \dots, 0) + \sum_{l=1}^n \sum_{\substack{k_1+v_1+\dots+k_l+v_l+\dots+v_n=m, \\ k_{l+1}=\dots=k_n=0, \quad k_j=0, \dots, m}} \prod_{j=1}^{l-1} \frac{(-1)^{\text{sign } k_j}}{(p-1)^{1-\text{sign } k_j}} \sum_{\beta_1=1}^{p-1} \dots \sum_{\beta_{l-1}=1}^{p-1} \sum_{\substack{r \in \mathbb{Z}^l, \\ r_l \neq 0}} \hat{f}(p^{k_1-1}(pr_1 + \text{sign } k_1), \dots, p^{k_{l-1}-1}(pr_{l-1} + \text{sign } k_{l-1}), p^{k_l}r_l, 0, \dots, 0).$$

Доказательство. Утверждение доказывается аналогично [20, теорема 3.1].

Рассмотрим функционал

$$P_{m,v}(\mu; f) = T_{m,v} \left(f e^{2\pi i \left(\frac{r}{p^{k+v}} \right)} \right) = \frac{1}{p^m} \sum_{\substack{k_1+v_1+\dots+k_n+v_n=m, \\ k_j=0, \dots, m}} \sum_{r_1=0}^{p^{k_1+v_1}-1} \dots \sum_{r_n=0}^{p^{k_n+v_n}-1} \prod_{j=1}^{n-1} \frac{e^{2\pi i \left(\frac{2lr_j}{p} + \text{sign } k_j \right)}}{(p-1)^{1-\text{sign } k_j}} \times f \left(\frac{r_1}{p^{k_1+v_1}}, \dots, \frac{r_n}{p^{k_n+v_n}} \right) e^{2\pi i \left(\frac{r_1 \mu_1}{p^{k_1+v_1}} + \dots + \frac{r_n \mu_n}{p^{k_n+v_n}} \right)},$$

где $p > 1$ – простое, $m \in \mathbb{N}, v = (v_1, \dots, v_n) \in \mathbb{Z}_+^n, v_1 + \dots + v_n \leq m, \mu \in \rho(v)$.

Из Леммы 4 следует выражение для функционала $P_{m,v}(\mu; f)$ в терминах коэффициентов Фурье (см. ниже).

Следствие 2. Пусть $p > 1$ – простое, $m, n \in \mathbb{N}, v = (v_1, \dots, v_n) \in \mathbb{Z}_+^n, v_1 + \dots + v_n \leq m, \mu \in \rho(v)$. Тогда

$$P_{m,v}(\mu; f) = \hat{f}(\mu_1, \dots, \mu_n) + \sum_{l=1}^n \sum_{\substack{k_1+v_1+\dots+k_l+v_l+\dots+v_n=m, \\ k_j=0, \dots, m, \quad j=1, \dots, l}} \prod_{j=1}^{l-1} \frac{(-1)^{\text{sign } k_j}}{(p-1)^{1-\text{sign } k_j}} \sum_{\beta_1=1}^{p-1} \dots \sum_{\beta_{l-1}=1}^{p-1} \sum_{\substack{r \in \mathbb{Z}^l, \\ r_l \neq 0}} \hat{f}(p^{k_1+v_1-1}(pr_1 + \beta_1 \text{sign } k_1) + \mu_1, \dots, p^{k_{l-1}+v_{l-1}-1}(pr_{l-1} + \beta_{l-1} \text{sign } k_{l-1}) + \mu_{l-1}, p^{k_l+v_l}r_l + \mu_l, \mu_{l+1}, \dots, \mu_n).$$

Следствие 3. Пусть $p > 1$ – простое, $m, n \in \mathbb{N}$. И пусть $S_{G_m}(x) = S_{G_m}(x; f)$ – полином со спектром из ступенчатого гиперболического креста. Тогда

$$P_{m,v}(\mu; S_{G_m}) = \hat{f}(\mu), \quad \mu \in \rho(v), \quad v_1 + \dots + v_n \leq m.$$

Доказательство. Из следствия 2 следует

$$P_{m,v}(\mu; S_{G_m}) = \hat{f}(\mu_1, \dots, \mu_n) + R_m.$$

Однако множество ступенчатого гиперболического креста G_m и множество индексов коэффициентов Фурье, входящих в R_m , не имеют общих точек, т.е. коэффициенты Фурье функции $S_{G_m} \widehat{S_{G_m}}(\mu) = 0$, где μ взяты из R_m . Из чего следует $R_m = 0$.

Теорема 1. Пусть $p > 1$ – простое, $m, n \in \mathbb{N}, v = (v_1, \dots, v_n) \in \mathbb{Z}_+^n, |v|_1 = v_1 + \dots + v_n \leq m, \mu \in \rho(v)$.

Если $1 \leq q \leq \infty, \alpha > 1/\tilde{q} = \max(1/q, 1/2), f \in W_q^\alpha$. Тогда

$$\sup_{\|f\|_{W_q^\alpha}=1} |\hat{f}(\mu_1, \dots, \mu_n) - P_{m,v}(\mu; f)| \leq c \frac{(\max(m - |v|_1, 1))^{\frac{n-1}{\tilde{q}}}}{p^{\alpha m}}. \quad (1)$$

В случае $2 \leq q < \infty$:

$$\sup_{\|f\|_{W_q^\alpha}=1} |\hat{f}(\mu_1, \dots, \mu_n) - P_{m,v}(\mu; f)| \asymp \frac{(\max(m - |v|_1, 1))^{\frac{n-1}{2}}}{p^{\alpha m}}, \quad (2)$$

где $A \asymp B$ означает $c_1 B \leq A \leq c_2 B$.

Определим функцию

Теорема 2. Пусть $p > 1$ – простое, $m, n \in \mathbb{N}$, $N \asymp m^{n-1} p^m$ – количество узлов сетки в $F_m(x; f)$.

Если $1 < q \leq 2 < \tau < \infty, \alpha > 1/q$, то

$$F_m(x; f) = \sum_{\substack{v_1 + \dots + v_n \leq m, \\ v_i \in \mathbb{Z}_+, i=1, \dots, n}} \sum_{\mu \in \rho(v)} P_{m,v}(\mu; f) e^{2\pi i(\mu, x)}. \quad (3)$$

Примечание. Идея поставить вместо коэффициентов Фурье частичной суммы функционал, основанный на значениях функции в некоторых точках, не новая. Альтернативным методом к формуле $F_m(x; f)$ является метод решеток [21] и приложение его к быстрому преобразованию Фурье [22, глава 8]. В работе [23] были получены верхние оценки погрешности восстанавливающей функции на методе решетках ранга-1 для пространств W_2^α в метрике L_2 , которая близка к оптимальной с вероятностью $1 - 4/|\Gamma(R)|$, $\Gamma(R) = \{k \in \mathbb{Z}^n: \prod_{j=1}^n \max(1, |k_j|) \leq R\}$ – гиперболический крест.

Преимуществом функции $F_m(x; f)$ является простота формулы и то, что по порядку погрешность функции $F_m(x; f)$ совпадает с ортопоперечником.

Доказательство теоремы 1. Воспользуемся следствием 2.

$$|\hat{f}(\mu_1, \dots, \mu_n) - P_{m,v}(\mu; f)| \leq \sum_{k \in H_{m,v}} |\hat{f}(k)|,$$

где

$$H_{m,v,\mu} = \bigcup_{l=1}^n \bigcup_{\substack{k_1 + v_1 + \dots + k_l + v_l + \dots + v_n = m, \\ k_j = 0, \dots, m, j=1, \dots, l}} \{ (p^{k_1 + v_1 - 1} (pr_1 + \beta_1 \operatorname{sign} k_1) + \mu_1, \dots, p^{k_{l-1} + v_{l-1} - 1} (pr_{l-1} + \beta_{l-1} \operatorname{sign} k_{l-1}) + \mu_{l-1}, p^{k_l + v_l} r_l + \mu_l, \mu_{l+1}, \dots, \mu_n) : \mu \in \rho(v), \beta_j = 1, \dots, p-1, j=1, \dots, l-1 \}.$$

Пусть $\bar{t} = \max(|t|, 1)$ и $1 < q \leq \infty$. Применим неравенство Гельдера, неравенство Хаусдорфа-Юнга и свойство пространств Лебега соответственно

$$\begin{aligned} |\hat{f}(\mu_1, \dots, \mu_n) - P_{m,v}(\mu; f)| &\leq \sum_{k \in H_{m,v,\mu}} \frac{\bar{k}^\alpha |\hat{f}(k)|}{\bar{k}^\alpha} \leq \left(\sum_{k \in H_{m,v,\mu}} \frac{1}{\bar{k}^{\alpha \tilde{q}}} \right)^{1/\tilde{q}} \left(\sum_{k \in H_{m,v,\mu}} (\bar{k}^\alpha |\hat{f}(k)|)^{\tilde{q}'} \right)^{1/\tilde{q}'} \\ &\leq \|f^\alpha\|_{L_{\tilde{q}}} \left(\sum_{k \in H_{m,v,\mu}} \frac{1}{\bar{k}^{\alpha \tilde{q}}} \right)^{1/\tilde{q}} \leq \|f^\alpha\|_{L_q} \left(\sum_{k \in H_{m,v,\mu}} \frac{1}{\bar{k}^{\alpha \tilde{q}}} \right)^{1/\tilde{q}} = \|f\|_{W_q^\alpha} \left(\sum_{k \in H_{m,v,\mu}} \frac{1}{\bar{k}^{\alpha \tilde{q}}} \right)^{1/\tilde{q}}. \end{aligned}$$

Оценим

$$\left(\sum_{k \in H_{m,v,\mu}} \frac{1}{\bar{k}^{\alpha \tilde{q}}} \right)^{1/\tilde{q}}.$$

Рассмотрим отдельно

$$\sum_{\beta_j=1}^{p-1} \overline{p^{k_j+v_j-1}(pr_j + \beta_j \operatorname{sign} k_j) + \mu_j}^{-\alpha \tilde{q}}, \quad j = 1, \dots, l-1.$$

Учитывая, что $p^{v_j-1} \leq |\mu_j| < p^{v_j}$, $j = 1, \dots, n$, получаем

$$\sum_{\beta_j=1}^{p-1} \overline{p^{k_j+v_j-1}(pr_j + \beta_j \operatorname{sign} k_j) + \mu_j}^{-\alpha \tilde{q}} \leq \sum_{\beta_j=1}^{p-1} \overline{p^{k_j+v_j-1}(pr_j + \beta_j \operatorname{sign} k_j) - p^{v_j}}^{-\alpha \tilde{q}} \leq (p^{k_j+v_j-2} \bar{r}_j)^{-\alpha \tilde{q}}.$$

Отсюда получаем

$$\left(\sum_{k \in H_{m,v,\mu}} \frac{1}{\bar{k}^{\alpha \tilde{q}}} \right)^{1/\tilde{q}} \leq c \frac{1}{p^{\alpha m}} \left(\sum_{l=1}^n \sum_{\substack{k_1+v_1+\dots+k_l+v_l+\dots+v_n=m, \\ k_j=0,\dots,m, \quad j=1,\dots,l}} \sum_{r \in \mathbb{Z}^l} \bar{r}^{-\alpha \tilde{q}} \right)^{1/\tilde{q}}.$$

Ряд $\sum_{r \in \mathbb{Z}^l} \bar{r}^{-\alpha \tilde{q}}$ ввиду условия $\alpha \tilde{q} > 1$ сходится и ввиду оценки [24]

$$\sum_{\substack{z_1+\dots+z_n=m, \\ z_j=0,\dots,m, \quad j=1,\dots,n}} 1 = \binom{m+n-1}{n-1} \asymp m^{n-1},$$

где $A \asymp B$ означает $c_1 B \leq A \leq c_2 B$, получаем

$$\left(\sum_{k \in H_{m,v,\mu}} \frac{1}{\bar{k}^{\alpha \tilde{q}}} \right)^{1/\tilde{q}} \leq c \frac{(\max(m - |v|_1, 1))^{\frac{n-1}{\tilde{q}}}}{p^{\alpha m}}.$$

Для случая $q = 1$ применим вложение $W_1^\alpha \hookrightarrow E^\alpha$, $\alpha > 0$, где E^α – пространство Коробова.

$$\begin{aligned} |\hat{f}(\mu_1, \dots, \mu_n) - P_{m,v}(\mu; f)| &\leq \sum_{k \in H_{m,v,\mu}} \frac{\bar{k}^\alpha |\hat{f}(k)|}{\bar{k}^\alpha} \leq \sup_{k \in \mathbb{Z}^n} \bar{k}^\alpha |\hat{f}(k)| \sum_{k \in H_{m,v,\mu}} \frac{1}{\bar{k}^\alpha} = \|f\|_{E^\alpha} \sum_{k \in H_{m,v,\mu}} \frac{1}{\bar{k}^\alpha} \\ &\leq \|f\|_{W_1^\alpha} \sum_{k \in H_{m,v,\mu}} \frac{1}{\bar{k}^\alpha} \leq c \frac{(\max(m - |v|_1, 1))^{n-1}}{p^{\alpha m}} \|f\|_{W_1^\alpha}. \end{aligned}$$

Объединяя полученные оценки, выполняется неравенство (1).

Пусть $2 \leq q < \infty$. Докажем (2). Оценка сверху следует из (1). Докажем оценку снизу.

Введем пробную функцию

$$f_0(x) = \frac{1}{p^{\alpha m} m - |v|_1^{(n-1)/2}} \sum_{\substack{k_1+v_1+\dots+k_n+v_n=m+1, \\ k_j=0,\dots,m}} e^{2\pi i \sum_{j=1}^n p^{k_j+v_j} j x_j},$$

где $\overline{m - |v|_1} = \max(m - |v|_1, 1)$.

Применим неравенства Литтлвуда-Пэли к норме

$$\|f_0\|_{W_q^\alpha} \asymp \left\| \left(\sum_{k \in \mathbb{Z}_+^n} \left| \sum_{\mu \in \rho(k)} \bar{\mu}^\alpha \hat{f}_0(\mu) e^{2\pi i(\mu, \cdot)} \right|^2 \right)^{1/2} \right\|_{L_q} \asymp 1.$$

Применяя теорему 1 и вложение $W_q^\alpha \hookrightarrow W_2^{\alpha-\varepsilon}$ где $\varepsilon = 1/q - 1/2, 1 < q \leq 2$, получим

$$\begin{aligned} I_2 &\leq c \left(\sum_{\substack{v_1+\dots+v_n \leq m, \\ v_j=0,\dots,m}} \sum_{\mu \in \rho(v)} \bar{\mu}^{\frac{2}{\tau}-1} \frac{\overline{m - |v|_1}^{n-1}}{p^{2(\alpha-\varepsilon)m}} \sum_{k \in H_{m,v,\mu}} (\bar{k}^\alpha |\hat{f}(k)|)^2 \right)^{1/2} \leq c \frac{1}{p^{(\alpha-\frac{1}{q}+\frac{1}{\tau})m}} \|f\|_{W_2^{\alpha-\varepsilon}} \\ &\leq c \frac{1}{p^{(\alpha-\frac{1}{q}+\frac{1}{\tau})m}} \|f\|_{W_q^\alpha} \asymp \left(\frac{\log^{n-1} N}{N} \right)^{\alpha-1/q+1/\tau} \|f\|_{W_q^\alpha}. \end{aligned}$$

Для I_1 (см. теорему А):

$$I_1 \asymp \left(\frac{\log^{n-1} N}{N} \right)^{\alpha-1/q+1/\tau} \|f\|_{W_q^\alpha}.$$

Для I_2 оценка снизу получается ввиду того, что $F_m(x; f)$ – полином со спектром из ступенчатого гиперболического креста, т.е.

$$\sup_{\|f\|_{W_q^\alpha}=1} \|f - F_m(\cdot; f)\|_{L_\tau} \geq E_{G_m}(W_q^\alpha)_\tau = \inf_{t \in T(G_m)} \sup_{\|f\|_{W_q^\alpha}=1} \|f - t\|_{L_\tau},$$

где $T(G_m)$ – множество полиномов со спектром из ступенчатого гиперболического креста. Но ввиду

$$E_{G_m}(W_q^\alpha)_\tau \asymp \frac{1}{p^{(\alpha-\frac{1}{q}+\frac{1}{\tau})m}} \asymp \left(\frac{\log^{n-1} N}{N} \right)^{\alpha-1/q+1/\tau}$$

получаем

$$\sup_{\|f\|_{W_q^\alpha}=1} \|f - F_m(\cdot; f)\|_{L_\tau} \geq \left(\frac{\log^{n-1} N}{N} \right)^{\alpha-1/q+1/\tau}.$$

Теорема доказана.

Заключение

В данной статье была приведена новая восстанавливающая функция для пространств Соболева с доминирующей смешанной производной. Было доказано, что восстанавливающая функция точна для полиномов со спектром из соответствующего ступенчатого гиперболического креста. Были получены оценки погрешности восстанавливающей функции, имеющей порядок ортопоперечника, что является оптимальным для пространств Соболева с доминирующей смешанной производной.

Информация о финансировании

Данная статья финансирована Комитетом науки Министерства науки и высшего образования Республики Казахстан (грант №АР 23488613).

ЛИТЕРАТУРА

- 1 Темляков В.Н. Приближение функций с ограниченной смешанной производной // Труды Математического института имени В.А. Стеклова. – 1986. – Т. 178. – С. 3–113.
- 2 Темляков В.Н. Поперечники некоторых классов функций нескольких переменных // Докл. АН СССР. – 1982. – Т. 267. – № 2. – С. 314–317.
- 3 Темляков В.Н. Приближение периодических функций нескольких переменных тригонометрическими полиномами и поперечники некоторых классов функций // Изв. АН СССР. Сер. матем. – 1985. – Т. 49, №5. – С. 986–1030. <https://doi.org/10.1070/IM1986v027n02ABEH001179>.
- 4 Темляков В.Н. Оценки асимптотических характеристик классов функций с ограниченной смешанной производной или разностью // Сборник трудов Всесоюзной школы по теории функций. – Тр. МИАН СССР – 1989. – Т. 189. – С. 138–168.
- 5 Темляков В.Н. Об оценках ε -энтропии и поперечников классов функций с ограниченной смешанной производной или разностью // Докл. АН СССР. – 1988. – Т. 301. – С. 288–291.
- 6 Галеев Э.М., Приближение суммами Фурье классов функций с несколькими ограниченными производными // Матем. заметки. – 1978. – Т. 23. – № 2. – С. 197–212. <https://doi.org/10.1007/BF01153149>.
- 7 Никольская Н.С. Приближение дифференцируемых функций многих переменных суммами Фурье в метрике L_p // Сиб. матем. журн. – 1974. – Т. 15. – № 2. – С. 395–412. <https://doi.org/10.1007/BF00968291>.
- 8 Byrenheid G. and Ullrich T. Optimal sampling recovery of mixed order Sobolev embeddings via discrete Littlewood–Paley type characterizations // Anal Math. – 2017. – Vol. 43. – P. 133–191. <https://doi.org/10.1007/s10476-017-0303-5>.
- 9 Dinh Dung and Ullrich T. Lower bounds for the integration error for multivariate functions with mixed smoothness and optimal Fibonacci cubature for functions on the square // Math. Nachr. – 2015. – Vol. 288. – P. 743–762. <https://doi.org/10.1002/mana.201400048>.
- 10 Krieg D. et al. Sampling recovery in L_2 and other norms // arXiv preprint arXiv:2305.07539. – 2023.
- 11 Kolomoitsev Y., Lomako T. and Tikhonov S., Sparse grid approximation in weighted Wiener spaces // J. Fourier Anal. Appl. – 2023. – Vol. 29. – No. 2. – P. 19. <https://doi.org/10.1007/s00041-023-09994-2>.
- 12 Byrenheid G., Sparse representation of multivariate functions based on discrete point evaluations : Diss. – Universitäts-und Landesbibliothek Bonn, 2019.
- 13 Kolomoitsev Y. Approximation by quasi-interpolation operators and Smolyak's algorithm // J. Complexity. – 2022. – Vol. 69. – P. 101601. <https://doi.org/10.1016/j.jco.2021.101601>.
- 14 Dinh Dung B-spline quasi-interpolant representations and sampling recovery of functions with mixed smoothness // J. Complexity. – 2011. – Vol. 27. – P. 541–467. <https://doi.org/10.1016/j.jco.2011.02.004>.
- 15 Triebel H. Bases in function spaces, sampling, discrepancy, numerical integration, European Math. Soc. Publishing House, Zürich, 2010.
- 16 Krieg D. and Ullrich M. Function values are enough for L_2 -approximation // Found. Comput. Math. – 2021. – Vol. 21. – No. 4. – P. 1141–1151. <https://doi.org/10.1007/s10208-020-09481-w>.
- 17 Нурсултанов Е.Д., Тлеуханова Н.Т. О восстановлении мультипликативных преобразований функций из анизотропных пространств // Сибирский математический журнал. – 2014. – Т. 55. – № 3. – С. 592–609. <https://doi.org/10.1134/S0037446614030100>.
- 18 Dũng D., Temlyakov V., Ullrich T. Hyperbolic cross approximation. – Springer, 2018.
- 19 Нурсултанов Е.Д. О мультипликаторах рядов Фурье по тригонометрической системе // Математические заметки. – 1998. – Т. 63. – № 2. – С. 235–247. <https://doi.org/10.4213/mzm1270>.
- 20 Bassarov S., Nursultanov E. New Cubature Formulas for Sobolev Spaces with Dominant Mixed Derivative // J. Math. Sci. – 2025. – Vol. 291. – P. 4–19. <https://doi.org/10.1007/s10958-025-07774-5>.
- 21 Dick J., Kritzer P., Pillichshammer F. Lattice rules. – Springer, Cham, 2022. <https://doi.org/10.1007/978-3-031-09951-9>.
- 22 Plonka G. et al. Numerical fourier analysis. – New York : Springer International Publishing, 2018. <https://doi.org/10.1007/978-3-031-35005-4>
- 23 Bartel F. et al. On the reconstruction of functions from values at subsampled quadrature points // Math. Comput. – 2024. – Vol. 93. – № 346. – P 785–809.
- 24 Sickel W., Ullrich T. Smolyak's algorithm, sampling on sparse grids and function spaces of dominating mixed smoothness. – Univ., 2006.

REFERENCES

- 1 Temlyakov, V.N. Approximations of functions with bounded mixed derivative, *Proc. Steklov Inst. Math.*, 178, 3–113 (1986).
- 2 Temlyakov, V.N. Widths of some classes of functions of several variables, *Soviet Math Dokl.*, 267(2), 314–317 (1982).
- 3 Temlyakov, V.N. Approximation of periodic functions of several variables by trigonometric polynomials, and widths of some classes of functions, *Math. USSR-Izv.*, 49(5), 986–1030 (1985). <https://doi.org/10.1070/IM1986v027n02ABEH001179>.
- 4 Temlyakov, V.N. Estimates for the asymptotic characteristics of classes of functions with bounded mixed derivative or difference, *Proc. Steklov Inst. Math.*, 189, 138–168 (1989).
- 5 Temlyakov, V.N. On Estimates of ε -Entropy and Widths of Classes of Functions with Bounded Mixed Derivative or Difference, *Soviet Math. Dokl.*, 301, 288–291 (1988).
- 6 Galeev, A.M., Approximation by Fourier sums of classes of functions with several bounded derivatives, *Math. Notes*, 23(2), 109–117 (1978),). <https://doi.org/10.1007/BF01153149>. (in Russian).
- 7 Nikolskaya, N.S. Approximation of differentiable functions of several variables by Fourier sums in the L_p metric. *Siberian Math. J.*, 15 (2), 282–295 (1974). <https://doi.org/10.1007/BF00968291>. (in Russian).
- 8 Byrenheid, G. and Ullrich, T. Optimal sampling recovery of mixed order Sobolev embeddings via discrete Littlewood–Paley type characterizations, *Anal Math.*, 43, 133–191 (2017). <https://doi.org/10.1007/s10476-017-0303-5>
- 9 Dinh Dung and Ullrich T., Lower bounds for the integration error for multivariate functions with mixed smoothness and optimal Fibonacci cubature for functions on the square, *Math. Nachr.*, 288, 743–762 (2015). <https://doi.org/10.1002/mana.201400048>.
- 10 Krieg, D. et al. Sampling recovery in L_2 and other norms, *arXiv preprint arXiv:2305.07539*, 2023.
- 11 Kolomoitsev, Y., Lomako, T. and Tikhonov, S. Sparse grid approximation in weighted Wiener spaces, *J. Fourier Anal. Appl.*, 29 (2), 1–19 (2023). <https://doi.org/10.1007/s00041-023-09994-2>.
- 12 Byrenheid, G., *Sparse Representation of Multivariate Functions Based on Discrete Point Evaluations*, Universitäts-und Landesbibliothek Bonn, 2019.
- 13 Kolomoitsev, Y. Approximation by quasi-interpolation operators and Smolyak's algorithm, *J. Complexity*, 69, 101601 (2022). <https://doi.org/10.1016/j.jco.2021.101601>.
- 14 Dinh Dung B-spline quasi-interpolant representations and sampling recovery of functions with mixed smoothness, *J. Complexity*, 27, 541–467 (2011). <https://doi.org/10.1016/j.jco.2011.02.004>.
- 15 Triebel, H. *Bases in Function Spaces, Sampling, Discrepancy, Numerical Integration*, European Math. Soc. Publishing House, Zürich, 2010.
- 16 Krieg, D. and Ullrich, M. Function values are enough for L_2 -approximation *Found. Comput. Math.*, 21(4), 1141–1151 (2021). <https://doi.org/10.1007/s10208-020-09481-w>.
- 17 Nursultanov, E.D., Tleukhanova, N.T. On the reconstruction of multiplicative transformations of functions from anisotropic spaces, *Siberian Math. J.*, 55(3), 482–497 (2014). <https://doi.org/10.1134/S0037446614030100> (in Russian).
- 18 Dũng, D., Temlyakov, V., Ullrich, T. *Hyperbolic Cross Approximation*. Springer, 2018.
- 19 Nursultanov, E.D. On the multipliers of Fourier series with respect to the trigonometric system, *Math. Notes*, 63(2), 205–214 (1998). <https://doi.org/10.4213/mzm1270>.
- 20 Bassarov, S., Nursultanov, E. New Cubature Formulas for Sobolev Spaces with Dominant Mixed Derivative. *J Math Sci.*, 291, 4–19 (2025). <https://doi.org/10.1007/s10958-025-07774-5>
- 21 Dick, J., Kritzer, P., Pillichshammer, F. *Lattice Rules*. Springer. Cham. 2022. <https://doi.org/10.1007/978-3-031-09951-9>.
- 22 Plonka, G. et al. *Numerical Fourier Analysis*. Springer International Publishing. New York. 2018. <https://doi.org/10.1007/978-3-031-35005-4>.
- 23 Bartel, F. et al. On the reconstruction of functions from values at subsampled quadrature points, *Math. Comput.*, 93 (346), 785–809 (2024).
- 24 Sickel, W., Ullrich, T. Smolyak's algorithm, sampling on sparse grids and function spaces of dominating mixed smoothness. *Univ.*, 2006.

¹*Басаров С.Ж.,

магистр, к.ғ.к., ORCID ID: 0000-0001-5422-2424,

*e-mail: bassarov.serzhan98@gmail.com

²Нурсултанов Е.Д.,

ф.-м.ғ.д., профессор, ORCID ID: 0000-0003-3879-2261,

e-mail: er-nurs@yandex.kz

¹«Geometry» ЖШС, Астана қ., Қазақстан²М.В. Ломоносов атындағы Мәскеу мемлекеттік университетінің

Қазақстан филиалы, Астана қ., Қазақстан

СОБОЛЕВ КЕҢІСТІКТЕРІ ҮШІН БАСЫМ АРАЛАС ТУЫНДЫСЫ БАР ҚАЛПЫНА КЕЛТІРЕТІН ФУНКЦИЯ

Аңдатпа

Сызықтық аппроксимация – бұл белгілі бір кластағы функцияны осы кластың бекітілген ақырлы өлшемді ішкі кеңістігінің элементтерімен жуықтау. Мысалы, бір айнымалылы периодты функциялар үшін мұндай элементтер ретінде тригонометриялық көпмүшелер алынады. Бізді қызықтыратын көпөлшемді жағдайда, яғни әр айнымалы бойынша периодты функциялар үшін, аталған ішкі кеңістік рөлін жиілігі баспалдақты гиперболалық крестен тұратын тригонометриялық көпмүшелер жиыны атқарады. Дегенмен, бұл көпмүшелердің коэффициенттерін таңдау мәселесі туындайды. Бұл мақалада берілген нүктелер бойынша басым аралас туындысы бар Соболев кеңістігіндегі функцияларды қалпына келтіру аппараты ұсынылып, қалпына келтіру қателігіне бағалаулар алынды. Әдіс қалпына келтіруші функцияны жиілігі баспалдақты гиперболалық креске жататын, ал коэффициенттері берілген нүктелер арқылы есептелетін көпмүше түрінде құруға негізделген. Жуықтау қателігінің реті ортогональ көлденеңмен шамалас, бұл – мұндай көпмүшелер үшін оңтайлы нәтиже. Ұсынылған әдіс жиілігі баспалдақты гиперболалық креске жататын көпмүшелер үшін дәл нәтиже береді. Сондай-ақ, аталған кеңістіктердегі функциялардың Фурье коэффициенттерін қалпына келтіретін функционал алынған. Қалпына келтіру функциясының нақты өрнегінің арқасында алынған формуланы қолданбалы есептерді шешуде пайдалануға болады.

Тірек сөздер: көп айнымалы, функциялар теориясы, Соболев кеңістігі, аралас туынды, гиперболалық жиілік.

¹*Bassarov S.Zh.,

MSc., j.r.a., ORCID ID: 0000-0001-5422-2424,

*e-mail: bassarov.serzhan98@gmail.com

²Nursultanov E.D.,

Dr. Phys.-Math. Sc., Professor, ORCID ID: 0000-0003-3879-2261,

e-mail: er-nurs@yandex.kz

¹«Geometry» LLP, Astana, Kazakhstan²Kazakhstan Branch of M. V. Lomonosov Moscow State University Branch,
Astana, Kazakhstan

RECOVERY FUNCTION FOR SOBOLEV SPACES WITH DOMINATING MIXED DERIVATIVE

Abstract

Linear approximation is the approximation of a function from a certain class by elements of a fixed finite-dimensional subspace of that same class. For instance, for one-dimensional periodic functions, such elements are trigonometric polynomials. In the multidimensional case of interest, for functions periodic in each variable, this subspace is the set of trigonometric polynomials with a spectrum from a step hyperbolic cross. However, the question of selecting coefficients for these polynomials arises. This paper presents an apparatus for recovering functions from

Sobolev spaces with a dominating mixed derivative based on given points, and establishes error estimates for the recovery. The method is based on constructing a recovery function in the form of a polynomial with a spectrum from a step hyperbolic cross, where the coefficients are calculated using the given points. The approximation error is of the order of the orthowidth, which is an optimal result for such polynomials. The proposed method is exact for polynomials with a spectrum from a step hyperbolic cross. Additionally, a functional that recovers the Fourier coefficients for functions from the indicated spaces is derived. Due to the explicit expression of the recovery function, the obtained formula can be used to solve applied problems.

Keywords: multiple dimension, function theory, Sobolev spaces, mixed derivative, hyperbolic cross.

Дата поступления статьи в редакцию: 08.10.2025

ӨОЖ 517.928
FTAXP 27.31.21

<https://doi.org/10.55452/1998-6688-2025-22-4-324-339>

^{1*}**Төлеубай А.М.,**
PhD, ORCID ID: 0000-0002-5490-8212,
*e-mail: altyn.15.94@mail.ru
²**Кәріпова А.Т.,**
докторант, ORCID ID: 0009-0008-7837-1801,
e-mail: karipova-aidana@mail.ru

¹«Шәкәрім университеті» КеАҚ, Семей қ., Қазақстан
²Л.Н. Гумилев атындағы Еуразия ұлттық университеті, Астана қ., Қазақстан

ПЕРФОРАЦИЯЛАНҒАН ОБЛЫСТА ЧЭФИ-ИНФАНТЕ ТЕҢДЕУЛЕРІНІҢ АТТРАКТОРЛАРЫНЫҢ ОРТАШАЛАНУЫ

Аңдатпа

Бұл жұмыста микро біртекті емес ортада Чэфи-Инфанте теңдеуінің орташалау есебі қарастырылады. Зерттеу барысында жылдам тербелмелі мүшелері мен диссипациясы бар Чэфи-Инфанте теңдеулері жан-жақты талданады. Модельге кіші параметр енгізіліп, оның шекті мәндеріне байланысты орташаланған (шекте) есеп алынады. Бастапқы теңдеудің траекториялық аттракторларының орташаланған теңдеудің аттракторына жинақталатыны дәлелденіп, тиісті теорема тұжырымдалады. Сонымен қатар, бірегей шешімдер жағдайында глобалды аттракторлардың жинақталуы зерттеледі, ол үшін теңдеудің сызықтық емес мүшелеріне қосымша шарттар қойылады. Жұмыста перфорацияланған облыстардағы және олардың қуыстарының шекараларындағы күрделі шекаралық шарттары бар сингулярлы есептер де қарастырылады. Мұндай жағдайда бастапқы теңдеуден құрылымы өзгеше орташаланған (шекте) теңдеу алынады. Бұл теңдеу жүйенің тиімді орташа сипаттамасын бейнелейді және оның құрамында қосымша потенциалдық мүше пайда болуы мүмкін. Кіші параметр нөлге ұмтылған кезде аттракторлардың Хаусдорф мағынасындағы жинақталуы дәлелденеді. Жалпы нәтиже ретінде, орташаланған аттракторлар сипатталып, бастапқы жүйенің аттракторларының шекте теңдеудің аттракторларына жинақталатыны көрсетіледі. Бұл жұмыста периодты перфорацияланған ортадағы Чэфи-Инфанте теңдеулерін орташалау мәселесі алғаш рет талданып, нәтижелердің ғылыми жаңалығы мен өзектілігі негізделеді. Бұл зерттеу дифференциалдық операторларды орташалау теориясының әдістерін дамытуға және сызықтық емес дифференциалдық теңдеулердің шешімдерін зерттеуде қолданыс табуы мүмкін.

Тірек сөздер: Чэфи-Инфанте теңдеулері, кеуекті орта, аттракторлар, перфорацияланған облыс, орташалау.

Кіріспе

Перфорацияланған материалдар мен кеуекті орталардағы процестерді математикалық модельдеу – азаматтық, авиациялық және ғарыштық инженерияда, сондай-ақ жаңа материалдар өндіру саласында маңызды ғылыми мәселе. Мұндай модельдер, әдетте, перфорацияланған облыстарда қойылған шекаралық есептердің асимптотикалық талдауына алып келеді. Бұл тәсілді қолдану үшін алдымен орташаланған (яғни, шекте) есепті құрып, шешімдердің жинақталу теоремаларын дәлелдеу қажет. Егер есептің шешімі бірегей болмаса, мәселе күрделене түседі – бұл жағдайда траекториялық аттракторлардың жинақталуын дәлелдеу қажет (толығырақ [1–4] қараңыз).

Мақалада перфорацияланған облыстар мен олардың қуыстарының шекарасында берілген шекаралық шарттары бар сингулярлы қиын есептер зерттеледі [5–7]. Мұндай жағдайда бастапқы теңдеуге қарағанда құрылымы өзгеше болатын орташаланған (шекте) теңдеу алы-

нады. Бұл теңдеу модельдің тиімді (орташа) сипаттамасын бейнелейді және оның құрамында қосымша потенциалдық мүше пайда болуы мүмкін. Осы шағын параметр нөлге ұмтылған кезде аттракторлардың Хаусдорф мағынасындағы жинақталуы дәлелденеді. Нәтижесінде, орташаланған аттракторлар анықталып, бастапқы жүйенің аттракторларының шектік (орташаланған) теңдеудің аттракторларына жинақталуы көрсетіледі.

Жұмыста перфорацияланған облыстардағы Чэфи-Инфанте теңдеулері үшін аралас (бастапқы-шекаралық) есеп қарастырылады. Соңғы жылдары кеуекті орталарда қойылған есептердің орташалануы мен асимптоталық талдауына арналған көптеген зерттеулер жарық көрді (мысалы, [8–11]). Периодты, локальды периодты, квази-периодты және кездейсоқ құрылымдар үшін әртүрлі орташалануы теоремалары дәлелденген. Бұл бағыттағы кешенді шолулар мен библиографияны [11–14] еңбектерінен табуға болады.

Аттракторлар теориясы туралы жан-жақты мәліметтер [14, 20–22] монографияларында және ондағы сілтемелерде ұсынылған. Аттракторларды орташалау мәселелері [15, 18–22] еңбектерінде, сондай-ақ [18, 19] жұмыстарында зерттелген. [23, 24] параболалық теңдеулер үшін орташалау әдістері жан-жақты талданып, олардың Чэфи-Инфанте типті теңдеуге қолданылу мүмкіндіктері көрсетілген. Бұған дейін Чэфи-Инфанте типті теңдеулер үшін орташалау тек классикалық біртекті облыстарда қарастырылған болатын. Бұл жұмыста алғаш рет перфорацияланған орта жағдайында орташаланған траекториялық аттракторлардың бар болуы және олардың шектік аттракторға жинақталуы дәлелденді.

Мақалада ε параметріне тәуелді перфорацияланған облыстағы Чэфи-Инфанте теңдеуі үшін анықталған траекториялық аттрактор $\mathcal{A}_\varepsilon$ -ның $\varepsilon \rightarrow 0$ кезінде әлсіз мағынада тиісті функционалдық кеңістіктегі орташаланған теңдеудің траекториялық аттракторына $\mathcal{A}$ жинақталатыны дәлелденеді. Мұндағы ε – перфорацияланған ортадағы қуыстардың диаметрі мен олардың арасындағы арақашықтықты сипаттайтын шағын параметр. Сонымен қатар, теңдеу коэффициенттеріне белгілі бір қосымша шарттар орындалған жағдайда глобалдық аттракторлардың да жинақталуы көрсетіледі. Сондай-ақ, шектік теңдеуде пайда болатын «бөгде мүше» – потенциалдық мүше – туралы айтылады (бұл құбылыс алғаш рет [1, 2] еңбектерінде зерттелген).

Материалдар мен әдістер

Зерттеу жұмысының бұл бөлімінде автономды эволюциялық теңдеулердің траекториялық аттракторлары туралы негізгі ұғымдар анықталады және теоремалар тұжырымдалады.

Бұл бөлімде автономды эволюциялық теңдеулердің траекториялық аттракторларын құрудың жалпы сызбасы көрсетілген. Келесі нәтижелерді алу үшін бұл нәтижелер перфорацияланған облыста Чэфи-Инфанте теңдеулерінің траекториялық аттракторларын сол теңдеулерде және шекаралық шарттарында тез тербелмелі мүшелері қарастырылады және сәйкес орташаланған теңдеумен зерттеу үшін қолданылады.

Абстракттілі автономды эволюциялық теңдеуді қарастырамыз:

$$\frac{\partial u}{\partial t} = A(u), t \geq 0 \quad (1)$$

Сызықтық емес оператор $A(\cdot): E_1 \rightarrow E_0$ берілген, мұнда E_1, E_0 – Банах кеңістіктері және $E_1 \subseteq E_0$. Мысалы, $A(u) = \Delta u - f(u) + g(\cdot)$ (1.2-ші бөлімді қараңыз).

(1) теңдеуінің шешімі деп L_2 -гі функцияларды атаймыз, егер оларды (1)-ге қойып алынған кезде D' жалпыланған функциялар мағынасында тепе-теңдік болып алынады.

Біз (1) теңдеудің $u(s)$ шешімдерін, толығымен $s \in \mathbb{R}_+$ айнымалының функциялары ретінде зерттейміз. Мұнда $s \equiv t$ уақыттың айнымалысын білдіреді. (1) теңдеудің шешімдерінің жиынын $\mathcal{K}^+$ траекторияларының кеңістігі деп атаймыз. $\mathcal{K}^+$ траекторияларының кеңістігін толығырақ сипаттайық.

Ең алдымен, $\mathbb{R}$ -гі бекітілген $[t_1, t_2]$ уақыт аралығында анықталған (1) теңдеудің $u(s)$ шешімдерін қарастырамыз.

t_1 мен t_2 -ге тәуелді, $\mathcal{F}_{t_1, t_2}$ Банах кеңістігіне жататын (1) теңдеуінің шешімдерін зерттейміз. $\mathcal{F}_{t_1, t_2}$ кеңістігі $f(s)$, $s \in [t_1, t_2]$ функцияларынан тұрады, мысалы, $f(s) \in E$ барлық дерлік үшін $s \in [t_1, t_2]$, мұндағы E – Банах кеңістігі. $E_1 \subseteq E \subseteq E_0$ деп саналады.

Мысалы, $\mathcal{F}_{t_1, t_2}$ кеңістігі $C([t_1, t_2]; E)$ кеңістігі немесе $L_p(t_1, t_2; E)$ кеңістігі болуы мүмкін, $p \in [1, \infty]$ үшін, немесе осындай кеңістіктердің қиылысуы (1.2 бөлімді қараңыз [25]).

$$\Pi_{t_1, t_2} \mathcal{F}_{\tau_1, \tau_2} \subseteq \mathcal{F}_{t_1, t_2}$$

және

$$\|\Pi_{t_1, t_2} f\|_{\mathcal{F}_{t_1, t_2}} \leq C(t_1, t_2, \tau_1, \tau_2) \|f\|_{\mathcal{F}_{\tau_1, \tau_2}}, \quad \forall f \in \mathcal{F}_{\tau_1, \tau_2}, \quad (2)$$

мұнда $[t_1, t_2] \subseteq [\tau_1, \tau_2]$ және $\Pi_{t_1, t_2} [t_1, t_2]$ кесіндісіндегі оператордың сығылуын білдіреді. $C(t_1, t_2, \tau_1, \tau_2)$ тұрақтысы f функциясына тәуелді емес. Әдетте кеңістіктердің біртекті жағдайы қарастырылады, онда $C(t_1, t_2, \tau_1, \tau_2) = C(t_2 - t_1, \tau_2 - \tau_1)$.

Айталық, $h \in \mathbb{R}$ үшін $S(h)$ жылжыту операторын білдіреді

$$S(h)f(s) = f(h + s).$$

Егер $f(\cdot)$ функцияның s аргументі $[t_1, t_2]$ кесіндісінде жататын болса, онда $h \in \mathbb{R}$ үшін $S(h)f(\cdot)$ функциясының s аргументі $[t_1 - h, t_2 - h]$ кесіндісінде жатады. Айталық, $S(h)$ бейнелеуі F_{t_1, t_2} – тің $F_{t_1 - h, t_2 - h}$ -гі изоморфизмі болып табылады және

$$\|S(h)f\|_{\mathcal{F}_{t_1 - h, t_2 - h}} = \|f\|_{\mathcal{F}_{t_1, t_2}}, \quad \forall f \in \mathcal{F}_{t_1, t_2}. \quad (3)$$

Бұл тұжырым толығымен орынды, мысалы, біртекті кеңістіктер үшін.

Егер $f(s) \in \mathcal{F}_{t_1, t_2}$, онда $A(f(s)) \in \mathcal{D}_{t_1, t_2}$, мұндағы $\mathcal{F}_{t_1, t_2} \subseteq \mathcal{D}_{t_1, t_2}$. $\frac{\partial f(t)}{\partial t} \in D'((t_1, t_2); E_0)$ мәндері E_0 жататын жалпыланған функция және $\mathcal{D}_{t_1, t_2} \subseteq D'((t_1, t_2); E_0)$ барлық $(t_1, t_2) \subset \mathbb{R}$ үшін орындалсын дейік. Егер $D'((t_1, t_2); E_0)$ кеңістігіндегі $\frac{\partial u}{\partial t}(s) = A(u(s))$ жалпыланған функциялар мағынасында орындалса, онда $u(s) \in \mathcal{F}_{t_1, t_2}$ функциясын (1) теңдеуінің $\mathcal{F}_{t_1, t_2}$ кеңістігіндегі шешімі $((t_1, t_2)$ интервалындағы) деп атаймыз.

Айталық,

$$\mathcal{F}_+^{loc} = \{f(s), s \in \mathbb{R}_+ \mid \Pi_{t_1, t_2} f(s) \in \mathcal{F}_{t_1, t_2}, \quad \forall [t_1, t_2] \subset \mathbb{R}_+\}. \quad (4)$$

Мысалы, егер $\mathcal{F}_{t_1, t_2} = C([t_1, t_2]; E)$, онда $\mathcal{F}_+^{loc} = C(\mathbb{R}_+; E)$, ал егер $\mathcal{F}_{t_1, t_2} = L_p(t_1, t_2; E)$, онда $\mathcal{F}_+^{loc} = L_p^{loc}(\mathbb{R}_+; E)$.

Егер $\Pi_{t_1, t_2} u(s) \in \mathcal{F}_{t_1, t_2}$ және $\Pi_{t_1, t_2} u(s)$ функциясы кез келген $[t_1, t_2] \subset \mathbb{R}_+$ уақыт аралығы үшін (1) теңдеудің шешімі болып табылса, онда $u(s) \in \mathcal{F}_+^{loc}$ функциясы $\mathcal{F}_+^{loc}$ -гі (1) теңдеуінің шешімі деп аталады.

Айталық, $\mathcal{K}^+$ бұл $\mathcal{F}_+^{loc}$ -не тиісті (1) теңдеуінің шешімдерінің кейбір жиыны болсын, $\mathcal{F}_+^{loc}$ -не тиісті болатын, бірақ (1) теңдеуінің барлық шешімдерінің жиыны емес болуы мүмкін. $\mathcal{K}^+$ кеңістігінің өзі бұл (1) теңдеуінің траекторияларының кеңістігі болса, онда $\mathcal{K}^+$ жиынының элементі траектория деп аталады.

$\mathcal{K}^+$ траекторияларының кеңістігі келесі мағынада трансляциялы-инвариантты деп саналады: егер $u(s) \in \mathcal{K}^+$ болса, онда кез келген $h \geq 0$ үшін $u(h + s) \in \mathcal{K}^+$. Бұл шарт біртекті кеңістіктерінде автономды теңдеу шешімдерінің нақты қасиеті болып табылады.

Енді $\mathcal{F}_+^{loc}$ – $\text{deg } S(h)$ жылжыту операторларын қарастырайық:

$$S(h)f(s) = f(s + h), \quad h \geq 0.$$

$\{S(h), h \geq 0\}$ бейнелеуінің жиыны $h_1, h_2 \geq 0$ үшін $\mathcal{F}_+^{loc}$: $S(h_1)S(h_2) = S(h_1 + h_2)$ жартылай группаны бейнелейді және $S(0) = I$, мұнда I тепе-тең бейнелеу болатыны айқын. h -айнымалысын t уақыт айнымалысына ауыстырамыз. $\{S(t), t \geq 0\}$ жартылай группасы жылжытулардың жартылай группасы деп аталады. Жасалған болжамға байланысты жылжытулардың жартылай группасы $\mathcal{K}^+$ траекториясының кеңістігін өз-өзіне бейнелейді:

$$S(t)\mathcal{K}^+ \subseteq \mathcal{K}^+, \quad \forall t \geq 0.$$

Әрі қарай, $\mathcal{K}^+ \subset \mathcal{F}_+^{loc}$ траекториялар кеңістігіне әсер ететін $\{S(t)\}$ жылжытулардың жартылай группасының тартылу қасиеттері зерттеледі. $\mathcal{F}_+^{loc}$ кеңістігінде кейбір топологияны анықтайық.

Айталық, $\rho_{t_1, t_2}(\cdot, \cdot)$ – барлық $[t_1, t_2] \subset \mathbb{R}$ кесінділері үшін $\mathcal{F}_{t_1, t_2}$ кеңістігінде анықталған метрика, төменгі шарттарды қанағаттандырады

$$\rho_{t_1, t_2}(\Pi_{t_1, t_2} f, \Pi_{t_1, t_2} g) \leq D(t_1, t_2, \tau_1, \tau_2) \rho_{\tau_1, \tau_2}(f, g),$$

$$\forall f, g \in \mathcal{F}_{\tau_1, \tau_2}, \quad [t_1, t_2] \subseteq [\tau_1, \tau_2],$$

$$\rho_{t_1 - h, t_2 - h}(S(h)f, S(h)g) = \rho_{t_1, t_2}(f, g),$$

$$\forall f, g \in \mathcal{F}_{t_1, t_2}, \quad [t_1, t_2] \subset \mathbb{R}, \quad h \in \mathbb{R}.$$

Θ_{t_1, t_2} арқылы $\mathcal{F}_{t_1, t_2}$ -гі кеңістігіндегі сәйкес метрикалық кеңістікті белгілейміз. Мысалы, $\rho_{t_1, t_2} - \mathcal{F}_{t_1, t_2}$ Банах кеңістігіндегі $\|\cdot\|_{\mathcal{F}_{t_1, t_2}}$ нормамен туындайтын метрика болуы мүмкін. Сонымен қатар, қосымшада $\mathcal{F}_{t_1, t_2}$ Банах кеңістігіндегі күшті жинақтылыққа қарағанда ρ_{t_1, t_2} метрикасы Θ_{t_1, t_2} топологиясын әлсіз мағынада туындатады.

Кез-келген $[t_1, t_2] \subset \mathbb{R}_+$ үшін Θ_{t_1, t_2} -ны локальды жинақты топологиямен жабдықталған $\mathcal{F}_+^{loc}$ кеңістігін Θ_+^{loc} арқылы белгілейміз. Дәлірек айтсақ, $\{f_k(s)\} \subset \mathcal{F}_+^{loc}$ функцияларының тізбегі анықтама бойынша $k \rightarrow \infty$ Θ_+^{loc} – да $f(s) \in \mathcal{F}_+^{loc}$ функциясына жинақталады, егер $(\rho_{t_1, t_2}(\Pi_{t_1, t_2} f_k, \Pi_{t_1, t_2} f) \rightarrow 0$ болса) $k \rightarrow \infty$ болса, кез келген $[t_1, t_2] \subset \mathbb{R}_+$ кесінді үшін. Мысалы, Θ_+^{loc} топологиясы Фреше метрикасы арқылы өлшенетінін дәлелдеу қиын емес

$$\rho_+(f_1, f_2) := \sum_{m \in \mathbb{N}} 2^{-m} \frac{\rho_{0, m}(f_1, f_2)}{1 + \rho_{0, m}(f_1, f_2)}.$$

Егер барлық Θ_{t_1, t_2} метрикалық кеңістіктері толық болса, онда Θ_+^{loc} метрикалық кеңістігі толық болатыны айқын.

$\{S(t)\}$ жылжытулардың жартылай тобы Θ_+^{loc} топологиясында үзіліссіз болатынына назар аударайық. Бұл Θ_+^{loc} топологиялық кеңістігінің анықтамасынан тікелей туындайды.

Сонымен қатар, келесідей Банах кеңістігін анықтаймыз

$$\mathcal{F}_+^b := \{f(s) \in \mathcal{F}_+^{loc} \mid \|f\|_{\mathcal{F}_+^b} < +\infty\}, \quad (5)$$

мұнда норма

$$\|f\|_{\mathcal{F}_+^b} := \sup_{h \geq 0} \|\Pi_{0,1} f(h+s)\|_{\mathcal{F}_{0,1}}. \quad (6)$$

Мысалы, егер $\mathcal{F}_+^{loc} = C(\mathbb{R}_+; E)$, онда $\mathcal{F}_+^b = C^b(\mathbb{R}_+; E)$ кеңістігі нормасымен $\|f\|_{\mathcal{F}_+^b} = \sup_{h \geq 0} \|f(h)\|_E$, ал егер, $\mathcal{F}_+^{loc} = L_p^{loc}(\mathbb{R}_+; E)$ болса, онда $\mathcal{F}_+^b = L_p^b(\mathbb{R}_+; E)$ нормасымен $\|f\|_{\mathcal{F}_+^b} = \left(\sup_{h \geq 0} \int_h^{h+1} \|f(s)\|_E^p ds \right)^{1/p}$ анықталады.

Келесі $\mathcal{F}_+^b \subseteq \Theta_+^{loc}$ орындалады. $\mathcal{F}_+^b$ Банах кеңістігі $\mathcal{K}^+$ траекториялық кеңістіктің ішіндегі шектеулі жиындарды анықтау үшін қажет. $\mathcal{K}^+$ траекториялық аттракторын салу барысында біз $\mathcal{F}_+^b$ Банах кеңістігінің топологиясындағы сәйкес бірмәнді жинақтылықты қарастырмаймыз. Оның орнына әлдеқайда әлсіз, Θ_+^{loc} локальды жинақтылық топологиясын пайдаланамыз.

$\mathcal{K}^+ \subseteq \mathcal{F}_+^b$, яғни (1) теңдеудің $u(s) \in \mathcal{K}^+$ кез-келген траекториясының ақырлы (8) нормасы бар деп есептейік. $\mathcal{K}^+$ — на әсер ететін $\{S(t)\}$ жылжытулардың жартылай группасының траекториялық аттракторының және тартылатын жиынның анықтамаларын тұжырымдаймыз.

Анықтама 1. $\mathcal{P} \subseteq \Theta_+^{loc}$ жиыны Θ_+^{loc} топологиясында $\mathcal{K}^+$ — на әсер ететін $\{S(t)\}$ жылжытулардың жартылай группасының тартатын жиыны деп аталады, егер Θ_+^{loc} топологиясында $\mathcal{F}_+^b$ -де кез келген $\mathcal{B} \subseteq \mathcal{K}^+$ шектелген жиынында $\mathcal{P}$ жиыны $S(t)\mathcal{B}$, $t \rightarrow +\infty$ тартса, яғни кез-келген $t \geq t_1$ $S(t)\mathcal{B} \subseteq O_\varepsilon(\mathcal{P})$ болатындай Θ_+^{loc} — ғы кез келген $O_\varepsilon(\mathcal{P})$ — ε аймағында $t_1 \geq 0$ табылса. $\mathcal{P}$ — ның тартатын қасиеті келесідей эквивалентті түрде тұжырымдалуы мүмкін:

$\mathcal{F}_+^b$ — де кез келген жиыны $\mathcal{B} \subseteq \mathcal{K}^+$ шенелген және кез келген $M > 0$ үшін

$$\text{dist}_{\Theta_{0,M}}(\Pi_{0,M} S(t)\mathcal{B}, \Pi_{0,M} \mathcal{P}) \rightarrow 0 \quad (t \rightarrow +\infty),$$

мұнда

$$\text{dist}_{\mathcal{M}}(X, Y) := \sup_{x \in X} \text{dist}_{\mathcal{M}}(x, Y) = \sup_{x \in X} \inf_{y \in Y} \rho_{\mathcal{M}}(x, y)$$

$\mathcal{M}$ метрикалық кеңістігіндегі X жиынынан Y жиынына дейінгі Хаусдорфтың жартылай кашықтығын білдіреді.

Анықтама 2. [1, б. 351–359]. Егер

(i) $\mathcal{A}$ жиыны $\mathcal{F}_+^b$ -де шенелген және Θ_+^{loc} -да компактті болса,

(ii) $\mathcal{A}$ жиыны жартылай группалардың жылжытуларына қатысты қатаң инвариантты болса :

$$S(t)\mathcal{A} = \mathcal{A} \quad \text{барлығы үшін } t \geq 0,$$

(iii) $\mathcal{A}$ жиыны Θ_+^{loc} топологиясында $\mathcal{K}^+$ -да, кез келген $M > 0$, $\{S(t)\}$ жартылай группаның жылжытуларының тартатын жиыны болса,

онда $\mathcal{A} \subseteq \mathcal{K}^+$ жиыны Θ_+^{loc} топологиясында $\mathcal{K}^+$ -да $\{S(t)\}$ жартылай группаның жылжытуларының траекториялық аттракторы деп аталады

$$\text{dist}_{\Theta_{0,M}}(\Pi_{0,M} S(t)\mathcal{B}, \Pi_{0,M} \mathcal{A}) \rightarrow 0 \quad (t \rightarrow +\infty).$$

Ескерту 1. [13, б. 635–677] жұмысының терминологиясын пайдаланып, егер $(\mathcal{F}_+^b, \Theta_+^{loc})$ - $\mathcal{K}^+$ -не әсер етуші, $\mathcal{A}$ аттракторы Θ_+^{loc} топологиясында $t \rightarrow +\infty$ кезде $S(t)\mathcal{B}$ жиынын тартса, $\{S(t)\}$ жартылай группаның жылжытуының $\mathcal{A}$ траекториялық аттракторы глобальды болады:

$$\text{dist}_{\Theta_+^{loc}}(S(t)\mathcal{B}, \mathcal{A}) \rightarrow 0 \quad (t \rightarrow +\infty).$$

мұнда $\mathcal{B} \subseteq \mathcal{K}^+$ -гі кез келген шенелген $(\mathcal{F}_+^b, \Theta_+^{loc})$ жиын.

(1)-ші теңдеудің траекториялық аттракторының құрылымын және негізгі бар болу теоремасын тұжырымдаймыз.

Теорема 1 ([14–18]). Айталық, $\mathcal{K}^+$ траекторияларының кеңістігі, (1) теңдеуге сәйкес, $\mathcal{F}_+^b$ -кеңістігінде тұйық және (5)-ші шарт орындалсын. Айталық, $\mathcal{F}_+^b$ – де шенелген және Θ_+^{loc} – да компакттілі, $\{S(t)\}$ жартылай группаның жылжытуының $\mathcal{P} \subseteq \mathcal{K}^+$ тартылатын жиыны бар болсын. Онда $\mathcal{K}^+$ – на әсер етуші $\{S(t), t \geq 0\}$ жылжытудың жартылай группасының $\mathcal{U} \subseteq \mathcal{P}$ траекториялық аттракторы бар болады. $\mathcal{U}$ жиыны $\mathcal{F}_+^b$ –де шенелген және Θ_+^{loc} – де компакттілі болады.

(1)-ші теңдеудің $\mathcal{U}$ траекториялық аттракторының құрылымын осы теңдеудің толық траекториялары тұрғысынан бейнелейік. Уақыттың барлық сандық осінде (1)-ші теңдеуді қарастырайық

$$\frac{\partial u}{\partial t} = A(u), \quad t \in \mathbb{R}. \quad (7)$$

Бұл анықтаманы барлық $\mathbb{R}$ осіне жалпылайық. Егер $f(s)$, $s \in \mathbb{R}$ функциясы барлық уақыт осінде берілсе, онда $S(h)f(s) = f(s+h)$ жылжытулары теріс h сандар үшін анықталады. Егер $\Pi_+ u(s+h) \in \mathcal{K}^+$ кез-келген $h \in \mathbb{R}$ үшін $u(s)$, $s \in \mathbb{R}$ функциясы (7) теңдеудің толық траекториясы деп аталады.

Мұнда $\Pi_+ = \Pi_{0,\infty}$ операторының жарты осьте $\mathbb{R}_+$ шектелуін білдіреді.

Сәйкесінше $\mathcal{F}_+^{loc}, \mathcal{F}_+^b$ – нен және Θ_+^{loc} – нан $\mathcal{F}^{loc}, \mathcal{F}^b$ және Θ^{loc} кеңістіктерін анықтаймыз:

$$\mathcal{F}^{loc} := \{f(s), s \in \mathbb{R} \mid \Pi_{t_1, t_2} f(s) \in \mathcal{F}_{t_1, t_2} \quad \forall [t_1, t_2] \subseteq \mathbb{R}\};$$

$$\mathcal{F}^b := \{f(s) \in \mathcal{F}^{loc} \mid \|f\|_{\mathcal{F}^b} < +\infty\},$$

мұнда

$$\|f\|_{\mathcal{F}^b} := \sup_{h \in \mathbb{R}} \|\Pi_{0,1} f(h+s)\|_{\mathcal{F}_{0,1}} \quad (8)$$

Θ^{loc} топологиялық кеңістік (жиын ретінде) $\mathcal{F}^{loc}$ – мен сәйкес келеді, және Θ^{loc} – нан анықтама бойынша $f_k(s) \rightarrow f(s)$ ($k \rightarrow \infty$), егер Θ_{t_1, t_2} – де кез келген $[t_1, t_2] \subseteq \mathbb{R}$ үшін $\Pi_{t_1, t_2} f_k(s) \rightarrow \Pi_{t_1, t_2} f(s)$ ($k \rightarrow \infty$). Θ^{loc} - Θ_+^{loc} сияқты метрикалық кеңістік болып табылатыны айқын

Анықтама 3. (9) теңдеуінің $\mathcal{F}^b$ кеңістігіндегі $\mathcal{K}$ өзегі (9) теңдеуінің $\mathcal{F}^b$ нормасында шенелген барлық толық $u(s)$, $s \in \mathbb{R}$ траекторияларының бірігуі болады:

$$\|\Pi_{0,1} u(h+s)\|_{\mathcal{F}_{0,1}} \leq C_u, \quad \forall h \in \mathbb{R}.$$

Теорема 2. Айталық, 1-ші теореманың шарттары орындалсын. Онда

$$\mathcal{U} = \Pi_+ \mathcal{K}$$

$\mathcal{K}$ жиыны Θ^{loc} топологиясында компакттілі және $\mathcal{F}^b$ кеңістігінде шенелген.

Толық дәлелдеуі [13, б. 635–677; 18, б. 215–235]-да ұсынылған.

Келесі бөлімде $\varepsilon > 0$ кіші параметріне тәуелді болатыны Чэфи-Инфанте теңдеулерінің және олардың траекториялық аттракторлары зерттеледі.

Анықтама 4. Θ_+^{loc} топологиялық кеңістігінде $\mathcal{U}_\varepsilon$ траекториялық аттракторы $\varepsilon \rightarrow 0$ ұмтылғанда $\overline{\mathcal{U}}$ траекториялық аттракторына жинақталады дейміз, егер Θ_+^{loc} – ның кез-келген $\mathcal{O}(\mathcal{U})$ маңайында $\varepsilon_1 \geq 0$ болса, кез-келген $\varepsilon < \varepsilon_1$ үшін $\mathcal{U}_\varepsilon \subseteq \mathcal{O}(\mathcal{U})$ болатындай табылса, яғни, кез-келген $M > 0$ үшін

$$\text{dist}_{\Theta_{0,M}}(\Pi_{0,M} \mathcal{U}_\varepsilon, \Pi_{0,M} \overline{\mathcal{U}}) \rightarrow 0 \quad (\varepsilon \rightarrow 0).$$

Есептің қойылымы

Келесі шенелген облысты қарастырайық:

$$\Omega \subset \mathbb{R}^n, \quad n \geq 3.$$

Айталық, $Y = \left(-\frac{1}{2}, \frac{1}{2}\right)^n$ түріндегі бекітілген ұяшық берілсін. Мұнда тегіс шекарасы бар $G_0 \subset Y$ – жиыны қарастырылады.

Кез келген $\delta > 0$ үшін $B \subset \mathbb{R}^n$ жиыны келесідей түрде енгізіледі:

$$\delta B = \{x \in \mathbb{R}^n: \delta^{-1}x \in B\}.$$

Ары қарай, $\varepsilon > 0$ жеткілікті аз деп есептейміз, сонда :

$$\varepsilon^{n/(n-2)} G_0 \subset \varepsilon Y.$$

$j \in \mathbb{Z}^n$ жиыны үшін төмендегідей түрде енгіземіз:

$$P_\varepsilon^j = \varepsilon j, Y_\varepsilon^j = P_\varepsilon^j + \varepsilon Y, G_\varepsilon^j = P_\varepsilon^j + \varepsilon^{n/(n-2)} G_0.$$

Төмендегідей белгілеулер енгізейік:

$$\tilde{\Omega}_\varepsilon = \{x \in \Omega: \rho(x, \partial\Omega) > \sqrt{n\varepsilon}\},$$

Және индекстердің рұқсат етілген жиыны:

$$Y_\varepsilon = \{j \in \mathbb{Z}^n: G_\varepsilon^j \cap \tilde{\Omega}_\varepsilon \neq \emptyset\}.$$

Сондай-ақ $|Y_\varepsilon| \approx d\varepsilon^{-n}$, мұндағы $d > 0$ – тұрақты.

Келесідей перфорацияланған облысты қарастыралық:

$$\Omega_\varepsilon = \Omega \setminus \overline{G_\varepsilon}, \text{ мұнда } G_\varepsilon = \bigcup_{k \in \mathbb{Z}^n} (\varepsilon k + \varepsilon^{n/(n-2)} G_0) \cap \Omega,$$

ал, сондай-ақ, $S_\varepsilon := \partial\Omega_\varepsilon \setminus \partial\Omega$ – ішкі шекара.

Келесідей белгілеулер енгіземіз:

$$Q_\varepsilon = \Omega_\varepsilon \times (0, +\infty), Q = \Omega \times (0, +\infty).$$

Келесідей аралас (бастапқы-шекаралық) есепті қарастырайық:

$$\begin{cases} \frac{\partial u_\varepsilon}{\partial t} = \Delta u_\varepsilon - \sum_{j=0}^N b_j \left(\frac{x}{\varepsilon}\right) u_\varepsilon^j + g(x), & x \in \Omega_\varepsilon, t > 0, \\ \varepsilon \frac{\partial u_\varepsilon}{\partial \nu} + \varepsilon^\theta q\left(x, \frac{x}{\varepsilon}\right) u_\varepsilon = 0, & x \in S_\varepsilon, t > 0, \\ u_\varepsilon = 0, & x \in \partial\Omega, \\ u_\varepsilon(x, 0) = U(x), & x \in \Omega_\varepsilon. \end{cases} \quad (9)$$

Мұнда, $N = 2\ell + 1$ -тақ сан.

$b_j(z) \in C_b(\mathbb{R}^n)$ ($z = (z_1, \dots, z_n)$) деп белгілейміз және

$$b_N(z) \geq \beta_N > 0 \quad \forall z \in \mathbb{R}^n \quad (10)$$

осындай шарт қоямыз.

Осыдан, төмендегідей орташалау орындалады.

$$b_j\left(\frac{x}{\varepsilon}\right) \rightarrow \bar{b}_j \quad L_{\infty, \text{weak}}(\Omega) \text{ — де } (j = 0, \dots, N). \quad (11)$$

$q\left(x, \frac{x}{\varepsilon}\right)$ функциясының $V' = H^{-1}(\Omega)$ – кеңістігінде $\bar{q}(x)$ орташалауы бар деп ұйғарамыз:

$q\left(x, \frac{x}{\varepsilon}\right) \rightarrow \bar{q}(x) \quad (\varepsilon \rightarrow 0+) \quad V'$ – де әлсіз мағынада.

Кеңістіктер үшін келесі белгілеулерді енгіземіз:

$$\begin{aligned} H &:= L_2(\Omega), & H_\varepsilon &:= L_2(\Omega_\varepsilon), \\ V &:= H_0^1(\Omega), & V_\varepsilon &:= \{v: H_0^1(\Omega_\varepsilon): v|_{\partial\Omega} = 0\}. \end{aligned}$$

Бұл кеңістіктердің нормалары, сәйкесінше, келесідей анықталады:

$$\begin{aligned} \|v\|_H^2 &= \int_\Omega |v(x)|^2 dx, & \|v\|_{H_\varepsilon}^2 &= \int_{\Omega_\varepsilon} |v(x)|^2 dx, \\ \|v\|_V^2 &= \int_\Omega |\nabla v(x)|^2 dx, & \|v\|_{V_\varepsilon}^2 &= \int_{\Omega_\varepsilon} |\nabla v(x)|^2 dx. \end{aligned}$$

Ескерту 2. $V' = H^{-1}(\Omega)$ - түйіндес кеңістігі, сәйкесінше $V'_\varepsilon = H^{-1}(\Omega_\varepsilon)$ түйіндес кеңістігі. $u_\varepsilon \in L^\infty_\infty(\mathbb{R}_+; H_\varepsilon) \cap L_2(\Omega)(\mathbb{R}_+; V_\varepsilon) \cap L_{2(l+1)}^\infty(\mathbb{R}_+; L_{2(l+1)}(\Omega_\varepsilon))$ функциясы (9) есебінің әлсіз шешімі деп аталады, егер кез келген $\psi \in C_0^\infty(\mathbb{R}_+; H_\varepsilon)$ үшін

$$\int_{Q_\varepsilon} u_\varepsilon \frac{\partial \psi}{\partial t} dx dt + \int_{Q_\varepsilon} \nabla u_\varepsilon \nabla \psi dx dt + \sum_{j=0}^N \int_{Q_\varepsilon} b_j \left(\frac{x}{\varepsilon} \right) u_\varepsilon^j \psi dx dt + \varepsilon^{\theta-1} \int_{S_\varepsilon} q \left(x, \frac{x}{\varepsilon} \right) u_\varepsilon \psi d\sigma dt = \int_{Q_\varepsilon} g(x) \psi dx dt. \quad (12)$$

u_ε – (9) мағынасындағы әлсіз шешім болсын. Онда $u_\varepsilon \in C(\mathbb{R}_+; H_\varepsilon)$ және дерлік барлық $t > 0$ үшін

$$\frac{1}{2} \frac{d}{dt} \|u_\varepsilon(t)\|_{H_\varepsilon}^2 + \int_{\Omega_\varepsilon} |\nabla u_\varepsilon|^2 dx + \int_{\partial\Omega} |u_\varepsilon|^2 d\sigma + \sum_{j=0}^N \int_{\Omega_\varepsilon} b_j \left(\frac{x}{\varepsilon} \right) |u_\varepsilon|^{j+1} dx = \int_{\Omega_\varepsilon} g(x) u_\varepsilon dx. \quad (13)$$

Осыдан стандартты тәсілмен априорлық бағалаулар алынады.

Ескерту 2. $U \in H_\varepsilon$ болсын және $\varepsilon > 0$ бекітейік. $u(x, 0) = U(x)$ болсын, онда (9) бастапқы-шекаралық есебі үшін әлсіз шешімнің бар екендігі дәлелденеді. Дәлелдеу стандартты Галеркин әдісіне негізделген (қараңыз, мысалы, [9]). Негізгі лемма (априорлық бағалауларды [10] қараңыз).

Лемма 1. Айталық, $u_\varepsilon(x, t) \in L_2^{loc}(\mathbb{R}_+; V_\varepsilon) \cap L_{2(l+1)}^{loc}(\mathbb{R}_+; L_{2(l+1)}(\Omega_\varepsilon))$ – Чэфи-Инфанте есептерінің әлсіз шешімі болсын. Сонда, барлық дерлік $T > 0$ үшін келесідей қасиеттер орындалады:

$$\int_{Q_\varepsilon} u_\varepsilon \frac{\partial \psi}{\partial t} dx dt + \int_{Q_\varepsilon} \nabla u_\varepsilon \nabla \psi dx dt + \sum_{j=0}^N \int_{\Omega_\varepsilon} b_j \left(\frac{x}{\varepsilon} \right) u_\varepsilon^j \psi dx dt + \varepsilon^{\theta-1} \int_{S_\varepsilon} q \left(x, \frac{x}{\varepsilon} \right) u_\varepsilon \psi d\sigma dt = \int_{Q_\varepsilon} g(x) \psi dx dt. \quad (14)$$

ε -ді бекітейік. Кейінгі талдауда, егер түсінікті болса, кеңістіктердің белгілерінде ε индексін тастап отырамыз. Енді біз 1-бөлімде сипатталған сызбаларды қолданып, (9) есеп үшін траекториялық аттрактор құрамыз; ол (1) түрге ие болады, егер $E_1 = L_{2l+1} \cap H_0^1(\Omega)$, $E_0 = H^{-r}$, $E = L_2(\Omega)$ және $A(u) = -\Delta u + f(u) - g(\cdot, \cdot)$.

(9)-ы есеп үшін $\mathcal{K}_\varepsilon^+$ траекториясының кеңістігін сипаттау кезінде 1-бөлімнің жалпы сызбасын пайдаланамыз және әр кесінді $[t_1, t_2] \in R$ үшін Банах кеңістігін анықтаймыз

$$\mathcal{F}_{t_1, t_2} := L_\infty(t_1, t_2; H) \cap L_2(t_1, t_2; V) \cap L_{2(l+1)}(t_1, t_2; L_{2(l+1)}(\Omega)) \quad (15)$$

нормамен

$$\|v\|_{\mathcal{F}_{t_1, t_2}} := \|v\|_{L_\infty(t_1, t_2; H)} + \|v\|_{L_2(t_1, t_2; V)} + \|v\|_{L_{2(l+1)}(t_1, t_2; L_{2(l+1)}(\Omega))}. \quad (16)$$

Алайда, (2) шарт (19)-ы нормасы үшін орындалады, ал жартылай группаның жылжытулары $\{S(h)\}$ (3)-і шартты қанағаттандырады.

$\mathcal{D}_{t_1, t_2} = L_2(t_1, t_2; V)$ қою арқылы, $\mathcal{F}_{t_1, t_2} \subseteq \mathcal{D}_{t_1, t_2}$ аламыз, ал, егер $u(s) \in \mathcal{F}_{t_1, t_2}$, онда $A(u(s)) \in \mathcal{D}_{t_1, t_2}$. Әрі қарай, (9)-ы есептің әлсіз шешімдерін 1-бөлімнің жалпы сызбасынан теңдеулер жүйесінің шешімі ретінде қарастыруға болады.

Кеңістікті анықтағаннан кейін (4) біз мынаны аламыз:

$$\begin{aligned}\mathcal{F}_+^{loc} &= L_\infty^{loc}(\mathbb{R}_+; H) \cap L_2^{loc}(\mathbb{R}_+; V) \cap L_{2(l+1)}^{loc}(\mathbb{R}_+; L_{2(l+1)}(\Omega)) \\ \mathcal{F}_{\varepsilon,+}^{loc} &= L_\infty^{loc}(\mathbb{R}_+; H_\varepsilon) \cap L_2^{loc}(\mathbb{R}_+; V_\varepsilon) \cap L_{2(l+1)}^{loc}(\mathbb{R}_+; L_{2(l+1)}(\Omega_\varepsilon)).\end{aligned}$$

$\mathcal{K}_\varepsilon^+$ арқылы (9) есептің барлық әлсіз шешімдерінің жиының белгілейміз. Назар аударайық, кез келген функция үшін $U \in H$ -ның кем дегенде $u(0) = U(x)$ орындалатындай, $u(\cdot) \in \mathcal{K}_\varepsilon^+$ бір траекториясы бар болады. Демек, (9) есептің $\mathcal{K}_\varepsilon^+$ траекториясының кеңістігі бос емес және жеткілікті үлкен.

$\mathcal{K}_\varepsilon^+ \subset \mathcal{F}_+^{loc}$ және $\mathcal{K}_\varepsilon^+$ траекториясының кеңістігі трансляционды-инвариантты болып табылатыны айқын, яғни, егер $u(s) \in \mathcal{K}_\varepsilon^+$, онда кез келген $h \geq 0$ үшін $u(h+s) \in \mathcal{K}_\varepsilon^+$. Демек,

$$S(h)\mathcal{K}_\varepsilon^+ \subseteq \mathcal{K}_\varepsilon^+, \quad \forall h \geq 0.$$

Әрі қарай, $L_2(t_1, t_2; H)$ кеңістігінің нормасын қолдана отырып, $\mathcal{F}_{t_1, t_2}$ кеңістігіндегі $\rho_{t_1, t_2}(\cdot, \cdot)$ метрикасын анықтаймыз және келесідей түрде жазамыз

$$\rho_{0,M}(u, v) = \left(\int_0^M \|u(s) - v(s)\|_H^2 ds \right)^{1/2}, \quad \forall u(\cdot), v(\cdot) \in \mathcal{F}_{0,M}.$$

Бұл метрикалар $\mathcal{F}_+^{loc}$ кеңістігінде Θ_+^{loc} топологияны туындатады (сәйкесінше $\mathcal{F}_{\varepsilon,+}^{loc}$ – $\Theta_{\varepsilon,+}^{loc}$). Назар аударыңыз, егер кез келген $M > 0$ $\|v_k(\cdot) - v(\cdot)\|_{L_2(0,M;H)} \rightarrow 0$ ($k \rightarrow \infty$) болса, $k \rightarrow \infty$ $\{v_k\} \subset \mathcal{F}_+^{loc}$ тізбегі $v \in \mathcal{F}_+^{loc}$ функциясына Θ_+^{loc} – да жинақталады. Θ_+^{loc} топологиясы өлшемді (қр. (6)) және тиісті метрикалық кеңістікте толық болады. Біз (9)-есептің кеңістігінің траекториясында $\mathcal{K}_\varepsilon^+$ топологияны қарастырамыз. $\mathcal{K}_\varepsilon^+$ әсер ететін жартылай группаның жылжытулары $\{S(t)\}$ қарастырылып отырған Θ_+^{loc} топологиясында үзіліссіз.

1-ші бөлімнің жалпы сызбасына сәйкес, $\mathcal{F}_+^b$ Банах кеңістігін қолдана отырып, шенелген жиынды $\mathcal{K}_\varepsilon^+$ анықтаймыз (қр. (7)).

$$\mathcal{F}^b = L_\infty^b(\mathbb{R}; H) \cap L_2^b(\mathbb{R}; V) \cap L_{2l+1}^b(\mathbb{R}; L_{2(l+1)}(\Omega)). \quad (17)$$

және $\mathcal{F}_+^b - \mathcal{F}_+^{loc}$ кеңістігінің ішкі кеңістігі екені айқын.

$\mathcal{K}_\varepsilon^+, S(t): \mathcal{K}_\varepsilon^+ \rightarrow \mathcal{K}_\varepsilon^+, t \geq 0$ – да $\{S(t)\}$ жартылай группаның жылжытуларын қарастырамыз.

Айталық, (9)-шы есептің барлық әлсіз шешімдерінен $u(s), s \in \mathbb{R}$ тұратын шенелген кеңістіктегі $\mathcal{K}_\varepsilon$ өзегін білдіреді,

$$\mathcal{F}^b = L_\infty^b(\mathbb{R}; H) \cap L_2^b(\mathbb{R}; V) \cap L_{2l+1}^b(\mathbb{R}; L_{2(l+1)}(\Omega)).$$

Бұл лемманың дәлелдеуі [24] және [18] ескере отырып, дербес жағдай үшін келтірілген [19] дәлелдеумен толығымен сәйкес келетіні туралы тұжырымдама бар.

Лемма 2. (9)-шы есептің Θ_+^{loc} топологиялық кеңістігінде $\mathcal{U}_\varepsilon$ траекториялық аттракторы бар. $\mathcal{U}_\varepsilon$ жиыны бірқалыпты ($\varepsilon \in (0, 1)$ – де) $\mathcal{F}_+^b$ кеңістігінде шенелген және Θ_+^{loc} топологиясында компактiлi. Әрі қарай,

$$\mathcal{U}_\varepsilon = \Pi_+ \mathcal{K}_\varepsilon,$$

мұнда (9) есебінің $\mathcal{K}_\varepsilon$ өзегі $\mathcal{F}^b$ кеңістігіне ($\varepsilon \in (0, 1)$ – де) бос емес және бірқалыпты шенелген.

Мұндағы, $\mathcal{F}_+^b$ кеңістігі және Θ_+^{loc} топологиясы кіші параметр ε -ға тәуелді болатынын ескереміз.

Негізгі нәтижелерді алу үшін «бөгде мүше» (потенциал) пайда болатын орташаланған Чэфи-Инфанте теңдеуін шығарып аламыз және берілген теңдеудің траекториялық аттракторлары әлсіз мағынада орташаланған теңдеудің траекториялық аттракторларына ұмтылатынын дәлелдейміз. Коэффициенттерге қосымша шарттар қойылған жағдайда, глобалдық аттрактордың да жинақталуы қолданылады.

Нәтижелер мен талқылау

Бұл бөлімде $\varepsilon \rightarrow 0$ кездегі перфорацияланған Ω_ε облысындағы Чэфи-Инфанте теңдеулерінің траекториялық аттракторлары $\mathcal{A}_\varepsilon$ -ның шектік жағдайын және олардың орташаланған теңдеудің траекториялық аттракторымен байланысын қарастырамыз.

Шектік теңдеудегі «бөгде мүше» (потенциалды) анықтау үшін келесідей көмекші есептер қарастырылады:

Потенциалды анықтау үшін көмекші есеп:

$$\begin{cases} -\Delta_y v(x, y) = 0, & y \in \mathbb{R}^n \setminus G_0, \\ \frac{\partial v}{\partial \nu_y}(x, y) + q(x, y)v(x, y) = q(x, y), & y \in \partial G_0, \\ v(x, y) \rightarrow 0, & |y| \rightarrow \infty, \end{cases}$$

Мұнда, x – баяу айнымалы.

Шектік потенциал төмендегідей түрде анықталады:

$$V(x) = \int_{\partial G_0} \frac{\partial v}{\partial \nu_y}(x, y) d\sigma_y.$$

Ішкі потенциал $V(x)$ – перфорацияланған облыстың микроструктуралық әсерін сипаттайтын коэффициент. Ол шектік есепте қосымша «бөгде мүше» ретінде шығады және бастапқы Чэфи-Инфанте теңдеуіндегі перфорациялардың әсерін модельдейді.

Орташаланған теңдеу келесідей түрде жазылады:

$$\begin{cases} \frac{\partial u}{\partial t} = \Delta u - \sum_{j=0}^N \bar{b}_j(x) u^j - V(x)u + g(x), & x \in \Omega, \\ u = 0, & x \in \partial\Omega, \\ u = U(x), & t = 0, \end{cases} \quad (18)$$

Мұндағы $\bar{b}_j(x)$ -орташаланған коэффициенттер. $V(x)u$ – перфорацияланған микрошекарадан туындайтын ішкі диссипация немесе потенциалдық тежеліс.

(18)-ші есептің Θ_+^{loc} топологиялық кеңістігінде $\mathcal{U}_\varepsilon$ траекториялық аттракторы бар. $\mathcal{U}_\varepsilon$ жиыны $\mathcal{F}_+^b$ кеңістігінде ($\varepsilon \in (0, 1)$ – де) бірқалыпты шенелген және Θ_+^{loc} топологиясында компактiлi. Әрi қарай,

$$\mathcal{U}_\varepsilon = \Pi_+ \mathcal{K}_\varepsilon,$$

мұнда (18) есебiнiң $\mathcal{K}_\varepsilon$ өзегi – $\mathcal{F}_+^b$ кеңістігi ($\varepsilon \in (0, 1)$ – де) бос емес және бiрқалыпты шенелген.

Көмекші есептер:

Келесі шексіз $t \in \mathbb{R}$ уақыт аралығында берілген есепті қарастырайық: $u_{\varepsilon_k}(s)$, мұндағы $s \in \mathbb{R}$ келесідей есептің шешімі болсын:

$$\begin{cases} \frac{\partial u_{\varepsilon_k}}{\partial t} = \Delta u_{\varepsilon_k} - \sum_{j=0}^N b_j\left(\frac{x}{\varepsilon_k}\right) u_{\varepsilon_k}^j + g(x), & x \in \Omega_{\varepsilon_k}, t > 0 \\ \frac{\partial u_{\varepsilon_k}}{\partial \nu} + \varepsilon_k^{\theta-1} q\left(x, \frac{x}{\varepsilon_k}\right) u_{\varepsilon_k} = 0, & x \in S_{\varepsilon_k}, t > 0, \\ u_{\varepsilon_k} = 0, & x \in \partial\Omega, t > 0 \\ u_{\varepsilon_k} = U(x), & x \in \Omega_{\varepsilon_k}, t = 0, \end{cases} \quad (19)$$

Шешімнің ε -ке біртекті бағалаулар алу үшін біз келесі леммаларды қолданамыз (қараңыз [21, III-тарау, § 5] және [18]).

Лемма 3. Егер

$$W(f, g) = \int_{\Omega_\varepsilon} \nabla f(x) \nabla g(x) dx + \int_{\Omega_\varepsilon} q(x) f(x) g(x) dx + \int_{\partial\Omega_\varepsilon} r(x) f(x) g(x) ds \quad (20)$$

екі сызықты форма V_ε кеңістігінде анықталған болса және $q(x) \geq 0, r(x) \geq 0$ ($q(x) \not\equiv 0$ немесе $r(x) \not\equiv 0$), онда $W(f, g)$ ішкі көбейтінді анықтайды және ол

$$(f, g)_{H^1(\Omega_\varepsilon)} = \int_{\Omega_\varepsilon} (\nabla f(x) \nabla g(x) + f(x)g(x)) \, dx$$

көбейтіндісіне эквивалентті.

Дәлелдеуі. Функционал $W(f, g)$ айқын түрде симметриялы және екі сызықты. Сонымен қатар, $W(f, f) \geq 0$. Егер $f \equiv 0$ болмаса, онда $q(x)$ немесе $r(x)$ о болғандықтан, $W(f, f) > 0$. Осылайша, $W(f, g)$ V_ε кеңістігінде ішкі көбейтінді анықтайды. Эквиваленттілік стандартты Пуанкаре типті теңсіздіктер арқылы дәлелденеді ([15] қараңыз).

Лемма 4. (18) есебінің коэрцитивтігі бастапқы (9) есептің коэрцитивтігін қамтамасыз етеді.

Осыдан, интегралдық теңсіздік (14) және Леммалар 1, 3, 4 көмегімен мына бағалауларды аламыз:

$$\|u_{\varepsilon_k}\|_{\mathcal{F}^b} \leq C, k \in \mathbb{N}, \quad (21)$$

мұндағы константа C ε -ке тәуелсіз.

Толық дәлелдеулері [15] келтірілген.

Лемма 5. Барлық $\varphi(x) \in H_\varepsilon$ және t үшін мына теңсіздік орындалады:

$$\left| \frac{n}{\varepsilon^{n-2}} \sum_{j \in Y_\varepsilon} \int_{\partial G_\varepsilon^j} b \left(x, \frac{x - p_\varepsilon^j}{\varepsilon^{n/(n-2)}} \right) \varphi(x) \, ds - \int_{\Omega} V(x) \varphi(x) \, dx \right| \leq M \varepsilon \|\varphi\|_{H_\varepsilon}, \quad (22)$$

мұндағы M константасы ε -ке тәуелсіз.

Дәлелдеуі. (22) теңсіздіктің дәлелденуі [18]-тегі (21) леммаға сүйене отырып дәлелденеді. ([18] қараңыз.)

(22) жинақталуын алу үшін $u_\varepsilon(x, t)$ -ні (14) теңдікте сынақ функция ретінде қойып, шенелгендікті $\varepsilon \rightarrow 0$ кезінде аламыз.

$u_\varepsilon(x, t)$ функциясын (14) теңдеудегі тесттік функция ретінде қою арқылы келесі біртекті шектеулер алынады:

$$\|\nabla u_\varepsilon\|_{H_\varepsilon} \leq K,$$

Сондай-ақ,

$$\sum_{j \in Y_\varepsilon} \int_{\partial G_\varepsilon^j} b \left(x, \frac{x - p_\varepsilon^j}{\varepsilon^{n/(n-2)}} \right) u_\varepsilon(x, t) \psi(x) \, d\sigma \leq K \varepsilon^{\frac{n}{n-2}},$$

Мұндағы, K ε -ден тәуелсіз тұрақты.

[17]-тегі 4 және 5 леммаларды, сондай-ақ $h_\varepsilon(x) \rightarrow h(x)$ әлсіз жинақталуын V кеңістігінде $\varepsilon \rightarrow 0$ кезінде қолдану арқылы шектік жағдайға көшу мүмкіндігі орындалады.

$k \rightarrow \infty$ кезінде (19) нөмірлі есептің теңдеуінде шекке өте отырып, кеңістік $D'(\mathbb{R}_+; H^{-r})$ – де Лемма 5-тің дәлелдеуін және (22)-ші формуланы пайдалана отырып $\bar{u}(x, s)$ функциясы келесі есептің шешімі болатыны дәлелденеді:

$k \rightarrow \infty$ болғандағы перфорацияланған Ω_ε облысындағы есеп:

$$\begin{cases} \frac{\partial u_\varepsilon}{\partial t} = \Delta u_\varepsilon - \sum_{j=0}^N b_j \left(\frac{x}{\varepsilon} \right) u_\varepsilon^j + g(x), & x \in \Omega_\varepsilon, \\ \varepsilon \frac{\partial u_\varepsilon}{\partial \nu} + \varepsilon^\theta q \left(x, \frac{x}{\varepsilon} \right) u_\varepsilon = 0, & x \in S_\varepsilon, \\ u_\varepsilon = U(x), & x \in \partial\Omega, t = 0, \end{cases} \quad (23)$$

орташалау және жинақталу арқылы келесі шектік есепке ауысады:

$$u(x, s) = \lim_{k \rightarrow \infty} u_{\varepsilon_k}(x, s). \quad (24)$$

Орташаланған (шектік) есеп мына түрде жазылады:

$$\begin{cases} \frac{\partial u}{\partial t} = \Delta u - \sum_{j=0}^N \bar{b}_j u^j + g(x), & x \in \Omega, \\ \frac{\partial u}{\partial \nu} + \bar{q}(x)u = 0, & x \in \partial\Omega, \\ u = U(x), & x \in \Omega. \end{cases} \quad (25)$$

Мұндағы, $\bar{b}_j$ – орташаланған (шектік) коэффициенттер мен функциялар. $\bar{q}(x)u$ – шектік теңдеуде «бөгде мүше» ретінде көрінетін, перфорацияланған микрошекаралық әсерді есепке алатын шекаралық потенциал.

Демек,

$$u(x, s) \in \mathcal{K}.$$

Төмендегідей келтірілген дәлелдеулерге сәйкес Θ_+^{loc} кеңістігінде:

$$u_{\varepsilon_k}(x, s) \rightarrow u(x, s) \quad k \rightarrow \infty \text{ үшін}$$

Тұжырым 1. Егер

$$u_{\varepsilon_k}(x, s) \notin \mathcal{O}'(\mathcal{K}) \text{ болса, онда} \\ u(x, s) \notin \mathcal{O}'(\mathcal{K})$$

және тіпті, $u(x, s) \notin \mathcal{K}$, болады, бұл алдыңғы тұжырымға қайшы келеді.

Тұжырым 2. $\mathcal{U}_\varepsilon$ және $\overline{\mathcal{U}}$ траекториялық аттракторлары $\varepsilon > 0$ -да $L_\infty(\mathbb{R}; L_\infty(\Omega)) = L_\infty(\Omega \times \mathbb{R})$ кеңістіктерінде бірқалыпты шенелген.

Дәлелдеуін қараңыз [8; т. IV, 6.2 тұжырым].

Бізге кейбір қосымша анықтамалар мен белгілеулер қажет болады.

Төмендегідей белгілеулер енгізейік: $Q_T = \Omega \times]T, T + 1[$, $T \in \mathbb{R}$, биіктігі 1-ге тең цилиндр.

(9) және (25) теңдеулерінің әлсіз шешімдерін $u(x, t)$ қарастырамыз, мұнда

$$u(x, s) \in W_p^r(Q_T) \cap \{u|_{\partial\Omega} = 0\} \equiv W_{p,0}^r(Q_T) \quad \forall T \in \mathbb{R}$$

Мұндағы, $r = (2, \dots, 2, 1)$, ал $W_p^r(Q_T)$ – Никольский кеңістігін білдіреді. Айта кетелік, $W_p^r(Q_T)$ нормасын мына өрнекпен белгіленеді:

$$\|u, Q_T\|_{r,p}^p = \int_{Q_T} \left(\sum_{|\alpha| \leq 2} |\partial_x^\alpha u(x, s)|^p + |\partial_t u(x, s)|^p \right) dx ds$$

Никольскийдің енгізулер теоремасы төмендегідей тұжырымдалады (кр. [18]):

Егер $p > (n+2)/2$, онда

$$W_{p,0}^r(Q_T) \subseteq C^{\kappa r}(\bar{Q}_T) \cap \{u|_{\partial\Omega} = 0\} \equiv C_0^{\kappa r}(\bar{Q}_T) \quad (26)$$

Мұндағы, $0 < \kappa < 1 - \frac{1}{n} \left(1 + \frac{n}{2}\right)$.

Демек, бекітілген $\delta, 0 < \delta < 1$ үшін, (26)-тен шығады, $\kappa = 1 - \delta$ үшін төмендегідей компактiлі енгізулер орындалады:

$$W_{p,0}^r(Q_T) \subseteq C_0^{2-2\delta, 1-\delta}(\bar{Q}_T) \quad \text{үшін } p > \frac{n+2}{2\delta}. \quad (27)$$

Бізге келесідей Солонников теоремасы қажет болады ([19]). Төмендегідей белгілеулер енгізейік $W_{2,0}^r(Q_0)$, $Q_0 = \Omega \times]0, 1[$ функцияларының V_0 кеңістігіндегі $\{t = 0\}$ -гі іздері.

Егер $w(x, t) \in L_\infty(0, 1; L_\infty(\Omega)) \cap L_2(0, 1; H_0^1(\Omega))$ функциясы жылу өткізгіштік теңдеуінің шешімі болсын:

$$\partial_t w - \Delta w = h(x, t), (x, t) \in Q_0,$$

және бастапқы шартты қанағаттандырса

$$w|_{t=0} = w_0(x).$$

$h(x, t) \in L_\infty(0, 1; L_\infty(\Omega)) \equiv L_\infty(Q_0)$ және $w_0 \in V_0$ болжайық. Онда $w(x, t) \in W_{2,0}^r(Q_0)$ және

$$\|w, Q_0\|_{r,p} \leq C_p \left(\|h\|_{L_p(Q_0)} + \|w_0\|_{V_0} \right). \quad (28)$$

$\mathcal{W}_n^{\text{loc}}$ кеңістігін келесідей функциямен $v(x, s)$ анықталады, егер

$$\Pi_{T, T+1} v(x, s) \in W_{p,0}^r(Q_T) \quad \forall T \in \mathbb{R}. \quad (29)$$

Әдеттегідей, $\mathcal{W}_n^b \subset \mathcal{W}_n^{\text{loc}}$ кеңістігі төмендегідей нормамен анықталады:

$$\|v\|_{\mathcal{W}_p^b} = \sup_{T \in \mathbb{R}} \|\Pi_{T, T+1} v, Q_T\|_{r,p}. \quad (30)$$

Соңында, $\mathcal{W}_{p,w}^{\text{loc}} - \mathcal{W}_p^{\text{loc}}$ кеңістігі, бірақ, онда кез келген $T \in \mathbb{R}$ үшін $W_{p,0}^r(Q_T)$ бойынша локалды әлсіз топологияда жинақталады. (27)-тен төмендегі шығады:

$$\Theta^{\text{loc}} \in C_{0,\text{loc}}^{2-2\delta, 1-\delta}(\bar{\Omega} \times \mathbb{R}) \quad \text{для } p > \frac{n+2}{2\delta} \quad (31)$$

Чэфи-Инфанте бастапқы-шекаралық есебі үшін аттракторлардың жинақталуы туралы негізгі теорема төмендегідей түрде тұжырымдалады:

Теорема 3. $\varepsilon > 0$ –да (9)-шы теңдеуінің $\mathcal{A}_\varepsilon$ траекториялық аттракторы бірқалыпты $\mathcal{W}_p^b$ –де шенелген және $\mathcal{W}_{p,w}^{\text{loc}}$ –де $\varepsilon \rightarrow 0 +$ ұмтылғанда $\mathcal{A}_\varepsilon$ аттракторы (25)-ші теңдеудің $\overline{\mathcal{A}}$ траекториялық аттракторына жинақталады. Атап айтқанда, (31)-ші енгізулерден шығады

$$\text{dist}_{C^{2-2\delta, 1-\delta}(\bar{\Omega} \times [0, M])}(\Pi_M \mathcal{A}_\varepsilon, \Pi_M \overline{\mathcal{A}}) \rightarrow 0 \quad (\varepsilon \rightarrow 0+) \quad \forall M > 0. \quad (32)$$

Дәлелдеуі. $u \in \mathcal{K}_\varepsilon$ және $\bar{u} \in \overline{\mathcal{K}}$ болсын. 2-ші тұжырымдаудан және $|f(v, z)| \leq C_1(|v|^{p-1} + 1) \quad \forall v \in \mathbb{R}, z \in \mathbb{R}^n$ теңсіздіктен төмендегідей теңдеулердің оң жақ бөліктері шығады:

$$\partial_t u - \Delta u = -b\left(u, \frac{x}{\varepsilon}\right), \quad u|_{\partial\Omega} = 0, \quad 0 < \varepsilon \leq 1, \quad (33)$$

$$\partial_t \bar{u} - \Delta \bar{u} = -\bar{b}(\bar{u}), \quad \bar{u}|_{\partial\Omega} = 0. \quad (34)$$

$\varepsilon > 0$ –да бірқалыпты $L_\infty(\Omega \times \mathbb{R})$ –де шенеледі. Сондықтан $\mathcal{K}_\varepsilon$ және $\overline{\mathcal{K}}$ ядролары мына жиынға жатады:

$$B_0 = \left\{ u: \|u\|_{\mathcal{W}_p^b} \leq R_0 \right\}, \quad (35)$$

мұндағы R_0 ε –нан тәуелсіз. Шынында да, (33) теңдеуін кесілген монотонды функцияға көбейтейік: $\Phi_T(t) = \Phi(t - T)$, мұнда $\Phi(s) \in C_b^1(\mathbb{R}; \mathbb{R}_+)$, $\Phi(s) = 0$ егер $s \leq -1$ және $\Phi(s) = 1$ егер $s \geq 0$. Енді $T \geq 1$ болсын. Төмендегідей теңдікті аламыз:

$$\partial_t (\Phi_T u) - \Delta (\Phi_T u) = -\Phi_T f\left(u, \frac{x}{\varepsilon}\right) - \Phi'_T u. \quad (36)$$

Себебі $\Phi_T(t)u(x, t) \equiv 0$ егер $t = T - 1$, онда Солонниковтың (28) бағалауын $Q_{T-1, T+1} = \Omega \times]T - 1, T + 1[$ облысында қолдануға болады, нәтижесінде мына теңсіздік шығады:

$$\begin{aligned} & \|u, Q_T\|_{r,p} \\ & \leq \|\Phi_T u, Q_{T-1, T+1}\|_{r,p} \leq C_p \left\| -\Phi_T f\left(u, \frac{x}{\varepsilon}\right) - \Phi'_T u \right\|_{L_p(Q_{T-1, T+1})} \leq R_0 \quad \forall T \in \mathbb{R}. \end{aligned}$$

Демек, B_0 жиыны (35) $\mathcal{K}_\varepsilon$ және $\overline{\mathcal{K}}$ –ті қамтиды. B_0 жиыны $\mathcal{W}_{p,w}^{\text{loc}}$ –та компактiлi болғандықтан, стандартты пайымдауларды қолданып, $\mathcal{K}_\varepsilon \rightarrow \overline{\mathcal{K}}$ екенін аламыз, мұнда $\varepsilon \rightarrow 0 +$ кезінде $sW_{p,w}^r(\Omega \times]-M, M[)$ –де, демек, $\mathcal{A}_\varepsilon \rightarrow \overline{\mathcal{A}}$ $\varepsilon \rightarrow 0 +$ да $W_{p,w}^r(\Omega \times]0, M[)$ –де орындалады, кез келген $M > 0$ үшін. Атап айтқанда, (32) орындалады.

Есептің жалғыз шешімі болған кезде Теорема 3-тен төмендегі Салдар 3 шығады:

Салдар 3: (9) және (25) теңдеулерінің $\mathcal{A}_\varepsilon$ және $\overline{\mathcal{A}}$ глобалды аттракторлары $C^{2-\delta}(\bar{\Omega})$ кеңістігінде жатады және

$$\text{dist}_{C^{2-\delta}(\bar{\Omega})}(\mathcal{A}_\varepsilon, \overline{\mathcal{A}}) \rightarrow 0 \quad (\varepsilon \rightarrow 0+) \quad \forall \delta \in]0, 1].$$

Глобалды аттракторлар туралы толығырақ мәліметті [10, 6, 21] әдебиеттерден табуға болады.

Қорытынды

Бұл жұмыста ε параметріне тәуелді перфорацияланған облыстағы Чэфи-Инфанте теңдеуі үшін анықталған траекториялық аттрактор $\mathcal{A}_\varepsilon$ -ның $\varepsilon \rightarrow 0$ кезінде әлсіз мағынада тиісті функционалдық кеңістіктегі орташаланған теңдеудің траекториялық аттракторына $\mathcal{A}$ жинақталатыны дәлелденді. Сонымен қатар, теңдеу коэффициенттеріне белгілі бір қосымша шарттар орындалған жағдайда глобалдық аттракторлардың да жинақталуы көрсетілді. Алынған нәтижелер – жаңа.

Жұмыста келесі негізгі нәтижелер алынды:

1. Перфорацияланған облыстағы Чэфи-Инфанте теңдеуінің орташаланған шектік есебі алынды;
2. Біртекті бағалаулар үшін көмекші леммалар тұжырымдалды;
3. Шектік теңдеудегі «бөгде мүше» (потенциалды) анықтау үшін көмекші есептер қарастырылды;
4. Чэфи-Инфанте теңдеуі үшін анықталған траекториялық аттрактор $\mathcal{A}_\varepsilon$ -ның $\varepsilon \rightarrow 0$ кезінде әлсіз мағынада тиісті функционалдық кеңістіктегі орташаланған теңдеудің траекториялық аттракторына $\mathcal{A}$ жинақталатыны дәлелденді.

Қаржыландыру туралы ақпарат

Бұл зерттеуді Қазақстан Республикасы Ғылым және жоғары білім министрлігінің Ғылым комитеті (Грант №АР22684340) қаржыландырады. Авторлар құнды пікірталастар үшін ф.-м.ғ.д., профессор К.А. Бекмаганбетов пен ф.-м.ғ.д., профессор Г.А. Чечкинге алғысын білдіреді.

ӘДЕБИЕТТЕР

- 1 Bekmaganbetov, K.A., Chechkin, G.A., Chepyzhov, V.V. Attractors and a 'strange term' in homogenized equation. *C R Mec.*, 348(5), 351–359 (2020). <https://doi.org/10.5802/crmeca.15>
- 2 Bekmaganbetov, K.A., Chechkin, G.A., Chepyzhov, V.V. 'Strange term' in homogenization of attractors of reaction-diffusion equation in perforated domain. *Chaos Solit Fractals.*, 140, Article ID 110208 (2020). <https://doi.org/10.1016/j.chaos.2020.110208>
- 3 Belyaev, A.G., Piatnitski, A.L., Chechkin, G.A. Asymptotic behavior of solution for boundary-value problem in a perforated domain with oscillating boundary. *Sib Math J.*, 39(4), 730–754 (1998). <https://doi.org/10.1007/BF02677919>
- 4 Cioranescu, D., Murat, F. Un terme étrange venu d'ailleurs I & II. In: Berzis H, Lions JL, editors. *Nonlinear Partial Differential Equations and their Applications. Collège de France Seminar, Volume II & III. Research Notes in Mathematics*, 60 & 70, London: Pitman; 1982, pp. 98–138 & 154–178.
- 5 Marchenko, V.A., Khruslov, E.Y. *Boundary value problems in domains with fine-grain boundary*. Kiev: Naukova Dumka, 1974.
- 6 Chechkin, G.A., Piatnitski, A.L., Shamaev, A.S. *Homogenization: methods and applications*. Providence (RI): American Mathematical Society, 2007. ISBN: 978-0-8218-4309-6
- 7 Jikov, V.V., Kozlov, S.M., Oleinik, O.A. *Homogenization of differential operators and integral functionals*. Berlin: Springer, 1994. <https://link.springer.com/book/10.1007/978-3-642-84659-5>
- 8 Sanchez-Palencia, É. *Homogenization techniques for composite media*. Berlin: Springer, 1987. <https://doi.org/10.1007/3-540-17616-0>
- 9 Babin, A.V., Vishik, M.I. *Attractors of evolution equations*. Amsterdam: Elsevier, 1992. [https://doi.org/10.1016/S0168-2024\(08\)70270-4](https://doi.org/10.1016/S0168-2024(08)70270-4)
- 10 Chepyzhov, V.V., Vishik, M.I. *Attractors for equations of mathematical physics*. Providence (RI): American Mathematical Society, 2002. <https://doi.org/10.1051/cocv:2002056>

- 11 Temam, R. Infinite-dimensional dynamical systems in mechanics and physics. New York (NY): Springer, 1988. (Applied Mathematics Series; 68). <https://doi.org/10.1007/978-1-4684-0313-8>
- 12 Efendiev, M., Zelik, S. Attractors of the reaction-diffusion systems with rapidly oscillating coefficients and their homogenization. *Ann Inst H Poincaré Anal Non Linéaire*, 19, 961–989 (2002). [https://doi.org/10.1016/S0294-1449\(02\)00115-4](https://doi.org/10.1016/S0294-1449(02)00115-4)
- 13 Ilyin, A.A. Averaging principle for dissipative dynamical systems with rapidly oscillating right-hand sides. *Sb Math.*, 187, 635–677 (1996). <https://doi.org/10.1070/SM1996v187n05ABEH000126>
- 14 Chechkin, G.A., Chepyzhov, V.V., Pankratov, L.S. Homogenization of trajectory attractors of Ginzburg-Landau equations with randomly oscillating terms. *Discrete Continuous Dyn Syst Ser B.*, 23(3), 1133–1154 (2018). <https://doi.org/10.3934/dcdsb.2018145>
- 15 Belyaev, A.G., Piatnitski, A.L., Chechkin, G.A. Averaging in a perforated domain with an oscillating third boundary condition. *Sb Math.*, 192(7), 933–949 (2001). <https://doi.org/10.1070/SM2001v192n07ABEH000576>
- 16 Lions, J.-L. Quelques méthodes de résolutions des problèmes aux limites non linéaires. Paris: Dunod, GauthierVillars, 1969. <https://www.scribd.com/document/482201537/96314508>
- 17 Diaz, J.I., Gomez-Castro, D., Shaposhnikova, T.A., et al. Classification of homogenized limits of diffusion problems with spatially dependent reaction over critical-size particles. *Appl Anal.*, 98(1–2), 232–255 (2018). <https://doi.org/10.1080/00036811.2018.1441997>
- 18 Chechkin, G.A., Piatnitski, A.L. Homogenization of boundary-value problem in a locally periodic perforated domain. *Appl Anal.*, 71(1–4), 215–235 (1999). <https://doi.org/10.1080/00036819908840714>
- 19 Chepyzhov, V.V., Vishik, M.I. Trajectory attractors for reaction-diffusion systems. *Top Meth Nonlinear Anal J Julius Schauder Center.*, 7(1), 49–76 (1996). <https://doi.org/10.12775/TMNA.1996.002>
- 20 Chepyzhov, V.V., Goritsky, A.Y., Vishik, M.I. Integral manifolds and attractors with exponential rate for nonautonomous hyperbolic equations with dissipation. *Russ J Math Phys.*, 12, 17–39 (2005).
- 21 Chepyzhov, V.V., Vishik, M.I., Wendland, W.L. On non-autonomous sine-Gordon type equations with a simple global attractor and some averaging. *Discrete Contin Dyn Syst.*, 12, 27–38 (2005). <https://doi.org/10.3934/dcds.2005.12.27>
- 22 Vishik, M.I., Chepyzhov, V.V. Approximation of trajectories lying on a global attractor of a hyperbolic equation with an exterior force that oscillates rapidly over time. *Sb Math.*, 194, 1273–1300 (2003). <https://doi.org/10.1070/SM2003v194n09ABEH000765>
- 23 Chueshov, I., Schmalfuß, B. Averaging of Attractors and Inertial Manifolds for Parabolic PDEs with Random Coefficients. *Applied Mathematics and Optimization*, 2005 (5), 461–492, 10.1515/ans-2005-0402
- 24 Carvalho, A.N., Langa, J.A., Robinson, J.C. Attractors for Infinite-Dimensional Non-Autonomous Dynamical Systems. Springer, 2013. V. 182. <https://doi.org/10.1007/978-1-4614-4581-4>
- 25 Solonnikov, V.A. On the boundary-value problems for linear parabolic systems of differential equations of the general type. *Trudi MIAN.* 1965. V.83.

^{1*}**Toleubay A.M.,**

PhD, ORCID ID: 0000-0002-5490-8212,

*e-mail: altyn.15.94@mail.ru

²**Karipova A.T.,**

PhD student, ORCIDID: 0009-0008-7837-1801,

e-mail: karipova-aidana@mail.ru

¹«Shakarim University» NJSC, Semey, Kazakhstan

²L.N. Gumilyov Eurasian National University, Astana, Kazakhstan

HOMOGENIZATION OF ATTRACTORS OF THE CHAFEE–INFANTE EQUATIONS IN A PERFORATED DOMAIN

Abstract

This work addresses the averaging problem for the Chafee–Infante equation in a micro-heterogeneous medium. It analyzes the equation with rapidly oscillating terms and dissipation. The model introduces a small parameter and formulates an averaged problem based on its limits. The study proves that the trajectory attractors of

the original equation converge to those of the averaged equation. It establishes the relevant theorem. In cases with unique solutions, it investigates the convergence of global attractors, adding further conditions on the nonlinear terms. The work also analyzes singular problems with nontrivial boundary conditions in perforated domains and on cavity boundaries. Here, the averaged equation differs from the original, reflecting the model's effective averaged characteristics and possibly an extra potential term. As the small parameter approaches zero, the study shows that the attractors converge in Hausdorff distance. The main result is that the attractors of the original Chafee–Infante equation converge to those of the averaged (limit) equation as the small parameter approaches zero. This work establishes this convergence for the first time in the context of homogenization in a periodically perforated medium, highlighting the scientific novelty and relevance of the findings. This research advances averaging methods for differential operators and may apply to solutions of nonlinear differential equations.

Keywords: Chafee-Infante equations, porous medium, attractors, perforated domain, homogenization.

^{1*}Толеубай А.М.,
PhD, ORCIDID: 0000-0002-5490-8212,
*e-mail: altyn.15.94@mail.ru
²Карипова А.Т.,
докторант, ORCIDID: 0009-0008-7837-1801,
e-mail: karipova-aidana@mail.ru

¹НАО «Шәкәрім университет», г. Семей, Казахстан

²Евразийский национальный университет им. Л.Н. Гумилева, г. Астана, Казахстан

УСРЕДНЕНИЕ АТТРАКТОРОВ УРАВНЕНИЙ ЧЭФИ-ИНФАНТЕ В ПЕРФОРИРОВАННОЙ ОБЛАСТИ

Аннотация

В данной работе рассматривается задача усреднения уравнения Чэфи-Инфанте в микронеоднородной среде. В ходе исследования подробно анализируются уравнения Чэфи-Инфанте с быстро осциллирующими членами и диссипацией. В модель вводится+ малый параметр, и в зависимости от его предельных значений формулируется усредненная (предельная) задача. Доказывается, что траекторные аттракторы исходного уравнения сходятся к аттракторам усредненного уравнения, и формулируется соответствующая теорема. Кроме того, показано, что при условиях единственности решений имеет место сходимость глобальных аттракторов, что достигается при наложении дополнительных условий на нелинейные члены. В работе также изучаются сингулярные задачи с нетривиальными граничными условиями в перфорированных областях и на границах их полостей. В этом случае получаемое усредненное (предельное) уравнение имеет иную структуру по сравнению с исходным. Оно отражает эффективные средние характеристики модели и может включать дополнительный потенциальный член. Доказано, что при стремлении малого параметра к нулю аттракторы сходятся в смысле Хаусдорфа. В качестве основного результата описаны усредненные аттракторы и показано, что аттракторы исходной системы стремятся к аттракторам предельного уравнения. В данной работе впервые рассматривается задача усреднения уравнений Чафе–Инфанте в периодически перфорированной среде, обосновываются научная новизна и актуальность полученных результатов. Данное исследование вносит вклад в развитие методов теории усреднения дифференциальных операторов и может найти применение при изучении решений нелинейных дифференциальных уравнений.

Ключевые слова: уравнения Чафе–Инфанте, пористая среда, аттракторы, перфорированная область, усреднение.

Мақаланың редакцияға түскен күні: 12.09.2025

**ФИЗИКАЛЫҚ ҒЫЛЫМДАР
PHYSICAL SCIENCES
ФИЗИЧЕСКИЕ НАУКИ**

УДК 524.3
МРНТИ 41.23.39

<https://doi.org/10.55452/1998-6688-2025-22-4-340-353>

^{1*}**Нодяров А.С.,**

PhD, ORCID ID: 0000-0002-0045-5457,

*e-mail: nodyarov.atilkhan@gmail.com

²**Расылова А.Е.,**

PhD, ORCID ID: 0000-0001-6964-8444,

e-mail: rassylova.ainash@gmail.com

¹**Омар А.Ж.,**

PhD, ORCID ID: 0000-0002-5604-3742,

e-mail: omaruzhan@gmail.com

¹Казахский национальный университет им. аль-Фараби,
г. Алматы, Казахстан

²Казахский национальный женский педагогический университет,
г. Алматы, Казахстан

СПЕКТРАЛЬНЫЙ АНАЛИЗ ЗВЕЗДЫ MWC 645**Аннотация**

В данной работе представлены результаты спектроскопических наблюдений звезды MWC 645 – представителя слабоизученного класса объектов типа FS CМа, характеризующегося наличием сильных эмиссионных линий и инфракрасного избытка, связанного с околозвездной пылью. На основе наблюдений высокого спектрального разрешения впервые обнаружен холодный вторичный компонент в системе. Определены фундаментальные параметры обоих компонентов: для горячей (В-типа) звезды – эффективная температура $18,000 \pm 2000$ К и светимость $\log(L/L_{\odot}) = 3.9 \pm 0.4$, для холодной (К-типа) – 4250 ± 250 К и $\log(L/L_{\odot}) = 3.1 \pm 0.4$. Система находится на расстоянии 6.5 ± 0.9 кпк и демонстрирует признаки активного взаимодействия, включая текущий перенос массы, о чем свидетельствуют сложные профили эмиссионных линий. Полученные результаты подтверждают двойственную природу MWC 645 и ее принадлежность к типу FS CМа. Работа подчеркивает необходимость дальнейших наблюдений для уточнения параметров системы и лучшего понимания процессов в околозвездной оболочке.

Ключевые слова: звезды, двойные звезды, спектральные двойные звезды, Ве звезды, звезды с эмиссионными линиями.

Введение

Класс звезд В[е]-типа характеризуется наличием разрешенных и запрещенных эмиссионных линий низкого возбуждения, таких как H I, Fe II, [Fe II], N II, O I и другие в спектрах

звезд В-типа в сочетании с сильным избытком в ближнем и среднем ИК-диапазоне из-за излучения околозвездной пыли. Явление наблюдается у звезд с широким диапазоном масс и эволюционных статуса. К последним относятся звезды Ae/BeХербига до главной последовательности, симбиотические двойные системы, протопланетные туманности и небольшая группа сверхгигантов (например, η Киля), как это было сделано авторами работы [1]. Также почти 30 объектов из первоначального списка звезд B[e] остались неклассифицированными из-за отсутствия наблюдений и сложных свойств, которые содержали предположений на более чем одну звездную группу с известной природой. Большинство из них стали членами новой группы типов FS CMa, определенной в работе [2], который предположил, что эти объекты, скорее всего, являются тесными двойными звездами, которые претерпевают или претерпели фазу массообмена между компонентами.

Природа большинства объектов данного типа до сих пор до конца не выяснена, хотя все они считаются подозреваемыми или подтвержденными двойными системами. В частности, звезды MWC 17, FS CMa, MWC 342 и MWC 645 входят в группу FS CMa. Среди них MWC 645 является одним из наименее изученных объектов и составляет основной предмет настоящего исследования. На рисунке 1 представлено положение звезды MWC 645 по данным цифрового обзора неба (DSS).

Авторы работы [3] изучили синюю область спектра MWC 645, используя спектры Куде (coude) со спектральным разрешением $20 \text{ \AA mm}^{-1}$, полученные в 1971 г. Они обнаружили, что все сильные эмиссионные линии FeII и [FeII] имеют двойной пик с разницей средней лучевой скорости (RV) в 150 км/с между более сильным красным и более слабым синим пиками, в то время как профили H γ и H δ демонстрируют три компонента.

Оптические спектры промежуточного разрешения MWC 645 в области от $3741 \text{ \AA}$ до $8790 \text{ \AA}$ с щелью между 5137 и $6332 \text{ \AA}$, снятые в 1991–1995 гг., были изучены авторами [4]. Эти авторы идентифицировали 88 запрещенных и разрешенных эмиссионных линий и измерили для этих линий среднюю гелиоцентрическую скорость (RV), равную 76 км/с , но не обнаружили линий поглощения, сформированных в звездной атмосфере. Сравнение их результатов с данными других авторов показало, что спектральные линии сильно различались по эквивалентной ширине (EW) вплоть до 2 раз, а также что некоторые линии не были обнаружены в спектрах, снятых в разное время.

Главная звезда показана крестом, звезды сравнение числами. Прямое восхождение звезды $21:53:27.486$, склонение звезды $+52:59:58.01$.

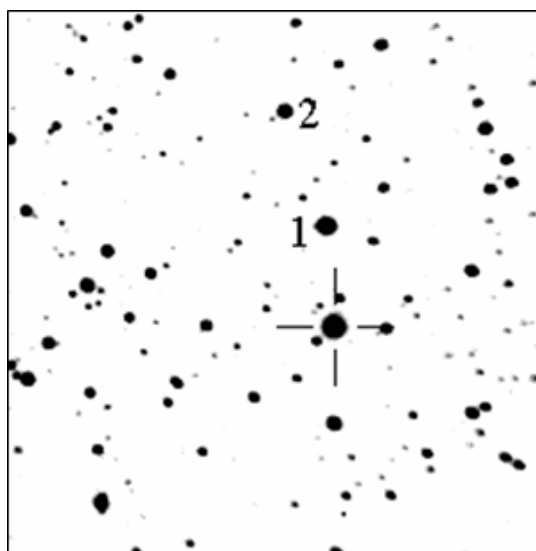


Рисунок 1 – Поле размером $10' \times 10'$ вокруг MWC 645 из цифрового обзора неба (DSS)

Ранее сообщалось, что профиль линии H α у звезды MWC 645 обладает необычной структурой, он состоит из широкого компонента с синим смещением и значительно более интенсивного узкого компонента с красным смещением. Это было замечено в спектрах низкого разрешения, полученных в работе [5] в 1981 г., а позже подтверждено в работе [6] автором, изучавшим узкие области спектра вокруг нескольких эмиссионных линий (например, H α и [OI] 6300 и 6363 Å) в 1986–1988 и 2000–2002 гг. Он пришел к выводу, что такой профиль не обнаружен ни у одного другого объекта выборки из 18 звезд B[e], наблюдавшихся им с высоким разрешением. Этот автор также описал профили всех разрешенных эмиссионных линий в спектрах MWC645 как имеющие характерную асимметрию с крутым красным флангом и крылом на синей стороне, в то время как запрещенные линии были узкими и однопиковыми.

Недавно авторы работы [7] исследовали двойную природу звезды типа FS CMa MWC 645 с помощью ближней инфракрасной и оптической спектроскопии из нескольких обсерваторий. Они идентифицировали различные атомные и молекулярные спектральные особенности и проанализировали их свойства. Молекулярные полосы поглощения CO (1,62 мкм и 2,3 мкм) были впервые обнаружены в MWC 645, что дало новое представление о вторичном компоненте. Полосы CO предполагают наличие компаньона G-типа с верхним пределом для его эффективной температуры 5200 К. Авторы также обнаружили изменения в интенсивности эмиссионной линии H α , коррелирующие с изменениями яркости, что указывает на события выброса массы или увеличенную потерю массы.

Поскольку все исследования этого объекта проводились давно и были безрезультатными в отношении его природы, MWC 645 требует более детального изучения. В частности, свойства околозвездного вещества, фундаментальные параметры, распределение энергии в спектре (SED) и кривая блеска подробно не изучались. В этой работе представляются результаты спектроскопии высокого и низкого разрешения звезды MWC 645, направленные на определение ее фундаментальных параметров и оценку ее природы и эволюционного статуса.

Материалы и методы

Для анализа использовались как фотометрические, так и спектроскопические методы. Фотометрические данные позволили оценить общие характеристики светимости объекта, выявить наличие инфракрасного избытка, а также проанализировать возможные вариации блеска во времени. Спектральный анализ был направлен на изучение профилей эмиссионных и абсорбционных линий, определение физических параметров звезды и характеристик околозвездной среды. Особое внимание уделялось эмиссионным линиям водорода и гелия, а также запрещенным линиям металлов, что позволило уточнить природу и структуру околозвездной оболочки.

Спектральные наблюдения звезды MWC 645 были получены на телескопах и инструментах, представленных в таблице 1.

Таблица 1 – Инструменты, использованные в спектральных наблюдениях

Обсерватория	Расположение	Диаметр телескопа, m	Диапазон Å	Разрешение	№
CFHT	Гавайи, США	3.6	3600 – 10500	65,000	4
McDonald	Техас, США	2.7	3600 – 10500	60,000	1
OAN-SPM	Сан-Педро-Мартир, Мексика	2.1	3600 – 7300	18,000	8
SAO	Нижний Архыз, Россия	6.0	5200 – 6800	60,000	1

В ходе исследования были получены оптические спектроскопические наблюдения звезды MWC 645 с высоким разрешением на 2.7-метровом телескопе Харлана Дж. Смита в обсерватории Макдональда (Техас, США, $R = 60\,000$) с помощью спектрографа Tull coude TS2 [8], 2.1-метрового телескопа Национальной астрономической обсерватории Сан-Педро-Мартин (ОАН СПМ, Нижняя Калифорния, Мексика, $R = 18\,000$) со спектрографом REOSC, 6-метровом телескопе Специальной астрофизической обсерватории (САО, Нижний Архыз, Россия, $R = 60\,000$) со спектрографом NES [9]) и 3.6-метровым канадско-французско-гавайском телескопе (CFHT, Маунакеа, Гавайи, США, $R = 65\,000$) со спектрополяриметром ESPaDOnS [10]. Наблюдения, полученные при CFHT, были обработаны программными пакетами Urena и Libre-ESpRIT [11]. Спектры других обсерваторий обрабатывались с помощью пакета echelle в IRAF. В таблице 2 представлена информация обо всех спектроскопических наблюдениях (14 спектров) с 2004 по 2021 гг.

Таблица 2 – Спектроскопические наблюдения

Дата	HJD	Эксп., сек	Диап., Å	Обс.
10.07.2004	3286.424	18000	5200-6800	SAO
10.12.2005	3655.633	900	4300-6800	SPM
12.21.2005	3725.595	1200	4050-10100	McD
11.17.2007	4421.634	6000	3900-6800	SPM
10.04.2008	4744.632	4800	4100-6900	SPM
10.08.2008	4748.635	6000	4000-6900	SPM
12.10.2008	4810.699	2400	4000-10000	CFHT
12.11.2008	4811.699	2400	4000-10000	CFHT
11.03.2009	5139.665	4800	4100-7100	SPM
10.18.2010	5487.748	6000	4100-7350	SPM
11.18.2012	6250.611	4800	3800-7350	SPM
10.18.2013	6583.705	3600	4550-8100	SPM
10.18.2016	7679.242	4686	3800-10450	CFHT
11.26.2021	9544.500a	10200	4000-10000	CFHT

Примечание: (1) Дата наблюдения (месяц/день/год). (2) Гелиоцентрическая юлианская дата (JD-2450000+). (3) Общее время экспозиции в секундах. (4) Информативный диапазон каждого спектра в Å. (5) Сокращенное наименование обсерватории.
а – Наблюдение CFHT в 2021 г. состоит из 10 комбинированных отдельных экспозиций, сделанных в течение трех ночей подряд (24–26 ноября 2021 г.).

Два спектра ближнего ИК-диапазона MWC645 с низким разрешением были получены на 3-метровом рефлекторе Шейна Ликской обсерватории (США, Калифорния, 0.46–2.5 мкм, $R \sim 700$) в июле 2005 г. и ноябре 2009 г. с помощью аэрокосмического спектрографа ближнего инфракрасного диапазона VNIRIS [12]. Охватываемый спектральный диапазон составлял 0.46–2.5 мкм в 2005 г. и 0.8–2.5 мкм в 2009 г.

Изучение закона межзвездного поглощения в направлении объекта

Для оценки межзвездного поглощения и покраснения звезды MWC 645 было использовано несколько методов.

Чтобы учесть межзвездное поглощение, были использованы опубликованные зависимости между эквивалентной шириной и диффузными полосами межзвездного поглощения (DIB) и избытком цвета $E(B-V)$. Калибровки авторов работ [13] и [14] для эквивалентные ширины

DIB-ов при 5780 Å и 6613 Å дают избыток цвета $E(B-V) = 1.0 \pm 0.1$ mag. Межзвездное поглощение $AV = 3.1 \times E(B-V) = 3.1$ mag.

Покраснение MWC 645 также можно оценить по соотношению линий O I при 0.8446, 1.1287 и 1.3165 мкм [15]. Из спектра Лика от 3 июля 2005 г. это дает $E(B-V) = 0.94$ mag для кривой покраснения из работы [16]. Оба этих метода дают почти одинаковый результат для избытка цвета, указывая на небольшое околозвездное покраснение.

Отдельным методом оценки межзвездного покраснения является использование 3D-карты покраснения на основе расстояний Gaia и фотометрических данных для проекционных близких звезд из каталогов Pan-STARRS 1 и 2MASS [17]. Это подразумевает покраснение $E(g-r)$; $E(B-V) = 0.78 \pm 0.03$ mag в направлении на объект (галактические координаты $l = 98^\circ$, $b = -1^\circ$) на расстоянии 6.5 ± 0.9 кпк.

Расхождение между методами может быть связано с разбросом данных DIB, собранных с широкого диапазона лучей зрения, или с отсутствием данных для звезд главной последовательности на расстояниях более 6 кпк, указанных на карте покраснения в направлении объекта. Последняя причина кажется более вероятной, поскольку покраснение, рассчитанное с помощью инструмента трехмерной карты, достигает $E(g-r) = 1.0$ mag всего в 0.2° в направлении галактической плоскости от положения объекта. Уверенность усиливается за счет очень хорошего совпадения значений $E(B-V)$, полученных с использованием методов DIB и линий покраснения.

Результаты и обсуждения

Эмиссионные линии звезды MWC 645

Оптический спектр MWC645 демонстрирует многочисленные эмиссионные линии, что согласуется с предыдущими исследованиями. Сильные разрешенные и запрещенные эмиссионные линии представлены водородом (серии Бальмера и Пашена), FeII, CrII, TiII, CaII, OI, NI, NI, CaII, [FeII], [NiII], [CaII], [OI] и несколько строк других элементов. RV эмиссионных линий определялись путем подгонки пиков линий к гауссиане. Линии были идентифицированы с использованием списка линий в звездных спектрах [18], а EW были измерены путем интегрирования в континуум-нормализованные спектры.

Список больше 226 эмиссионных линий и их лучевые скорости в диапазоне длин волн 3966–10407 Å, идентифицированных в спектре (CFHT) 2008 г., представлены в работе [19], где лучевые скорости двухпиковых линий связаны с наиболее интенсивным (смещенным в красную сторону) пиком, линии с синим смещением трудно измерить из-за их слабости в большинстве линий, за исключением самых сильных бальмеровских линий. Параметры эмиссионных линий H α и H β приведены в таблице 3.

Таблица 3 – Параметры линий H α и H β в спектрах MWC 645

H α						
Дата	Обс.	RV_blue	RV_red	I_blue	I_red	EW
10.07.2004	SAO	-250.4	-33.7	85.5	95.2	-561.9
10.12.2005	SPM	-222.5	-18.6	135.8	448.2	-1603.5
12.21.2005	McD	-195.1	2.9	-	-	-
11.17.2007	SPM	-221.6	-5.3	197.2	347.2	-1557.5
10.04.2008	SPM	-240.8	-21.8	207.6	292.4	-1833.9
10.08.2008	SPM	-241.2	-21.8	207.2	294.7	-1345.5
12.10.2008	CFHT	-261.3	-30.9	194.5	280.1	-1609.0

Продолжение таблицы 3

12.11.2008	CFHT	-260.9	-30.9	192.9	273.4	-1595.7
11.03.2009	SPM	-213.8	-7.6	180.3	483.4	-1765.9
10.18.2010	SPM	-222.5	-8.5	156.2	492.2	-1771.7
11.18.2012	SPM	-236.2	-8.5	298.8	359.6	-1951.4
10.18.2013	SPM	-215.2	-17.2	213.0	633.9	-1998.8
10.18.2016	CFHT	-241.7	-29.1	237.1	350.2	-1631.8
11.26.2021	CFHT	-244.9	-27.3	111.7	356.1	-1192.5
H β						
10.07.2004	SAO	-	-	-	-	-
10.12.2005	SPM	-	-25.4	-	15.9	-31.2
12.21.2005	McD	-	-1.4	-	20.9	-44.6
11.17.2007	SPM	-230.9	2.3	32.1	56.8	-93.7
10.04.2008	SPM	-252.5	-24.2	23.8	47.8	-66.4
10.08.2008	SPM	-254.4	-25.4	23.7	46.8	-81.2
12.10.2008	CFHT	-280.3	-35.3	37.8	50.1	-75.1
12.11.2008	CFHT	-280.9	-34.7	38.5	50.9	-75.6
11.03.2009	SPM	-278.4	-2.6	7.5	29.8	-30.6
10.18.2010	SPM	-195.7	-11.8	18.1	45.6	-67.8
11.18.2012	SPM	-243.3	-10.0	37.1	48.6	-83.8
10.18.2013	SPM	-	-	-	40.5	-61.8
10.18.2016	CFHT	-241.4	-29.7	33.8	45.2	-100.3
11.26.2021	CFHT	-282.1	-30.4	4.3	19.8	-70.7
Примечание: Информация в столбце: (1) Дата наблюдения (месяц/день/год). (2) Сокращенное название обсерватории. (3), (4) Гелиоцентрические RV или эмиссионные пики, смещенные в синюю (RVblue) и красную (RVred) сторону спектра. (5), (6) Интенсивности пиков в единицах континуума. (7) Эквивалентные ширины эмиссионных линий в Å. Следует отметить, что линия H α была близка к краю спектрального порядка в спектре 2005 г., полученном в обсерватории Макдональда для определения приемлемого уровня континуума.						

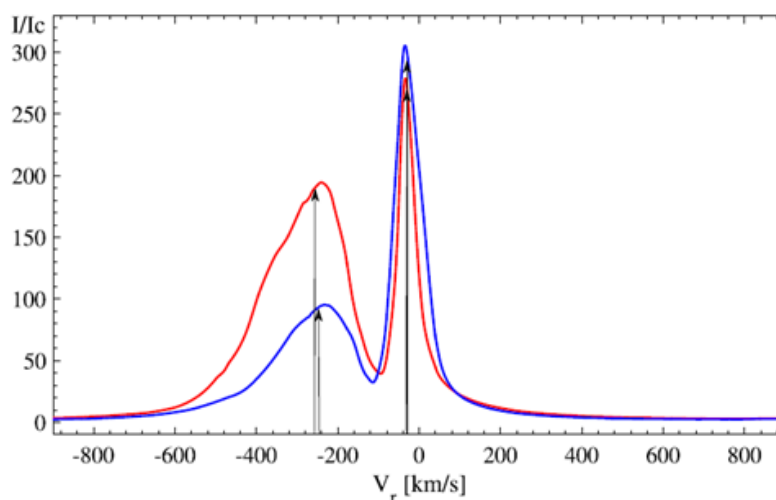
Средняя RV эмиссионных линий, определенная во всех наблюдаемых спектрах самого высокого разрешения, полученных в SAO (2004 г.), McD (2005 г.) и CFHT (2008, 2016, 2021 гг.), составляет -61.0 ± 4.3 км/с. Лучевые скорости (RV или) эмиссионных линий были определены по формуле эффекта Доплера (1):

$$V_r = \frac{\lambda_{obs} - \lambda_{lab}}{\lambda_{lab}} \cdot c \quad (1)$$

где, $\lambda_{obs} \cdot \lambda_{lab}$ – наблюдаемые и лабораторные длины волны, c – скорость света в вакууме в км/с.

Сравнивая спектры высокого разрешения и высокого отношения сигнал/шум, не было обнаружено исчезновения ни одной из спектральных линий, упомянутых [4]. Хотя в спектрах, полученных в разные годы, интенсивность линий разная, не были обнаружены сильные перемены лучевых скоростей эмиссионных линий в отдельных спектрах.

На рисунке 2 показан пример вариаций профиля линии H α . Его очень интенсивный и широкий пик с синим смещением имеет среднее значение RV -252 ± 9 км/с и ширину на полувысоте (FWHM) 224 ± 10 км/с, в то время как узкий пик с красным смещением имеет среднее значение RV -30 ± 2 км/с. и FWHM 60 ± 15 км/с. Используются все спектры с самым высоким разрешением, полученные в SAO и CFHT, для получения этих средних данных.



Стрелки показывают лучевые скорости пиков. Интенсивность нормирована на континуум, а лучевые скорости являются гелиоцентрическими.

Рисунок 2 – Сравнение профилей линии Na, полученных в 2008 г. (красная линия) и 2021 г. (синяя линия) в CFHT

Эти значения аналогичны значениям, указанным в работе [6]: $RV -218$ км/с и -30 км/с и FWHM 229 км/с и 59 км/с соответственно. Следовательно, пики профиля Na не претерпевают резких изменений за последние четыре десятилетия, что отражает относительную стабильность околозвездной среды. Параметры линий Na и H β представлены в таблице 3. Большинство разрешенных эмиссионных линий имеют сходные профили, тогда как расщепление пиков профилей запрещенных линий не очевидно.

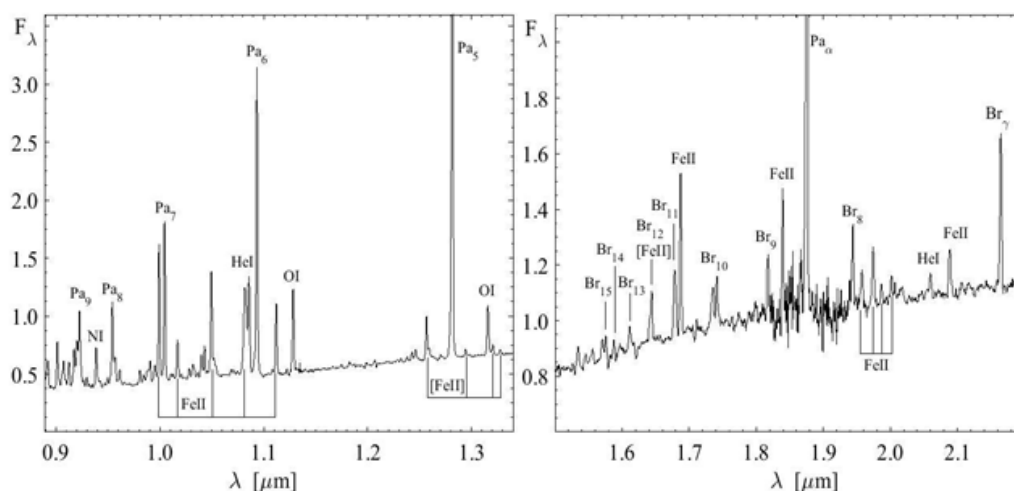
Эмиссионная линия He I $\lambda 6678$ Å, о наличии которой сообщается в работе [4] на основе спектра, полученного в 1995 г., отсутствовала в спектре 1987 г., представленного в работе [6]. В рамках настоящего исследования также не было обнаружено явных признаков присутствия данной линии в доступных наблюдательных спектрах.

Кроме того, автор работы [6] счел идентификацию эмиссионной линии около $\lambda 6665$ Å неопределенной, предполагая, что это может быть смесь линий FeI около $\lambda 6667$ Å. Однако в спектре MWC 645 присутствуют эмиссионные линии FeII, а не FeI. Эта линия была идентифицирована как линия [NiII] $\lambda 6666.80$ Å, которая наблюдается в спектрах некоторых других объектов типа FS CMa, таких как MWC 17 [20].

Эмиссионные линии, идентифицированные в спектрах ближнего ИК-диапазона, снятых в Ликской обсерватории по ранее опубликованным данным, снятым на том же оборудовании (например, [21, 22]), показаны на рисунке 3. Основной спектральный состав в диапазоне 0.8 – 2.5 мкм включает линии водорода серий Пашена и Брэкетта, а также линии FeII, NI и OI. Две эмиссионные линии HeI на 1.083 мкм и 2.058 мкм с умеренной интенсивностью присутствуют в обоих спектрах ближнего ИК, снятых в 2005 и 2009 гг.

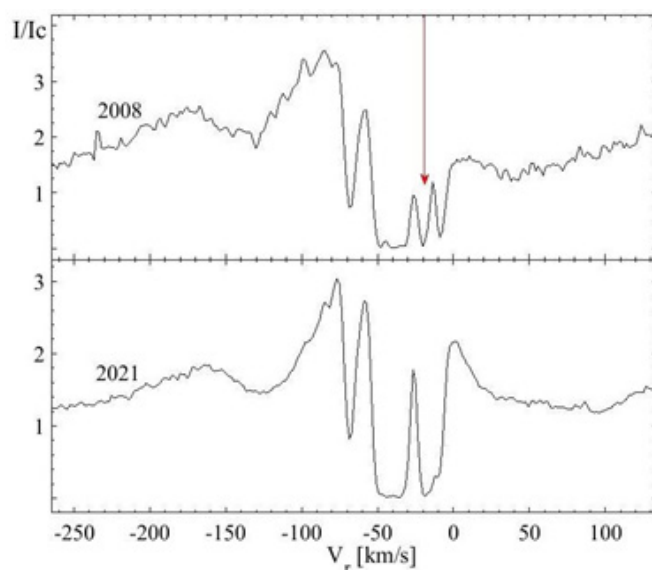
Линии поглощения звезды MWC 645

Спектр MWC645 содержит ряд линий поглощения, созданных в межзвездной среде, в том числе сильные диффузные межзвездные полосы (DIB), такие как 5780 , 5796 , 6613 Å. Профили линий NaI при 5889 и 5895 Å имеют сложную структуру (рисунок 4), состоящую из нескольких компонентов. Две более коротковолновые компоненты (узкая при $RV -70$ км/с и широкая при RV от -52 до -32 км/с) формируются, по-видимому, в межзвездной среде, а более длинноволновая (между RV от -22 до -12 км/с), положение которой близко к положению смещенного в красную сторону пика линии Na, изменяется по форме и может содержать некоторый вклад околозвездного вещества.



Длины волн даны в микронах, а потоки даны в единицах 10^{-12} Вт м $^{-2}$ мкм $^{-1}$.

Рисунок 3 – Части ближнего ИК-спектра MWC 645 с низким разрешением, полученные в Ликской обсерватории 3 июля 2005 г.



RV показаны относительно лабораторной длины волны линии (5889,92 Å). Интенсивность нормирована на континуум. Красная стрелка показывает среднее значение RV смещенного в красную сторону пика линии NaI.

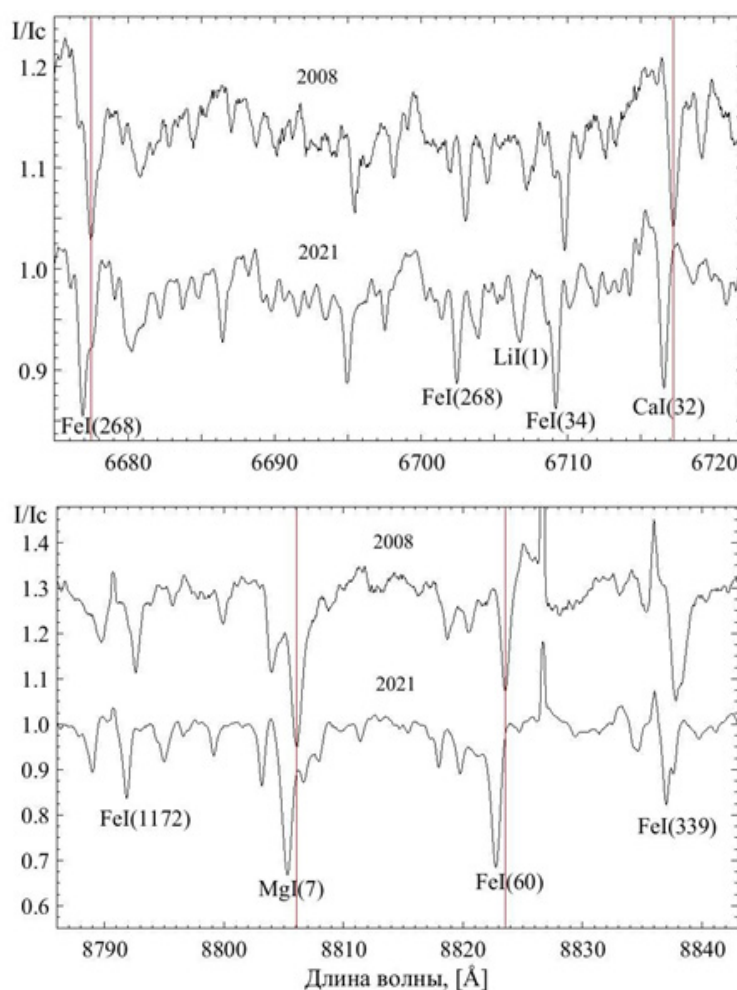
Рисунок 4 – Часть оптических спектров высокого разрешения MWC 645, полученных на CFHT в 2008 г. (верхняя панель) и 2021 г. (нижняя панель), которая показывает структуру линии NaI D2

Ни в одном из наблюдаемых спектров не обнаружено линий поглощения, типичных для атмосферы В-звезды. Это может быть связано с сочетанием высокой скорости вращения горячей звезды, вуалированы околосредным континуумом и относительно низким отношением сигнал/шум в синей области спектра, где эффекты вуали меньше. Однако впервые был обнаружен ряд слабых линий поглощения нейтральных металлов, типичных для спектров холодных звезд, в наблюдаемых спектрах с максимальным отношением сигнал/шум, снятых на CFHT в

2008 и 2021 гг. Эти линии теряются в шумах во всех остальных спектрах из нашей коллекции. Линии поглощения, идентифицированные по каталогу [18], вместе с RV показаны в работе [19].

Чтобы обеспечить надежную идентификацию, наблюдаемые спектры CFHT с самым высоким отношением сигнал-шум MWC645 были сравнены со спектрами MWC 623, двойной системы с компонентами В-типа и К-типа [23]. MWC 623 была включена [24] в первоначальный список звезд В[e] и признана членом группы FS CMa [2].

Средняя RV линий поглощения в спектре 2008 г. составляет -23.8 ± 3.6 км/с, тогда как в спектре 2021 г. она составляет -50.2 ± 3.7 км/с (рисунок 5). Этот результат может быть признаком орбитального движения холодного компонента. Линии поглощения в спектрах ближнего ИК-диапазона не обнаружены, скорее всего, из-за их вуалирования излучением околозвездной пыли.



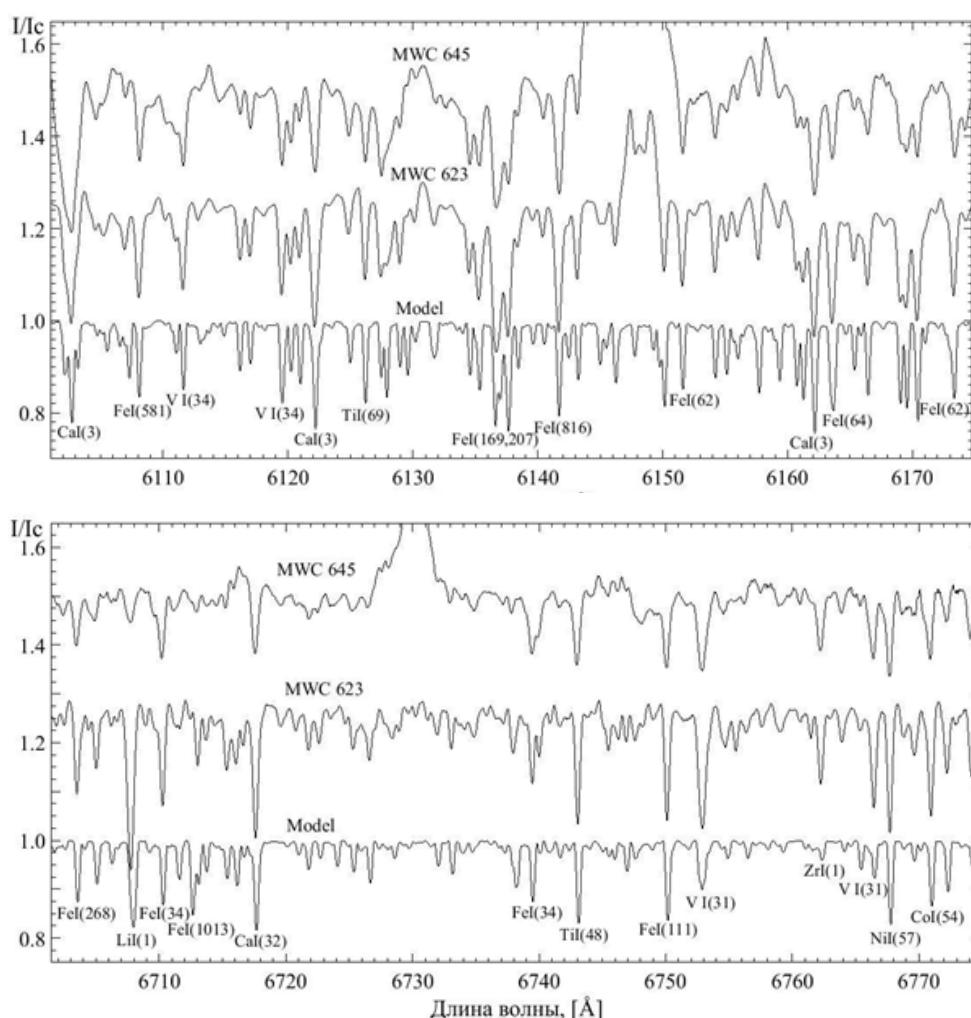
Вертикальные линии показывают положения некоторых линий в спектре 2008 г., чтобы обозначить сдвиг по сравнению с линиями в спектре 2021 г.

Рисунок 5 – Сравнение части спектров с линиями поглощения, снятыми 2008 и 2021 гг. в CFHT

Чтобы разделить вклады горячей и холодной компоненты, были рассчитаны синтетические спектры в нескольких узких диапазонах длин волн и искали наилучшее соответствие между наблюдаемыми и синтетическими спектрами, содержащими различные доли дополни-

тельного континуума. Последний представляет собой вклад горячего компонента без линий поглощения, которые не обнаруживаются в наблюдаемых спектрах. Такой подход оправдан пренебрежимо малым изменением вкладов компонент в диапазонах длин волн, используемых в данном анализе.

Для синтетических спектров были загружены модели атмосфер, предоставленные Ф. Каstellли, и рассчитаны спектры с помощью программы SPECTRUM, разработанной Gray & Corbally ([25]; рисунок 6) для T_{eff} между 3500 и 5000 K, $\log g = 1.5$ и 2.0, и проекции скорости вращения $v \sin i$ между 10 км/с и 20 км/с. Эти спектры сравнивались с наблюдаемыми спектрами, полученными на CFHT в 2008 и 2021 гг. Сравнивая интенсивности и ширины линий поглощения в синтетическом и наблюдаемом спектрах путем их деления и поиска наименьшей разницы между ними, были получены наиболее вероятные вклады компоненты. Относительные вклады горячей и холодной компонент оказались равными 85%/15% в области 5540–5580 Å, 75%/25% в области 6100–6175 Å и 70%/30% в области 6700–6775 Å соответственно.



Модель атмосферы была рассчитана для $T_{\text{eff}} = 4250$ K, силы тяжести на поверхности $\log g = 1.5$ и проецируемой скорости вращения $v \sin i = 10$ км/с. Вклад холодной компоненты составляет 25% в диапазоне 6100–6175 Å и 30% в диапазоне 6700–6775 Å. Интенсивность нормирована на локальный континуум, а шкала длин волн гелиоцентрическая.

Рисунок 6 – Сравнение спектров линий поглощения
MWC 645 и MWC 623 с модельным спектром

Заключение

Настоящая работа посвящена исследованию звезд с околозвездными оболочками, одной из наименее изученных категорий которых являются объекты типа FS CMa. Эта группа, включающая лишь несколько десятков известных представителей, характеризуется наличием ярких эмиссионных линий в оптических спектрах, включая запрещенные переходы, а также выраженным инфракрасным избытком, обусловленным присутствием околозвездной пыли. Несмотря на накопленные наблюдательные данные, природа формирования таких оболочек остается до настоящего времени неясной.

Исследования, проведенные в данной работе, привели к следующим основным результатам.

В работе представлены данные спектроскопических наблюдений высокого разрешения объекта типа FS CMa MWC 645. Основным результатом данного исследования является обнаружение холодного вторичного компонента.

Фундаментальные параметры MWC 645 составляют $T_{\text{eff}} = 18,000 \pm 2000$ К и светимость $\log(L/L_{\odot}) = 3.9 \pm 0.4$ для горячего компонента, $T_{\text{eff}} = 4250 \pm 250$ К и светимость $\log(L/L_{\odot}) = 3.1 \pm 0.4$ для холодного компонента, межзвездное поглощение $AV = 3.1 \times E(B-V) = 3.1$ mag, проекционная скорость вращения для холодного компонента $v \sin i = 10$ км/с, расстояние до звезды $D = 6.5 \pm 0.9$ кпк.

Результаты анализа спектральных данных ясно показывают, что MWC 645 представляет собой двойную систему, состоящую из звезд В-типа и К-типа, окруженную большим количеством околозвездного газа и пыли.

MWC 645 можно классифицировать как объект типа FS CMa на основании его двойственности и положения компонентов системы на HR-диаграмме. Хотя соответствующие эволюционные модели не точно описывают свойства компонентов, наблюдаемый спектр с сильными эмиссионными линиями согласуется с сильным текущим переносом массы между компонентами. Необычные профили многих эмиссионных линий (широкий пик с синим смещением и более сильный и узкий пик с красным смещением) могут служить дополнительным свидетельством продолжающегося переноса.

Более тщательное изучение, включающее более высокий уровень сигнала к шуму, спектроскопию высокого разрешения на больших телескопах и многоцветный фотометрический мониторинг, позволило бы определить больше свойств системы, таких как орбитальный период и наклонение, а также свойства ее околозвездной оболочки. В частности, ограничение цикличности возможных вариаций блеска помогло бы расширить теоретические исследования двойных систем с переносом массы на изучение структуры околозвездной среды.

ЛИТЕРАТУРА

- 1 Lamers, H.J., Zickgraf, F.J., de Winter, D., Houziaux, L., & Zorec, J. An improved classification of B[e]-type stars. *Astronomy and Astrophysics*, 340, 117–128 (1998).
- 2 Miroshnichenko, A.S. Towards understanding the B[e] phenomenon. I. Definition of the galactic FS CMa stars. *The Astrophysical Journal*, 667 (1), 497 (2007). <http://dx.doi.org/10.1086/520798>.
- 3 Swings, J.P., Allen, D.A. MWC 645 and MWC 819: Two Stars Resembling Eta Carinae. *Astrophysical Letters*, 14, 65 (1973).
- 4 Jaschek, M., Andrillat, Y., Jaschek, C. B[e] stars. III. MWC 645. *Astronomy and Astrophysics Supplement Series*, 120 (1), 99–105 (1996). <http://dx.doi.org/10.1051/aas:1996277>.
- 5 Swings, J.P., Andrillat, Y. Line widths in peculiar emission line objects. *Astronomy and Astrophysics*, 103, L3-L4 (1981).

- 6 Zickgraf, F.J. Kinematical structure of the circumstellar environments of galactic B[e]-type stars. *Astronomy & Astrophysics*, 408 (1), 257–285 (2003). <http://dx.doi.org/10.1051/0004-6361:20030999>.
- 7 Torres, A.F. et al. New insight into the FS CMa system MWC 645 from near-infrared and optical spectroscopy. *Galaxies*, 11 (3), 72 (2023). <http://dx.doi.org/10.3390/galaxies11030072>.
- 8 Tull, R.G., MacQueen, P.J., Sneden, C., & Lambert, D. L. The high-resolution cross-dispersed echelle white pupil spectrometer of the McDonald Observatory 2.7-m telescope. *Publications of the Astronomical Society of the Pacific*, 107 (709), 251 (1995). <http://dx.doi.org/10.1086/133548>.
- 9 Panchuk, V.E., Klochkova, V.G., Yushkin, M.V. The high-resolution Echelle Spectrograph of the 6-m telescope of the special astrophysical observatory. *Astronomy Reports*, 61 (9), 820–831 (2017). <http://dx.doi.org/10.1134/S1063772917080091>.
- 10 Manset, N., & Donati, J.-F., in *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, Vol. 4843, Proc. SPIE, ed. S. Fineschi, 2003, pp. 425–436
- 11 Donati, J.F., Semel, M., Carter, B.D., Rees, D.E., & Cameron, A.C. Spectropolarimetric observations of active stars. *Monthly Notices of the Royal Astronomical Society*, 291 (4), 658–682 (1997). <http://dx.doi.org/10.1093/mnras/291.4.658>.
- 12 Rudy, R.J., Puetter, R.C., Mazuk, S. Paschen lines and the reddening of the radio galaxy 3c 109. *The Astronomical Journal*, 118 (2), 666 (1999). <http://dx.doi.org/10.1086/300980>.
- 13 Herbig, G.H. The diffuse interstellar bands. IX-Constraints on the identification. *The Astrophysical Journal*, 407, 142–156 (1993). <http://dx.doi.org/10.1086/172500>.
- 14 Kos, J., Zwitter, T. Properties of diffuse interstellar bands at different physical conditions of the interstellar medium. *The Astrophysical Journal*, 774 (1), 72 (2013). <http://dx.doi.org/10.1088/0004-637X/774/1/72>.
- 15 Rudy, R., Erwin, P., Rossano, G.S., & Puetter, R.C. The 1.0-1.3 micron spectrum of LkH-alpha 101. *ApJ*, 383, 344–350 (1991). <http://dx.doi.org/10.1086/170792>.
- 16 Schlafly, E.F., Finkbeiner, D.P. Measuring reddening with Sloan Digital Sky Survey stellar spectra and recalibrating SFD. *The Astrophysical Journal*, 737 (2), 103 (2011). <http://dx.doi.org/10.1088/0004-637X/737/2/103>.
- 17 Green, G.M., Schlafly, E., Zucker, C., Speagle, J.S., & Finkbeiner, D. A 3D dust map based on Gaia, Pan-STARRS 1, and 2MASS. *The Astrophysical Journal*. 887 (1), 93 (2019). <http://dx.doi.org/10.3847/1538-4357/ab5362>.
- 18 Coluzzi, R. VizieR Online Data Catalog: Revised version of the ILLSS Catalogue (Coluzzi 1993–1999). *VizieR Online Data Catalog*, 1999, p.VI/71A.
- 19 Nodyarov, A.S. et al. VizieR Online Data Catalog: Toward Understanding the B [e] Phenomenon. IX (Nodyarov+, 2022). *VizieR Online Data Catalog*, 193, J/ApJ/936/129 (2024).
- 20 Klochkova, V.G., Chentsov, E.L. Detailed optical spectroscopy of the B[e] Star MWC17, *Astrophysical Bulletin*, 71(1), 33–42 (2016). <http://dx.doi.org/10.1134/S1990341316010041>.
- 21 Venturini, C.C., Rudy, R.J., Lynch, D.K., Mazuk, S., & Puetter, R.C. Near-Infrared Spectroscopy of Nova Sagittarius 1999 (V4444 Sagittarii). *The Astronomical Journal*, 124 (5), 3009 (2002). <http://dx.doi.org/10.1086/343832>.
- 22 Miroshnichenko, A.S., Levato, H., Bjorkman, K.S., Grosso, M., Manset, N., Men'shchikov, A.B., Rudy, R.J., Lynch, D.K., Mazuk, S., Venturini, C.C., Puetter, R.C., Perry, R.B. Properties of galactic B[e] supergiants-III. MWC 300. *Astronomy & Astrophysics*, 417 (2), 731–743 (2004).
- 23 Zickgraf, F.J. New spectroscopic observations of the B[e]/K binary system MWC 623. *Astronomy & Astrophysics*, 375 (1), 122–129 (2001). <http://dx.doi.org/10.1051/0004-6361:20010812>.
- 24 Allen, D.A., Swings, J.P. The spectra of peculiar Be stars with infrared excesses. *Astronomy and Astrophysics*, 47, 293–302 (1976).
- 25 Gray, R.O., Corbally, C.J. The calibration of MK spectral classes using spectral synthesis. 1: The effective temperature calibration of dwarf stars. *The Astronomical Journal*, 107, 742–746 (1994). <http://dx.doi.org/10.1086/116893>.

¹*Нодяров А.С.,

PhD, ORCID ID: 0000-0002-0045-5457,

*e-mail: nodyarov.atilkhan@gmail.com

²Расылова А.Е.,

PhD, ORCID ID: 0000-0001-6964-8444,

e-mail: rassylova.ainash@gmail.com

¹Омар А.Ж.,

PhD, ORCID ID: 0000-0002-5604-3742,

e-mail: omaruzhan@gmail.com

¹Әл-Фараби атындағы Қазақ ұлттық университеті, Алматы қ., Қазақстан²Қазақ ұлттық қыздар педагогикалық университеті, Алматы қ., Қазақстан

MWC645 ЖҰЛДЫЗЫНЫҢ СПЕКТРЛІК ТАЛДАУЫ

Аңдатпа

Бұл жұмыста жеткіліксіз зерттелген FS CMa типті нысандардың бірі – MWC 645 жұлдызына жүргізілген жоғары рұқсатты спектроскопиялық бақылаулардың нәтижелері ұсынылған. Бұл типтегі жұлдыздар күшті эмиссиялық сызықтарымен және жұлдыз маңындағы шаңның әсерінен пайда болатын инфрақызыл артық сәулеленуімен сипатталады. Алғаш рет бұл жүйеде салқын екінші компонент анықталды. Екі компоненттің де негізгі параметрлері анықталды: ыстық (B-типті) жұлдыз үшін тиімді температура $18\,000 \pm 2000$ К және жарқырауы $\log(L/L_{\odot}) = 3.9 \pm 0.4$; ал салқын (K-типті) жұлдыз үшін $T_{\text{eff}} = 4250 \pm 250$ К және $\log(L/L_{\odot}) = 3.1 \pm 0.4$. Жүйе 6.5 ± 0.9 кпк қашықтықта орналасқан және белсенді өзара әрекеттесудің айқын белгілерін көрсетеді. Бұл белгілер көптеген эмиссиялық сызықтардың күрделі профильдері арқылы байқалатын зат алмасу үдерістерімен түсіндіріледі. Алынған нәтижелер MWC 645 жұлдызының қосжұлдыздық табиғатын және оның FS CMa типіне жататынын растайды. Бұл жұмыс жүйенің параметрлерін нақтылау, сондай-ақ оның жұлдыз маңындағы ортасының құрылымы мен шығу тегін тереңірек түсіну үшін әрі қарайғы бақылаулардың қажеттілігін көрсетеді.

Тірек сөздер: жұлдыздар, қос жұлдыздар, спектралды қос жұлдыздар, Be жұлдыздар, эмиссиялық сызықты жұлдыздар.

¹*Nodyarov A.S.,

PhD, ORCID ID: 0000-0002-0045-5457

*e-mail: nodyarov.atilkhan@gmail.com

²Rassylova A.E.,

PhD, ORCID ID: 0000-0001-6964-8444

e-mail: rassylova.ainash@gmail.com

¹Omar A.Zh.,

PhD, ORCID ID: 0000-0002-5604-3742

e-mail: omaruzhan@gmail.com

¹Al-Farabi Kazakh National University, Almaty, Kazakhstan²Kazakh National Women's Pedagogical University, Almaty, Kazakhstan

SPECTRAL ANALYSIS OF THE STAR MWC 645

Abstract

This study presents results of high-resolution spectroscopic observations of the star MWC 645, a representative of the poorly studied FS CMa-type objects. These stars are characterized by strong emission lines and significant infrared excess caused by circumstellar dust. For the first time, a cool secondary component has been detected in this system. Fundamental parameters were determined for both components: for the hot (B-type) star, an effective

temperature of $18,000 \pm 2000$ K and luminosity $\log(L/L_{\odot}) = 3.9 \pm 0.4$; and for the cool (K-type) star, $T_{\text{eff}} = 4250 \pm 250$ K and $\log(L/L_{\odot}) = 3.1 \pm 0.4$. The system is located at a distance of 6.5 ± 0.9 kpc and shows clear signs of active interaction, including ongoing mass transfer indicated by the complex profiles of many emission lines. The results confirm the binary nature of MWC 645 and its classification as an FS CMa-type object. This work highlights the need for further observations to refine the system's parameters and improve our understanding of the structure and origin of its circumstellar environment.

Keywords: stars, binary stars, spectroscopic binary stars, Be stars, Emission line stars.

Дата поступления статьи в редакцию: 17.07.2025

UDC 533.9.01; 533.9...1
IRSTI 29.27.03; 29.27.47

<https://doi.org/10.55452/1998-6688-2025-22-4-354-364>

^{1,2,3*}**Ismagambetova T.N.,**

PhD, Senior Lector, ORCID ID: 0000-0003-4889-7526,

*e-mail: ismagambetova@physics.kz

^{1,2,4}**Muratov M.M.,**

PhD, Associate Professor, ORCID ID: 0000-0001-7270-9834,

e-mail: mukhit.muratov@gmail.com

⁵**Ussenov Y.A.,**

PhD, Associate Research Physicist, ORCID ID: 0000-0001-7903-8674,

e-mail: yussenov@pppl.gov

^{2,3,4}**Gabdullin M.T.**

PhD, Cand.Phys.-Math.Sc., Professor, ORCID ID: 0000-0003-4853-3642,

e-mail: gabdullin@physics.kz

¹National Nanotechnology Laboratory of Open Type (NNLOT),

Al-Farabi Kazakh National University, Almaty, Kazakhstan

²Institute of Applied Sciences and IT, Almaty, Kazakhstan

³Kazakh British Technical University, Almaty, Kazakhstan

⁴Kazakh Physical Society, Almaty, Kazakhstan

⁵Princeton Plasma Physics Laboratory, Princeton, NJ, USA

EFFECT OF IONIC CORE ON THE PROPERTIES OF NON-ISOTHERMAL PLASMAS

Abstract

This study investigates the effects of ionic cores on non-isothermal plasmas using a novel ion-ion interaction potential that incorporates screening effects from both ion cores and exchange-correlation interactions. Our findings indicate that with increasing distance, the effective potential approaches a Yukawa-like screening potential, while at shorter distances, strong electron binding weakens screening. The different values of the cutoff radius r_{cut} and core edge steepness α significantly influence the potential behavior and radial distribution functions (RDFs). Higher coupling parameters (Γ_i) strengthen the electron-ion interactions, leading to deeper potential wells and more pronounced non-ideality corrections. Increasing Θ decreases the absolute values of non-ideality corrections, indicating fewer interactions in the system. A larger cutoff radius r_{cut} at a fixed parameter α also reduces corrections due to weaker screening effects. As Γ_i increases, non-ideality corrections grow, reflecting stronger coupling. The results show the importance of taking into account the ion core effects in dense non-isothermal plasma research.

Keywords: dense plasmas, ion core, non-isothermal plasmas, ion-ion interaction potential, radial distribution functions, thermodynamic properties.

Introduction

The dense plasmas, characterized by a significant mass difference between ions and electrons, show non-isothermal behavior, making theoretical modeling more complex due to the slower response of ions to perturbations [1]. To model interactions accurately, researchers use the generalized Poisson-Boltzmann equation and derive effective interaction potentials from dielectric response functions [2–9].

Charge screening plays a key role in such dense plasmas, forming a polarization cloud of electrons that surrounds ions, compensating for the Coulomb potential. Excluding core electrons can be modeled with an empty core electron-ion pseudo-potentials [10–11]. Recent studies show that

ionic cores affect electron-ion scattering transport and temperature relaxation with strongly bound electrons altering the effective charge of ions and impacting the plasma's structural and thermodynamic properties for novel screened electron-ion, and the novel screened interaction potential for ion-ion pair [12–16].

These results showed the importance of considering ionic core effects in understanding dense plasma properties. Therefore, in this study, we want to continue the previous research based on the novel ion-ion interaction potential by analyzing the impact of the ionic core effects in non-isothermal plasmas. In this potential, both the screening effects due to the ion cores and from the exchange-correlation on the structural and thermodynamic properties are taken into account.

The paper is organized as follows: in Section 2, we describe the method used to compute the radial distribution functions (RDFs) and non-ideality corrections to the thermodynamic properties (internal energy and equation of state). In Section 3, the results and discussion of these calculations are presented.

Materials and methods

Ion–Ion Interaction Potential

In work (p. 2), a novel ion-ion screened potential that takes into account the ion core effect using a pseudo-potential model with strongly bound electrons was introduced by numerically solving the following formula [16–17]:

$$\Phi_{ii}(r) = \frac{Z^2 e^2}{r} + \int \frac{d^3 k}{(2\pi)^3} |\widetilde{\varphi_{ei}}(k)|^2 \chi_e(k) e^{i\vec{k}\vec{r}}, \quad (1)$$

where as a micro potential $\varphi_{ei}(r)$ the potential with soft empty cores, proposed by Gericke et al. was used [18]:

$$\varphi_{ei}(r) = \frac{Z e^2}{r} \left[1 - \exp\left(-\frac{r^\alpha}{r_{cut}^\alpha}\right) \right], \quad (2)$$

where the parameter α controls the core edge steepness ($\alpha = 2$ and 6), and r_{cut} is the core's cutoff radius ($r_{cut} = 0.75 \text{ \AA} \simeq 1.42 a_B$, where $a_B = \hbar^2/(m_e e^2)$ is the first Bohr radius) – for beryllium ion. Their values align with Kohn-Sham DFT (Density Functional Theory), which accurately describes dense plasmas in partial degeneracy, while other values can be used for various interaction regimes.

$\chi_e(k)$ is the static electron-density response function which we used with the parameter γ in the form from work that corrects for non-ideality in electronic screening due to static local field effects, equivalent to the local density approximation in DFT, accurately describing screening in the uniform electron gas (UEG) and modeling the exchange-correlation part of the electron free energy density, $f_{xc}(n, T_e)$ [19–21]. The importance and novelty of use in the interaction potential of parameter γ were explained in more detail in the work (p. 3) [16].

Therefore, we used the following form of the inverse screening length:

$$k_e^2 = \frac{k_{id}^2}{1 - k_{id}^2 \gamma}, \quad k_{id}^2 = k_{TF}^2 \theta^{\frac{1}{2}} I_{-\frac{1}{2}}(\eta)/2, \quad \gamma = -\frac{k_F^2}{4\pi e^2} \frac{\partial^2 n f_{xc}(n, T_e)}{\partial n^2}, \quad (3)$$

where $I_{-1/2}$ is the Fermi integral of the order $-1/2$, $\eta = \mu/k_B T_e$ is the reduced chemical potential, $k_{TF}^2 = 3\omega_p^2/v_F$ is the Thomas-Fermi wavenumber (ω_p is the plasma frequency, v_F is the Fermi velocity).

The following dimensionless parameters were used for non-isothermal plasma ($T_e \neq T_i$): the ion coupling parameter $\Gamma_{ii} \leq 10$ ($\Gamma_{ii} = \frac{Z_i^2 e^2}{a_{kB} T_i} \left(\frac{n_i}{n_e}\right)^{1/3} = \Gamma_{ee} Z_i^{5/3} \left(\frac{T_e}{T_i}\right) = \Gamma_{ei} Z_i^{2/3} \sqrt{\frac{T_e}{T_i}}$), where

$T_{ee} = T_e, T_{ii} = T_i$ and $T_{ei} = \sqrt{T_e T_i}$ [22], the degeneracy parameter $\Theta = \frac{k_B T}{E_F}$, where E_F is the Fermi energy, the density parameter $r_s = \frac{a}{a_B}$, where $a = (3/(4\pi n))^{1/3}$ is the mean inter-electronic distance (Wigner-Seitz radius), $R = r/a_B$ is measured in units of the first Bohr radius a_B .

To analyze the impact of the ion core, the potential (1) was compared with another screened Coulomb potential, known as the Yukawa potential, taking into account the exchange-correlation effect as described in equation (3) but without the ion core effect, and therefore, it will be denoted as simply the screened potential:

$$\Phi_Y(r) = \frac{Ze^2}{r} \exp(-k_e r). \quad (4)$$

The radial distribution functions of ions

The influence of the ionic core on structural properties was analyzed through the integral Ornstein-Zernike equation [23]:

$$h(r) = c(r) + n \int c(r') h(|\vec{r} - \vec{r}'|) d\vec{r}', \quad (5)$$

in the hyper-netted chain (HNC) approximation:

$$g_{HNC}(r) = \exp[-\beta\Phi(r) + h(r) - c(r)], \quad (6)$$

where $\beta = (k_B T)^{-1}$ and $\Phi(r)$ is the interaction potential, $h(r)$ and $c(r)$ are the total and direct correlation functions. The details about solving the equations (5–6) can be found in works [20, 24–25] (p. 14).

The radial distribution functions (RDFs) calculated in the HNC approximation use an effective screened ion-ion potential that accounts for the use of the electron polarization function but does not double-count ion screening. This is confirmed by comparison with MD results in work [26], where both methods obtain similar RDFs.

Thermodynamic properties

The internal energy and equation of state were obtained using potential (2) and RDFs (5–6) [27]:

$$E = E_{id} - \pi \sum_{\alpha=i,e} n_{\alpha} \sum_{\beta=i,e} n_{\beta} \int_0^{\infty} g^{\alpha\beta}(r) \varphi^{\alpha\beta}(r) r^2 dr = E_{id} - \Delta E, \quad (7)$$

$$P = P_{id} - \frac{2}{3} \pi \sum_{\alpha=i,e} n_{\alpha} \sum_{\beta=i,e} n_{\beta} \int_0^{\infty} \frac{\partial \varphi_{\alpha\beta}(r)}{\partial r} g_{\alpha\beta}(r) r^3 dr = P_{id} - \Delta P, \quad (8)$$

where N is the number of particles in the system, $E_{id} = 3/2 N k_B T$ is the internal energy of an ideal plasma, $P_{id} = n k_B T$ is the equation of state of an ideal plasma, and $\varphi(r)$ is the micropotential used as the basis of the considered screened interaction potential.

Results and discussion

The screened interaction potential for ion–ion pair for non-isothermal plasma, calculated using Equation (1), based on the potential (2) to account for the ionic core and the density response function of the UEG for plasma electron screening.

Figures 1 and 2 show the interaction potentials for i-i pairs at $r_s = 1$. In Figure 1 the results for potential (1) at $\theta = 1$ and $\Gamma_i = 5$ are presented: the black solid line represents $\alpha=2$ with $r_{cut} = 0.75$,

while the black dashed line corresponds to $\alpha=6$ with the same r_{cut} . Additionally, the red solid line indicates $\alpha=2$ with $r_{cut} = 0.25$, and the red dashed line represents $\alpha=6$ for the same r_{cut} . The blue solid line shows the results for the screening potential (4).

In Figure 2 the interaction potentials (1) are displayed with various parameters. The black solid line ($\alpha=2$) and dashed line ($\alpha=6$) represent $\theta = 1$, $\Gamma_i = 1$, and $r_{cut} = 0.25$. The red solid and dashed lines correspond to $\theta = 1$, $\Gamma_i = 5$, both with $r_{cut} = 0.25$. Similarly, the green solid and dashed lines show $\theta = 1$, $\Gamma_i = 10$, also with $r_{cut} = 0.25$. Finally, the dotted lines represent the results for $\theta = 0.5$, $\alpha=2$, and $r_{cut} = 0.25$, with black indicating $\Gamma_i = 1$, red for $\Gamma_i = 5$, and green for $\Gamma_i = 10$.

Figures 1 and 2 show that with increasing distance, potential (1) tends to the screening potential (4) (Yukawa), while at smaller distances, the strongly bound electrons shield the ion core, weakening the screening compared to the Yukawa potential. Increasing the r_{cut} at a fixed value of α leads to a softer potential (1), while decreasing α at a fixed r_{cut} results in a steeper potential (1) at short distances.

Higher Γ values indicate stronger coupling and deeper potential wells, meaning electron-ion interactions in the micro potential used as the basis become more significant as the system deviates from ideal gas behavior.

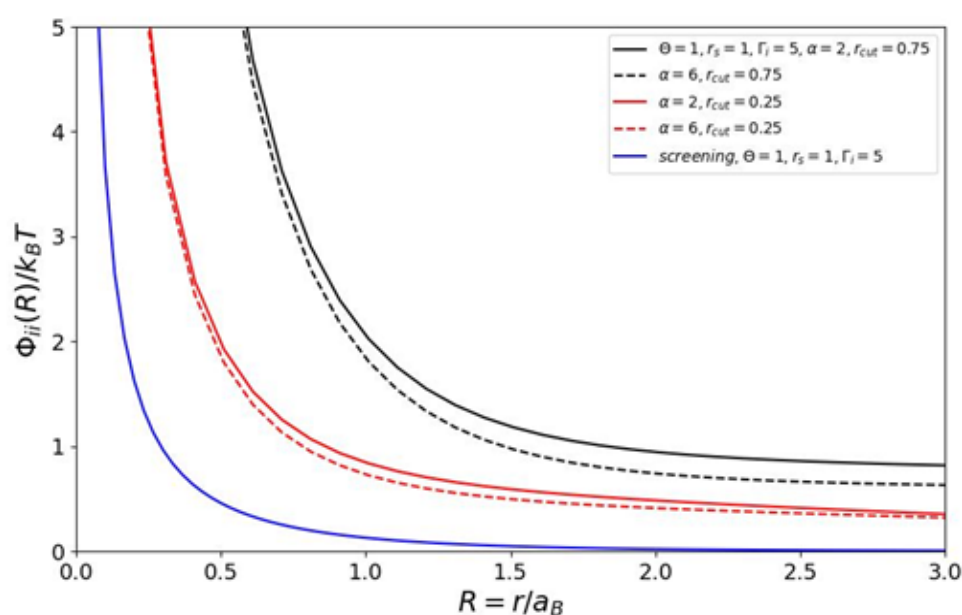


Figure 1 – The interaction potentials for i-i pairs at $r_s = 1$ for potential (1) $\theta = 1$, $\Gamma_i = 5$,
black solid line – $\alpha=2$, $r_{cut} = 0.75$, black dashed line – $\alpha=6$, $r_{cut} = 0.75$,
red solid line – $\alpha=2$, $r_{cut} = 0.25$, red dashed line – $\alpha=6$, $r_{cut} = 0.25$,
blue solid line – for screening potential (4)

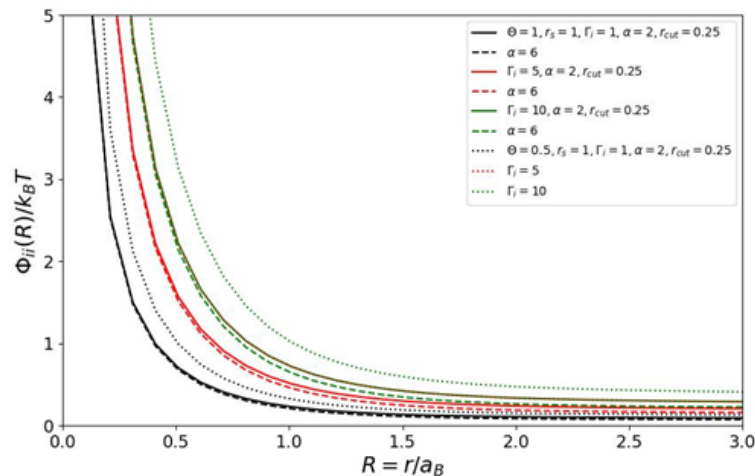


Figure 2 – The interaction potentials for i-i pairs at $r_s = 1$ for potential (1) black solid ($\alpha=2$) and dashed ($\alpha=6$) line – $\theta = 1, \Gamma_i = 1, r_{cut} = 0.25$, red solid ($\alpha=2$) and dashed ($\alpha=6$) line – $\theta = 1, \Gamma_i = 5, r_{cut} = 0.25$, green solid ($\alpha=2$) and dashed ($\alpha=6$) line – $\theta = 1, \Gamma_i = 10, r_{cut} = 0.25$, dotted lines represent results for $\theta = 0.5, \alpha=2, r_{cut} = 0.25$ – black ($\Gamma_i = 1$), red ($\Gamma_i = 5$), green ($\Gamma_i = 10$)

Figure 3 shows the interaction potentials for ion-ion (i-i) pairs at a degeneracy parameter $\theta = 0.5$ and a density parameter $r_s = 1$, with $\alpha = 2$ for potential (1). The interaction potentials are plotted for various ratios of electron temperature to ion temperature (T_e/T_i). Specifically, the black solid line represents $T_e/T_i = 1$ with a cut-off radius $r_{cut} = 0.25$, while the black dashed line corresponds to $r_{cut} = 0.75$. The red lines depict $T_e/T_i = 2$, with the solid line for $r_{cut} = 0.25$ and the dashed line for $r_{cut} = 0.75$. Similarly, the blue lines represent $T_e/T_i = 3$, with the solid line again for $r_{cut} = 0.25$ and the dashed line for $r_{cut} = 0.75$.

As the ion temperature T_i decreases, the screening in the potentials increases. This increase is larger for a larger cut-off radius of $r_{cut} = 0.75$ at a constant $\alpha = 2$. This behavior is attributed to a strengthening in the attraction between bound electrons and ions. In contrast, as the values of T_e/T_i increase, indicating stronger ion coupling, the screening of the potentials weakens.

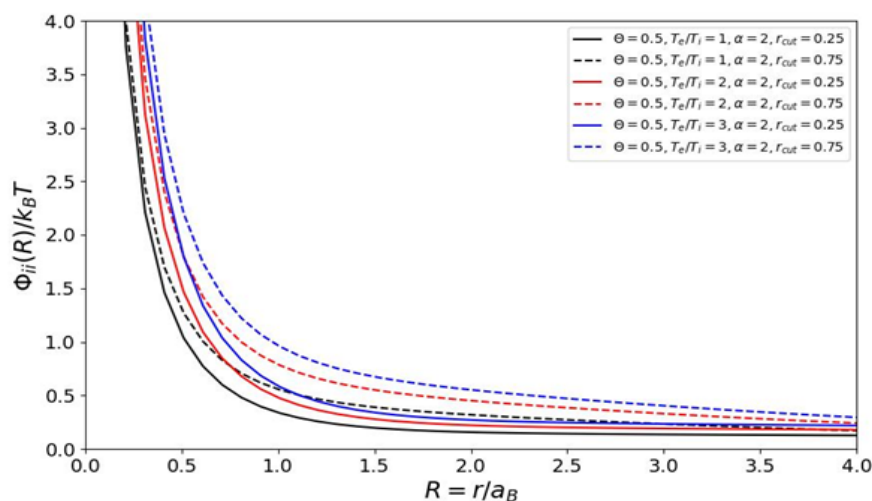


Figure 3 – The interaction potentials for i-i pairs at $\theta = 0.5, r_s = 1$ with $\alpha=2$ for potential (1) for different values of $\frac{T_e}{T_i}$: $\frac{T_e}{T_i} = 1$ – black solid ($r_{cut} = 0.25$) and dashed ($r_{cut} = 0.75$) line, $\frac{T_e}{T_i} = 2$ – red solid ($r_{cut} = 0.25$) and dashed ($r_{cut} = 0.75$) line, $\frac{T_e}{T_i} = 3$ – blue solid ($r_{cut} = 0.25$) and dashed ($r_{cut} = 0.75$) line

Figures 4 and 5 show the radial distribution functions (RDFs) for i-i pairs at $r_s = 1$. In Figure 3 the results for potential (1) at $\theta = 1$ are depicted with varying values of Γ_i and r_{cut} . For $\Gamma_i = 1$ and $r_{cut} = 0.25$, the black solid line corresponds to $\alpha=2$, while the dashed line represents $\alpha=6$. For $\Gamma_i = 5$ with the same r_{cut} , the lines are shown in red. When $\Gamma_i = 10$ is considered, the green lines indicate the results for $r_{cut} = 0.25$ with the same line styles. Additionally, the blue lines show the screening potential (4), with solid lines for $\Gamma_i = 1$, dashed lines for $\Gamma_i = 5$, and dotted lines for $\Gamma_i = 10$.

In Figure 5 the RDFs for potential (1) at $\Gamma_i = 10$ and $\theta = 0.5$ are shown. Here, the black solid lines ($\alpha=2$) represent $r_{cut} = 0.25$, while the dashed lines correspond to $r_{cut} = 0.75$. The red solid lines for $\theta = 1$ and $r_{cut} = 0.25$ are paired with dashed lines for $r_{cut} = 0.75$. The blue lines again depict the screening potential (4), with solid lines for $\theta = 0.5$ and dashed lines for $\theta = 1$.

In Figures 4 and 5 the radial distribution function (RDF) curves show how decreasing the r_{cut} strengthens the screening effects. The RDF for $r_{cut} = 0.25$ is positioned higher than that for $r_{cut} = 0.75$, indicating stronger attractive interactions between electrons and ions due to deeper negative minimum values of the non-screened electron-ion potential. Conversely, for fixed r_{cut} , increasing α shifts the potential minimum closer to the ion, making it steeper and resulting in a slightly higher RDF for $\alpha=6$ compared to $\alpha=2$. However, the difference between these RDFs is minor, while the distinction between potential (1) and the screening potential (4) is more pronounced, as the Yukawa model neglects the ion core effect, causing its RDFs (blue lines) to reach unity more quickly.

Because this is a non-isothermal plasma, RDF peaks increase with higher Γ_i values due to the difference between electron and ion temperatures, but overall curves with lower Γ_i values lie higher, as with the increasing coupling parameter, the probability of finding particles with the same charge (i.e., ions) decreases. And the difference in screening effects becomes more pronounced at lower θ values, indicating stronger coupling and strengthened electron-ion interactions.

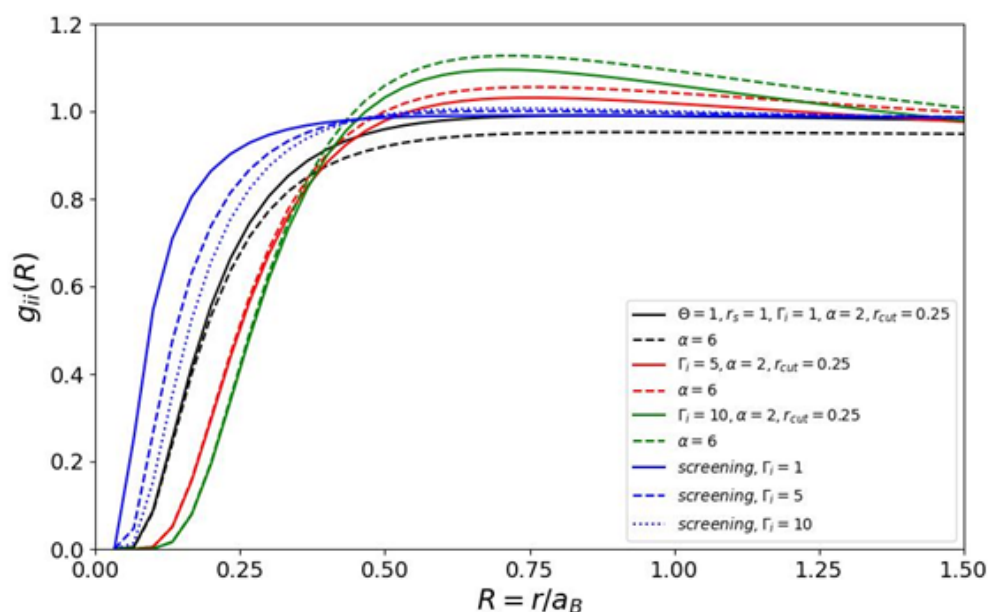


Figure 4 – The radial distribution functions for i-i pairs at $r_s = 1$ for potential (1) at $\theta = 1$ with $\Gamma_i = 1$, $r_{cut} = 0.25$ – black solid ($\alpha=2$) and dashed ($\alpha=6$) lines, $\Gamma_i = 5$, $r_{cut} = 0.25$ – red solid ($\alpha=2$) and dashed ($\alpha=6$) lines, $\Gamma_i = 10$, $r_{cut} = 0.25$ – green solid ($\alpha=2$) and dashed ($\alpha=6$) lines, blue lines – for screening potential (4) with solid ($\Gamma_i = 1$), dashed ($\Gamma_i = 5$) and dotted ($\Gamma_i = 10$)

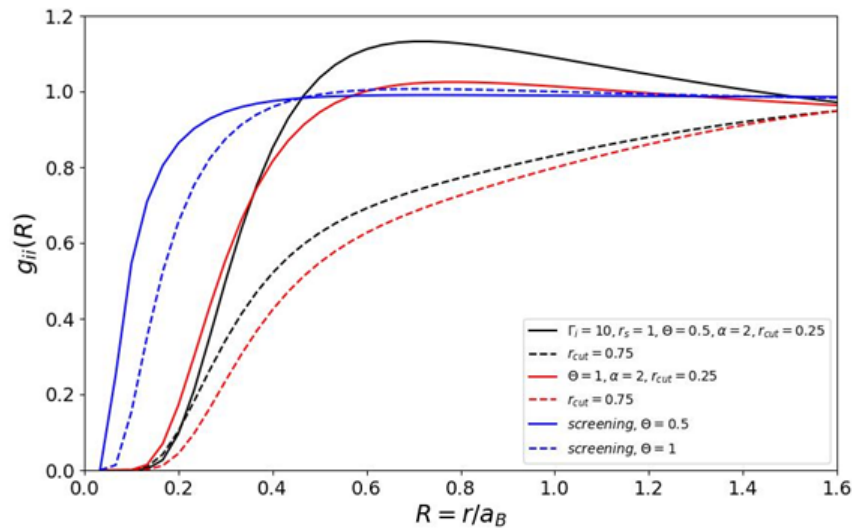


Figure 5 – The radial distribution functions for i-i pairs at $r_s = 1$ for potential (1) at $\Gamma_i = 10$ with $\theta = 0.5$, $\alpha = 2$ – black solid ($r_{cut} = 0.25$) and dashed ($r_{cut} = 0.75$) lines, $\theta = 1$, $\alpha = 2$ – red solid ($r_{cut} = 0.25$) and dashed ($r_{cut} = 0.75$) lines, blue lines – for screening potential (4) with solid ($\theta = 0.5$) and dashed ($\theta = 1$)

Figures 6 and 7 show correlation energy and equation of state (i.e., the ionic non-ideality corrections to the thermodynamic properties), respectively, calculated by solving equations (7–8) for the potential (1) and using potential (2) as micro potential $\varphi(r)$. The correlation energy and equation of state depend on the ratio of electron to ion temperatures, T_e/T_i , at a density parameter $r_s = 1$. Solid lines represent the case for $\theta = 1$, while dashed lines correspond to $\theta = 0.5$. The graphs include different values of α and r_{cut} : black lines indicate an interaction at $\alpha = 2$ with $r_{cut} = 0.25$, while red lines represent the same $\alpha = 2$ but with $r_{cut} = 0.75$. Additionally, green lines show the impact of increasing the $\alpha = 6$ with a $r_{cut} = 0.25$, and blue lines show the $r_{cut} = 0.75$.

As θ (and consequently temperature T) rises, the absolute values of non-ideality corrections decrease, showing a reduction in interactions within the system. The lower absolute values of the corrections to non-ideality are also observed for a larger cutoff radius of $r_{cut} = 0.75$ at a fixed α , as well as for a smaller $\alpha = 2$ at a fixed cutoff, due to a decrease in screening effects as the attraction between bound electrons and ions weakens. Conversely, as the values of Γ_i increase, signifying stronger coupling, the absolute values of the non-ideality corrections also rise.

Conclusions

In this study, we employed a novel ion-ion interaction potential that accounts for both screening from ionic cores and exchange-correlation interactions to analyze how the ion core effects impact the structural and thermodynamic properties of non-isothermal plasmas.

It should be noted that this study prioritizes investigating ion core effects in non-isothermal plasma over broad comparisons with other theoretical models or experimental data, as previous research has adequately addressed such comparisons [15, 16]. Our focus allows for a deeper exploration of ion-ion interactions without added complexity. While we did not include calculations with and without exchange-correlation interactions, prior work [15] has established that the differences in screening effects decrease at higher θ , becoming negligible at $\theta = 4$, which means in the regime of weakly or non-degenerate electrons, the effects of electronic exchange and correlation must be taken into account.

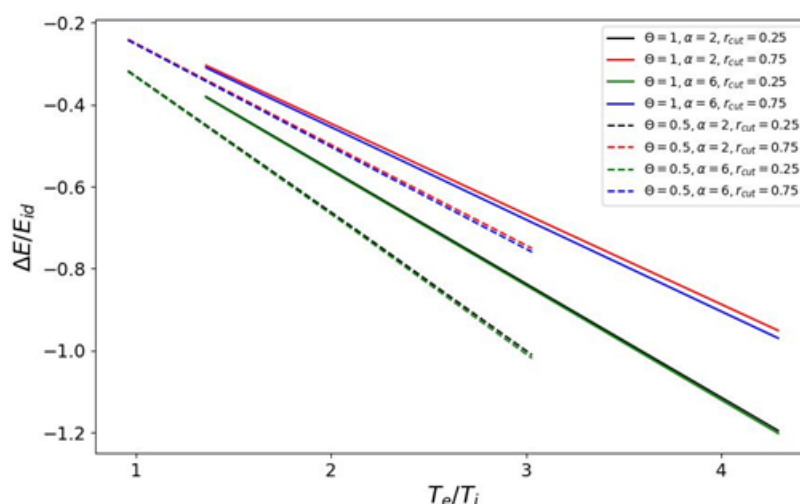


Figure 6 – The correlation energy for the potential (1) depending on relation between temperatures of electrons and ions $\frac{T_e}{T_i}$ using values of $\Gamma_i = 1, 5, 10$, at $r_s = 1$ for $\theta = 1$ ($T_e = 580\,000\text{ K}$, solid lines) and $\theta = 0.5$ ($T_e = 290\,000\text{ K}$, dashed lines): $\alpha=2$, $r_{cut} = 0.25$ – black lines, $\alpha=2$, $r_{cut} = 0.75$ – red lines, $\alpha=6$, $r_{cut} = 0.25$ – green lines, $\alpha=6$, $r_{cut} = 0.75$ – blue lines

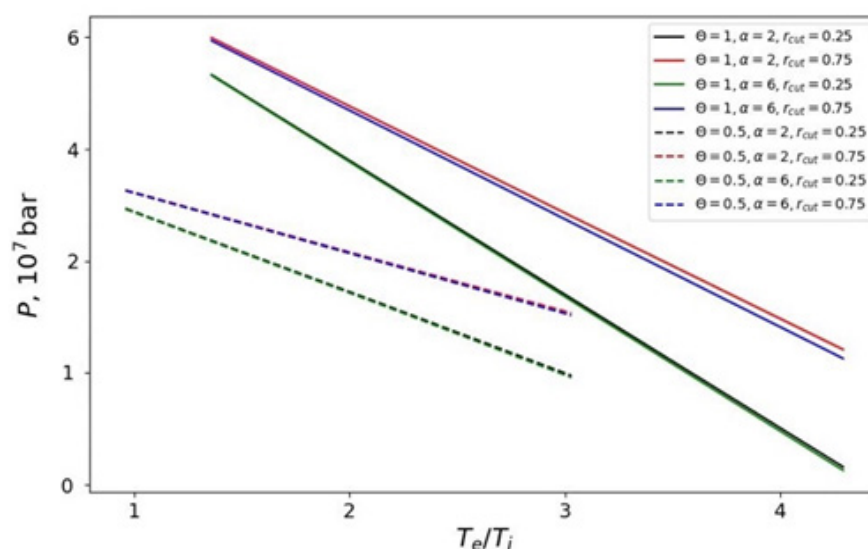


Figure 7 – The equation of state for the potential (1) depending on relation between temperatures of electrons and ions $\frac{T_e}{T_i}$ using values of $\Gamma_i = 1, 5, 10$, at $r_s = 1$ for $\theta = 1$ ($T_e = 580\,000\text{ K}$, solid lines) and $\theta = 0.5$ ($T_e = 290\,000\text{ K}$, dashed lines): $\alpha=2$, $r_{cut} = 0.25$ – black lines, $\alpha=2$, $r_{cut} = 0.75$ – red lines, $\alpha=6$, $r_{cut} = 0.25$ – green lines, $\alpha=6$, $r_{cut} = 0.75$ – blue lines

Our analysis demonstrates that as the distance between ions and electrons increases, the effective potential tends to converge towards a Yukawa-like screening potential. However, at shorter distances, the influence of strongly bound electrons creates a shielding effect that decreases the screening. The radial distribution functions (RDFs) further show the significance of adjusting parameters such as cutoff radius r_{cut} and core edge steepness α . For instance, a larger r_{cut} leads to a softer potential, while a smaller α results in a steeper potential at short distances.

Our analysis demonstrates that as the distance between ions and electrons increases, the effective potential tends to converge towards a Yukawa-like screening potential. However, at shorter distances, the influence of strongly bound electrons creates a shielding effect that decreases the screening. The radial distribution functions (RDFs) further show the significance of adjusting parameters such as cutoff radius r_{cut} and core edge steepness α . For instance, a larger r_{cut} leads to a softer potential, while a smaller α results in a steeper potential at short distances.

Additionally, we found that increasing Θ (and T), the absolute values of non-ideality corrections decrease, indicating fewer interactions in the system. The smaller absolute corrections to non-ideality are also seen with a larger cutoff radius of $r_{\text{cut}} = 0.75$ at a fixed α , as well as for a smaller $\alpha = 2$ at a fixed cutoff, due to reduced screening effects resulting from weaker attraction between bound electrons and ions weakens. On the other hand, as Γ_i increases (indicating stronger

Acknowledgments

This research has been funded by the Ministry of Science and Higher Education of the Republic of Kazakhstan (Grant No. AP19677200).

REFERENCES

- 1 Lifshitz, E.M., and Pitaevskii, L.P. *Fizicheskaya Kinetika [Physical Kinetics]* (Moscow: Physmathlit, 2002), pp. 96–97. (in Russian).
- 2 Ecker, G. *Theory of Fully Ionized Plasmas* (New York: Academic Press, 1972), pp. 132–134.
- 3 Ramazanov, T.S., Dzhumagulova, K.N., and Gabdullin, M.T. Effective Potentials for Ion-Ion and Charge-Atom Interactions of Dense Semiclassical Plasma. *Physics of Plasmas*, 17(4), 042703 (2010). <https://doi.org/10.1063/1.3381078>.
- 4 Ismagambetova, T.N., Moldabekov, Z.A., Amirov, S.M., et al. Dense Plasmas With Partially Degenerate Semiclassical Ions: Screening and Structural Properties. *Japanese Journal of Applied Physics*, 59, SHHA10 (2020). <https://doi.org/10.35848/1347-4065/ab75b5>.
- 5 Ramazanov, T.S., Dzhumagulova, K.N., and Moldabekov, Z.A. Generalized Pair Potential Between Charged Particles in Dense Semiclassical Plasma. *Physical Sciences and Technology*, 1(1), 48–54 (2018). <https://doi.org/10.26577/phst-2014-1-114>.
- 6 Ramazanov, T.S., Dzhumagulova, K.N., Gabdullin, M.T., Moldabekov, Z.A., and Ismagambetova, T.N. Development of Effective Potentials for Complex Plasmas. *Physical Sciences and Technology*, 6(3–4), 44–53 (2019). <https://doi.org/10.26577/phst-2019-2-p6>.
- 7 Ramazanov, T., and Moldabekov, Z. Dynamical collision frequency and conductivity of dense plasmas. *Physical Sciences and Technology*, 2(2), 53–57 (2016). <https://doi.org/10.26577/2409-6121-2015-2-2-53-57>.
- 8 Ramazanov, T.S., Kodanova, S.K., Issanova, M.K., Orazbayev, S.A., and Yelubaev, D.Ye. Temperature anisotropy relaxation processes in dense plasma. *Recent Contributions to Physics*, 75(4), 30–36 (2020). <https://doi.org/10.26577/RCPH.2020.v75.i4.04>.
- 9 Pines, D., and Nozieres, P. *The Theory of Quantum Liquids* (New York: Benjamin, 1966), pp. 277–278.
- 10 Ashcroft, N.W., and Stroud, D. *Theory of the Thermodynamics of Simple Liquid Metals*. *Solid State Physics*, 33, 1–81 (1977). [https://doi.org/10.1016/S0081-1947\(08\)60468-3](https://doi.org/10.1016/S0081-1947(08)60468-3).
- 11 Ashcroft, W., and Mermin, N.D. *Solid State Physics* (Philadelphia: Saunders College Publishing, 1976), p. 764.
- 12 Ramazanov, T.S., Kodanova, S.K., Nurusheva, M.M., Issanova, M.K. Ion core effect on scattering processes in dense plasmas. *Phys. Plasmas*, 28, 092702 (2021). <https://doi.org/10.1063/5.0059297>.
- 13 Ramazanov, T.S., Issanova, M.K., Aldakul, Y.K., Kodanova, S.K. Ion core effect on transport characteristics in warm dense matter. *Phys. Plasmas*, 29, 112706 (2022). <https://doi.org/10.1063/5.0102528>.
- 14 Ramazanov, T.S., Kodanova, S.K., Issanova, M.K., Kenzhegulov, B.Z. Influence of the ion core on relaxation processes in dense plasmas. *Contrib. Plasma Phys.* 64, e202300127 (2024). <https://doi.org/10.1002/ctpp.202300127>.
- 15 Ismagambetova, T.N., Muratov, M.M., Gabdullin, M.T., and Ramazanov, T.S. Influence of Ion Core on Structural and Thermodynamic Properties of Dense Plasma. *Contributions to Plasma Physics*, e70034 (2025). <https://doi.org/10.1002/ctpp.70034>.
- 16 Ismagambetova, T., Muratov, M., et al. The Influence of the Ionic Core on Structural and Thermodynamic Properties of Dense Plasmas. *Plasma*, 7(4), 858–866 (2024). <https://doi.org/10.3390/plasma7040046>.

- 17 Hansen, J.-P., McDonald, I.R. Theory of Simple Liquids (London, UK: Academic Press, 2000), pp. 100–102.
- 18 Gericke, D.O., Vorberger, J., et al. Screening of ionic cores in partially ionized plasmas within linear response. Phys. Rev. E, 81, 065401 (2010). <https://doi.org/10.1103/PhysRevE.81.065401>.
- 19 Moldabekov, Z.A., Dornheim, T. et al. Screening of a test charge in a free-electron gas at warm dense matter and dense non-ideal plasma conditions. Contrib. Plasma Phys., 62(2), e202000176 (2022). <https://doi.org/10.1002/ctpp.202000176>.
- 20 Moldabekov, Z.A., Groth, S., et al. Structural characteristics of strongly coupled ions in a dense quantum plasma. Phys Rev E, 98(2–1), 023207 (2018). <https://doi.org/10.1103/PhysRevE.98.023207>.
- 21 Groth, S., Dornheim, T., et al. Ab initio Exchange-Correlation Free Energy of the Uniform Electron Gas at Warm Dense Matter Conditions. Phys. Rev. Lett., 119, 135001 (2017). <https://doi.org/10.1103/PhysRevLett.119.135001>.
- 22 Bredow, R., Bornath, T., Kraeft, W.-D., and Redmer, R. Hypernetted Chain Calculations for Multi-Component and Non-Equilibrium Plasmas. Contributions to Plasma Physics, 53, 276–284 (2013). <https://doi.org/10.1002/ctpp.201200117>.
- 23 Kittel, C. Introduction to Solid State Physics, 8th Ed. (New York, USA: John Wiley and Sons, 2005), p. 696.
- 24 Springer, J.F., Pokrant, M.A., et al. Integral equation solutions for the classical electron gas. J. Chem. Phys., 58(11), 4863–4867 (1973). <https://doi.org/10.1063/1.1679070>.
- 25 Ng, K.-C. Hypernetted chain solutions for the classical one-component plasma up to $\Gamma=7000$. J. Chem. Phys., 61(7), 2680–2689 (1974). <https://doi.org/10.1063/1.1682399>.
- 26 Ramazanov, T.S., Ismagambetova, T.N., et al. The Influence of the Effects of the Bound Electrons on the Ion Structural and Thermodynamic Properties. IEEE Trans. Plasma Sci., 51(5), 1208–1211 (2023). <https://doi.org/10.1109/TPS.2023.3267845>.
- 27 Ishihara, A. Statistical physics (New York, United States: Academic Press, 1971), p. 287.

^{1,2,3*}**Исмагамбетова Т.Н.,**

PhD, сениор-лектор, ORCID ID: 0000-0003-4889-7526,

*e-mail: ismagambetova@physics.kz

^{1,2,4}**Муратов М.М.,**

PhD, қауымдастырылған профессор, ORCID ID: 0000-0001-7270-9834,

e-mail: mukhit.muratov@gmail.com

⁵**Усенов Е.А.,**

PhD, қауымдастырылған физик-зерттеуші, ORCID ID: 0000-0001-7903-8674,

e-mail: yussenov@pppl.gov

^{2,3,4}**Габдуллин М.Т.,**

PhD, ф.-м.ғ.к., профессор, ORCID ID: 0000-0003-4853-3642,

e-mail: gabdullin@physics.kz

¹Ашық Түрдегі Ұлттық Нанотехнологиялық Зертхана (АТҰНЗ),

Әл-Фараби атындағы Қазақ ұлттық университеті, Алматы қ., Қазақстан

²Қолданбалы ғылымдар және ақпараттық технологиялар институты, Алматы қ., Қазақстан

³Қазақстан-Британ техникалық университеті, Алматы қ., Қазақстан

⁴Қазақстандық физикалық қоғам, Алматы қ., Қазақстан

⁵Принстон плазма физикасы зертханасы, Принстон, Нью-Джерси, АҚШ

ИОНДЫҚ НЕГІЗДІҢ ИЗОТЕРМИЯЛЫҚ ЕМЕС ПЛАЗМАНЫҢ ҚАСИЕТТЕРІНЕ ӘСЕРІ

Аңдатпа

Бұл жұмыста иондық негізден (қаңқа) де, алмасу-корреляциялық өзара әрекеттесуден де пайда болатын экрандалуды ескеретін иондық-иондық өзара әрекеттесудің жаңа потенциалын қолдану арқылы иондық негіздің изотермиялық емес плазмаға әсері зерттеледі. Нәтижелер қашықтықтың өсуімен тиімді потенциал Юкава түрінің экрандалу потенциалына жақындағанын көрсетеді, ал сонымен қатар одан да қысқа қа-

шықтықта электрондардың күшті байланысы экрандалуды әлсіретеді. r_{cut} кесу радиусының және α қаңқа жиегінің тік болуының әртүрлі мәндерінде потенциалдың және радиалды таралу функциялардың (РТФ) күй-өзгерісіне айтарлықтай әсер етеді. Байланыс параметрлерінің (Γ_i) жоғары мәндері электрон-иондық өзара әрекеттесуді күшейтеді, бұл өз алдына тереңірек потенциалды шұңқырларға және идеалдылыққа неғұрлым айқын түзетулерге әкеледі. Θ мәнінің өсуі идеалдылыққа түзетулердің абсолютті мәндерін азайтады, бұл жүйеде өзара әрекеттесудің мөлшерінің аздығын көрсетеді. Белгіленген бір α параметріндегі үлкен r_{cut} кесу радиусы әлсіз экрандалу әсеріне байланысты түзетулерді азайтады. Γ_i ұлғаюымен идеалдылыққа түзетулер байланыстың күшеюін көрсете отырып артады. Нәтижелер тығыз, изотермиялық емес плазманы зерттеу кезінде иондық негіздің әсерін есепке алудың маңыздылығын көрсетеді.

Тірек сөздер: тығыз плазма, иондық негіз, изотермиялық емес плазма, иондардың өзара әрекеттесу потенциалы, радиалды таралу функциялары,

^{1,2,3*}**Исмагамбетова Т.Н.,**

PhD, сениор-лектор, ORCID ID: 0000-0003-4889-7526,

*e-mail: ismagambetova@physics.kz

^{1,2,4}**Муратов М.М.,**

PhD, ассоциированный профессор, ORCID ID: 0000-0001-7270-9834,

e-mail: mukhit.muratov@gmail.com

⁵**Усенов Е.А.,**

PhD, ассоциированный физик-исследователь, ORCID ID: 0000-0001-7903-8674,

e-mail: yussenov@pppl.gov

^{2,3,4}**Габдуллин М.Т.,**

PhD, к.ф.-м.н., профессор, ORCID ID: 0000-0003-4853-3642,

e-mail: gabdullin@physics.kz

¹Национальная нанотехнологическая лаборатория открытого типа (ННЛОТ),
Казахский национальный университет им. аль-Фараби, г. Алматы, Казахстан

²Институт прикладных наук и информационных технологий, г. Алматы, Казахстан

³Казахстанско-Британский технический университет, г. Алматы, Казахстан

⁴Казахское физическое общество, г. Алматы, Казахстан

⁵Принстонская лаборатория физики плазмы, г. Принстон, Нью-Джерси, США

ВЛИЯНИЕ ИОННОГО ОСТОВА НА СВОЙСТВА НЕИЗОТЕРМИЧЕСКОЙ ПЛАЗМЫ

Аннотация

В данной работе изучается влияние ионного остова на неизоотермическую плазму с использованием нового потенциала ион-ионного взаимодействия, учитывающего экранирующие эффекты как от ионного остова, так и от обменно-корреляционных взаимодействий. Результаты показывают, что с увеличением расстояния эффективный потенциал приближается к экранирующему потенциалу типа Юкавы, в то время как на более коротких расстояниях сильная связь электронов ослабляет экранирование. Различные значения радиуса обрезания r_{cut} и крутизны края остова α существенно влияют на поведение потенциала и радиальных функций распределения (РФР). Более высокие значения параметров связи (Γ_i) усиливают электрон-ионные взаимодействия, что приводит к более глубоким потенциальным ямам и более выраженным поправкам на неидеальность. Увеличение Θ уменьшает абсолютные значения поправок на неидеальность, что указывает на меньшее количество взаимодействий в системе. Большой радиус обрезания r_{cut} при фиксированном параметре α также уменьшает поправки из-за более слабого экранирующего эффекта. С увеличением Γ_i поправки на неидеальность возрастают, отражая усиление связи. Результаты показывают важность учета эффектов ионного остова при исследовании плотной неизоотермической плазмы.

Ключевые слова: плотная плазма, ионный остов, неизоотермическая плазма, потенциал взаимодействия ионов, радиальные функции распределения, термодинамические свойства.

Article submission date: 23.09.2025

УДК 621.315.592
МРНТИ 29.15.05; 29.17.27; 29.19.03

<https://doi.org/10.55452/1998-6688-2025-22-4-365-373>

^{1*}**Шонғалова А.Қ.,**

PhD, старший научный сотрудник ФТИ, ORCID ID: 0000-0002-7352-9007,

*e-mail: shongalova.aigul@gmail.com

²**Токмолдин Н.,**

PhD, ORCID ID: 0000-0002-0663-0228, e-mail: tokmoldin@pdi-berlin.de,

³**Өміртай М.,**

магистрант, ORCID ID: 0009-0003-2275-9018, e-mail: m.omirtay@gmail.com

¹Физико-технический институт, Satbayev University, г. Алматы, Казахстан

²Институт физики и астрономии, Потсдамский университет, г. Потсдам, Германия

³Satbayev University, г. Алматы, Казахстан

ПОЛУЧЕНИЕ И ИССЛЕДОВАНИЕ ПЛЕНКИ СЕЛЕНИДА СУРЬМЫ ПУТЕМ СЕЛЕНИЗАЦИИ

Аннотация

В данной работе представлены метод синтеза пленки селенида сурьмы (Sb_2Se_3), а также исследования его морфологии, структурных и оптических свойств. Метод синтеза состоял из двух этапов. На первом этапе с помощью магнетронного распыления был получен прекурсор из сурьмы. На втором этапе была проведена селенизация в парах селена при температуре 400 °C в течение 10 минут. Морфология полученных пленок была исследована с помощью сканирующей электронной микроскопии. Результаты исследования морфологии показали, что пленка имеет поликристаллическую структуру с хорошей адгезией к кремниевой подложке. С помощью энергодиспергирующей спектроскопии (ЭДС) был исследован элементный состав пленки. По результатам ЭДС соотношение атомного процента Se/Sb составило 1,59, что указывает на получение пленки, близкой к стехиометрическому. Данные, полученные ЭДС, были подтверждены исследованием фазового состава, выполненным методом рентгенофазного исследования. Было выяснено, что пленка кристаллизуется в орторомбической сингонии (Pnma). Дополнительных фаз в структуре не было обнаружено. Для исследования оптоэлектронных свойств пленки был снят спектр отражения. Из спектра отражения методом Таука была определена ширина запрещенной зоны, равная 1,69 эВ, что является оптимальным для применения в оптоэлектронных устройствах.

Ключевые слова: селенид сурьмы, селенизация, магнетронное распыление, оптоэлектроника, пленки.

Введение

Одним из важных направлений разработки новых материалов является поиск и создание экономически эффективных возобновляемых источников энергии. В число таких перспективных технологий входят ветровая, био-, гидро-, солнечная энергетика. В глобальном масштабе на данный момент мощность всех солнечных энергостанций в мировом электроэнергетическом балансе составляет лишь 1% [1]. Основными факторами небольшого спроса на данный момент являются цена и относительно низкая эффективность фотоэлементов, присутствующих на рынке солнечной энергетике. По этой причине исследователи изучают новые недорогие материалы, находящиеся в избытке в земной коре, оптимизируют существующие структуры и разрабатывают инновационные концепты структуры солнечных элементов, превосходящие обычный одиночный p-n-переход с точки зрения стоимости и эффективности.

В настоящее время солнечные батареи на основе поли- и монокристаллического кремния составляют более 80% рынка солнечной энергетике, по данным 2024 г. [2]. Остальные составляют фотоэлементы, изготовленные по тонкопленочной технологии из аморфного кремния

(a-Si), селенида меди, индия, галлия (Cu(In,Ga)Se_2 (CIGS)) [3], теллурида кадмия (CdTe) [4]. Эффективность солнечных элементов на основе CIGS и CdTe составляет свыше 20%, у a-Si – выше 13% [5]. Однако технологии CIGS и CdTe имеют существенные недостатки: в структурах CIGS вследствие ограниченности In, а в CdTe вследствие содержания тяжелых металлов, что вызывает озабоченность по причине воздействия на окружающую среду [6–7]. Как альтернатива появились новые материалы, состоящие из относительно распространенных элементов. В последнее десятилетие такие соединения, как $\text{Cu}_2\text{ZnSn(Se,S)}$ [8], CuSbS [9] и $\text{Sb}_2(\text{Se,S})_3$ [10], исследуются в качестве перспективных материалов для СЭ. Кроме того, набирает популярность новое поколение фотоэлектрических материалов, в том числе из гибридов органических и неорганических материалов – органо-неорганических перовскитов (например, $\text{CH}_3\text{NH}_3\text{PbI}_3$) с эффективностью выше 20% [11].

Селенид сурьмы (Sb_2Se_3) – полупроводниковый халькогенид группы $A_m^{\text{VI}}B_n^{\text{VI}}$, кристаллизующийся в орторомбической сингонии, пространственная группа $Pnma$ (62), обладающий оптимальной шириной запрещенной зоны в диапазоне 1,04–1,8 эВ для Sb_2Se_3 , высоким коэффициентом поглощения ($\alpha > 10^4 \text{ cm}^{-1}$) в видимой области спектра [12]. На основании электрохимических, опто- и термоэлектрических свойств были опубликованы более 100 трудов-исследований о возможностях применении этих материалов в качестве анодных материалов для литий-ионных батарей и для хранения водорода, в устройствах оптической записи, в опто- и термоэлектрических охлаждающих устройствах, в солнечных и фотоэлектрохимических элементах.

Технология синтеза имеет прямое влияние на структуру получаемых пленок, тем самым влияя на эффективность элемента. Известно три наиболее распространенных метода получения соединения $\text{Sb}_2(\text{S,Se})_3$: методы химического, термического и электрического осаждения [13].

Магнетронное распыление относится к методам распыления материалов ионной бомбардировкой. Нанесение покрытий в вакууме с помощью магнетронных систем заключается в распылении твердой мишени напыляемого материала ионами инертного газа, образующимися в плазме аномального тлеющего разряда при наложении на него магнитного поля, силовые линии которого ортогонально пересекают силовые линии магнитного поля. К основным достоинствам магнетронного распыления относятся: легкая контролируемость химического состава в тонких пленках и хорошая воспроизводимость материала катода, невысокая температура подложки, возможность нанесения однородных пленок, высокая адгезия, однородность толщины пленки, возможность управления структурой. Метод магнетронного распыления широко используется для получения CIGS структур. Как показывают результаты измерений характеристик пленок, полученных по данной методике, они обладают высоким качеством кристаллической структуры и высокой эффективностью по сравнению с пленками, полученными другими методами [14]. Недостатками метода магнетронного распыления является высокая энергоемкость процесса.

В данной работе представлен синтез поликристаллической пленки с использованием магнетронного распыления, и исследования ее характеристик.

Материалы и методы

Синтез. Получение пленки селенида сурьмы состояло из двух этапов. На первом этапе методом магнетронного распыления были получены пленки сурьмы. Далее пленка сурьмы проходила этап селенизации в трубчатой печи. В качестве источника сурьмы была использована мишень собственного производства. Мишень сурьмы была получена из кусков сурьмы путем их дробления до состояния крупнозернистого песка и спеканием в трехзонной печи быстрого отжига в атмосфере аргона при температуре ~650 °С. Перед спеканием полученная масса была подвержена очистке от грязи и жира: обрабатываемая дистиллированной водой, уксусом, спиртом в ультразвуковой ванне в течение 15 мин. последовательно, далее сушилась в

сушильном шкафу при температуре 100 оС до высыхания. Форма для мишени была изготовлена из кварцевого стекла. Диаметр полученной формы составляет ~55 мм. Перед спеканием полученная форма была очищена по вышеприведенной схеме. После подготовки материалы были помещены в печь. Мишень спекалась в течение 20 мин., поток аргона во время процесса был одинаков – 300 ст. см³/мин. После спекания мишень была доведена до необходимого диаметра катода магнетрона – 50 мм путем распиливания лишних краев. На рисунке 1а) приведен элементный анализ, полученный методом энергодисперсионной рентгеновской спектроскопией спеченной мишени. Мишень состоит исключительно из сурьмы, сторонних элементов в спектре не наблюдается.

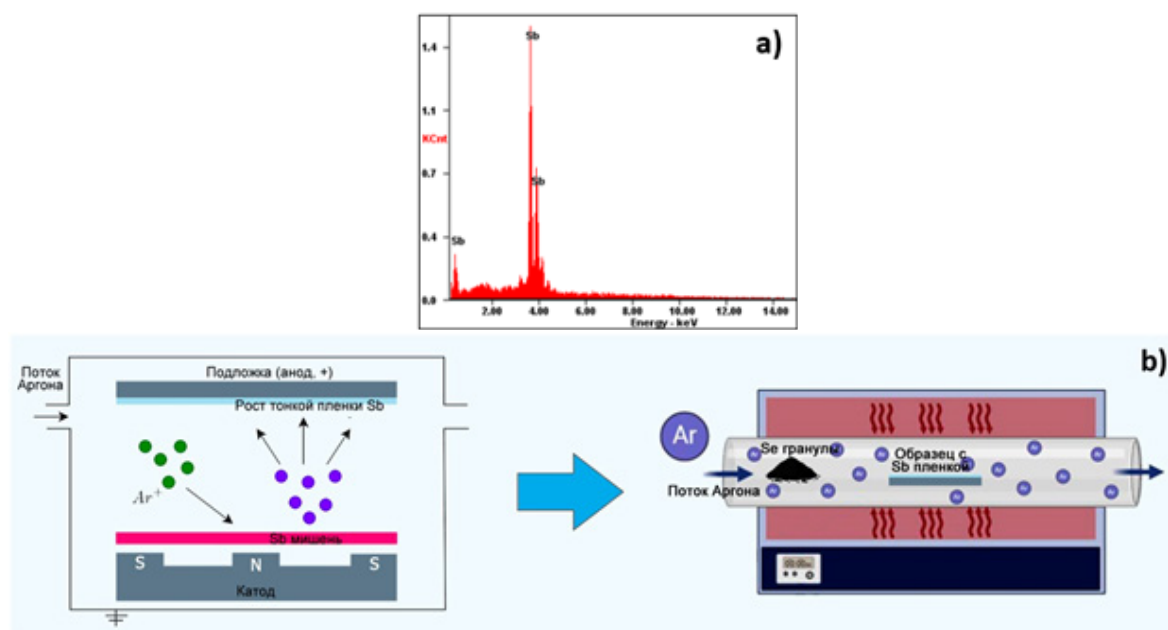


Рисунок 1 – Схема получения пленки селенида сурьмы: напыление пленки сурьмы магнетронным распылением и дальнейшая селенизация

Пленка сурьмы осаждалась на подложки из кремния n-типа методом радиочастотного магнетронного распыления в атмосфере аргона при рабочем давлении $1,8 \times 10^{-3}$ Торр. Пленка осаждалась в течение 20 мин.

Селенизация пленок сурьмы, осажденных на кремний, проводилась в кварцевой печи. Печь состоит из системы нагрева, состоящей из проволоки никель-хромового сплава, намотанной вокруг трубки, которая подсоединена к лабораторному автотрансформатору. Температура нагрева трубы регулировалась путем подачи напряжения к концам проволоки с помощью лабораторного автотрансформатора.

Процесс нагрева печи состоит из нескольких этапов. Предварительно объем печи наполняли аргоном, чтобы очистить атмосферу от кислорода. Затем печь с двух концов закрывали для предотвращения попадания воздуха во время отжига. В качестве источника селена были использованы гранулы селена. Гранулы были помещены в открытую емкость из никель-хромовой фольги. При помощи лабораторного автотрансформатора подавалось напряжение, и температура печи постепенно повышалась. Для достижения необходимой температуры селенизации печь нагревалась в течение 30–40 мин. Скорость нагрева составляла ~ 15 °С/мин. Далее образцы удерживались при фиксированной температуре 400 °С в течение 10 мин. Температура селенизации, представленная в данной работе, является одной из многих, испытанных в ранней нашей работе [15]. После селенизации образцы охлаждались в печи до комнатной температуры в атмосфере селена.

Исследование характеристик. Были исследованы морфология и элементный состав пленки селенида сурьмы на сканирующем электронном микроскопе (СЭМ) и энергодиспергирующем спектро스코пе (ЭДС) (JEOL). Структурный и фазовый состав изучались с помощью рентгеноструктурного анализа и спектроскопии комбинационного рассеяния света (КРС). Измерения спектроскопии КРС проводили с использованием Solver Spectrum. Поверхность образца была сфокусирована с помощью объектива 100× с диаметром лазерного пятна ~ 1 мкм. В качестве источника возбуждения использовался HeNe лазер (632,8 нм). Рентгеноструктурный анализ пленок проводился на рентгеновском дифрактометре RIGAKU MiniFlex 600 с программным обеспечением PPXL2.0.

Результаты и обсуждение

На рисунке 2 приведены изображения пленки сурьмы: (а) на кремнии и после селенизации (b). До селенизации пленка сурьмы имеет зеркальную поверхность, после селенизации, т.е. пленка Sb_2Se_3 имеет матовую поверхность. У полученной пленки хорошая адгезия.

При исследовании элементной структуры пленок по анализу были идентифицированы три элемента: Sb, Se (элементы пленки) и Si (материал подложки). Соотношение селена к сурьме по результатам элементного анализа с помощью ЭДС составило 1,59, что очень близко к стехиометрическому значению – 1,5. По данным ЭДС, содержание элементов в пленке составляет 61 ат.% селена и 39 ат.% сурьмы при исключении кремния.

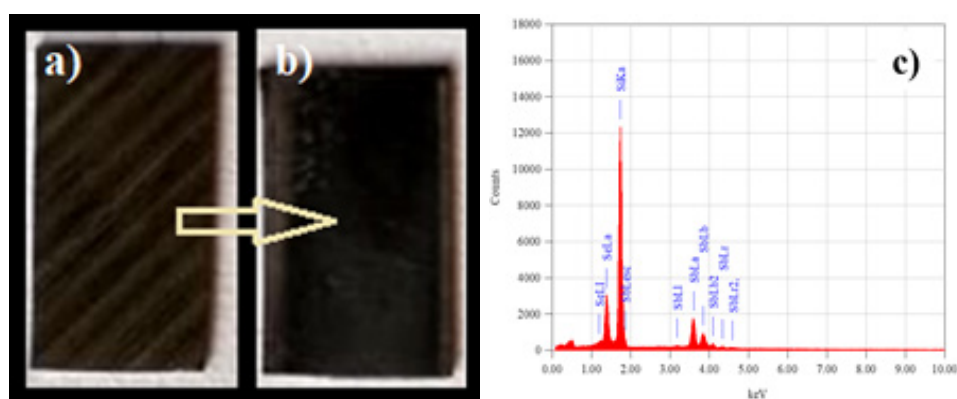


Рисунок 2 – а) пленка сурьмы на кремнии; b – пленка селенида сурьмы;
с) спектры энергодисперсионной спектроскопии для пленки селенида сурьмы

Изображения поверхности пленки селенида сурьмы с различным увеличением представлены на рисунке 3. Поверхность пленки при низком увеличении показывает относительно однородное покрытие без выраженных дефектов (рисунок 3а). А при более высоком увеличении пленка демонстрирует зернистую структуру с разными формами, что указывает на поликристаллический характер (рисунок 3b). Поперечное сечение пленки показывает столбчатый рост кристаллитов перпендикулярно поверхности подложки, с хорошей адгезией слоя к подложке (рисунок 3с).

Результаты спектроскопии КРС полученный при нормальных условиях представлены на рисунке 4. В спектре наблюдаются характерные для Sb_2Se_3 пики, расположенные при ~100, ~120, ~153, ~190, ~210 cm^{-1} . На спектре не обнаружено наличие другой фазы.

Результаты рентгеноструктурного анализа (рисунок 4) показали, что пленки Sb_2Se_3 кристаллизуются в орторомбической сингонии, пространственная группа $Pnma$ (62). Фазовый анализ проводился на основании банка данных Международного центра дифракционных данных (ICDD, карта PDF № 01-083-7430). Согласно данным рентгеноструктурного анализа, кро-

ме фазы селенидов сурьмы, других фаз в образце не определено. Показания рентгеноструктурного анализа соответствуют результатам спектроскопии КРС.

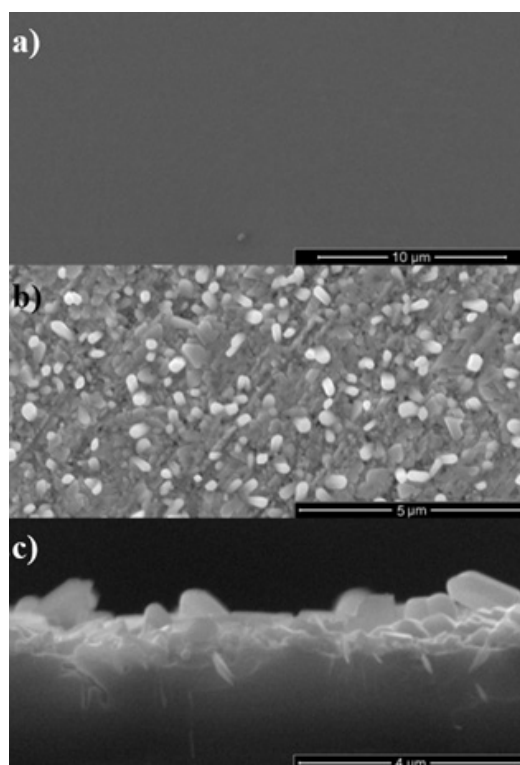


Рисунок 3 – Морфология пленки селенида сурьмы, полученного методом сканирующей электронной микроскопии

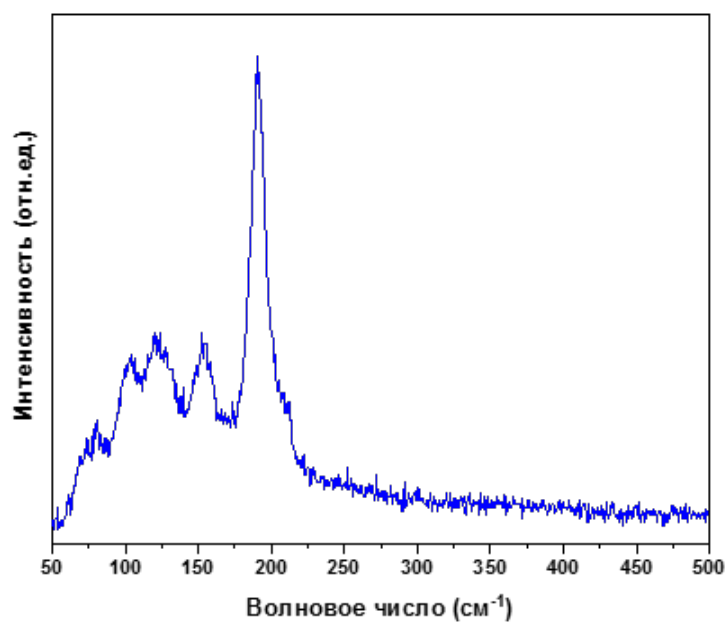


Рисунок 4 – Спектр комбинационного рассеяния света пленки селенида сурьмы

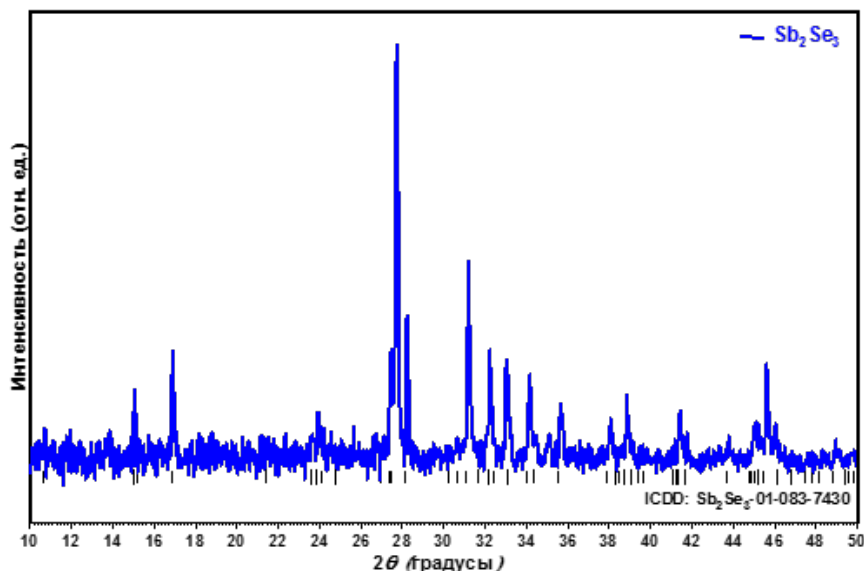


Рисунок 5 – Спектры рентгеноструктурного анализа для пленок селенида сурьмы

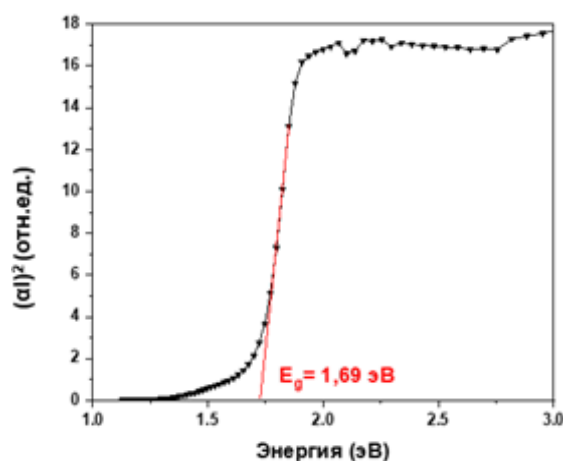


Рисунок 6 – Определение ширины запрещенной зоны пленки Sb₂Se₃ методом Таука

Для исследования оптоэлектронных характеристик был снят спектр отражения пленки. Ширина запрещенной зоны была определена методом Таука исходя из зависимости $A(h\nu - E_g) = (\alpha h\nu)^2$, где показатель $n = 2$, соответствует прямозонному переходу от энергии фотонов (рисунок 6). Путем экстраполяции зоны кривой была определена ширина запрещенной зоны, равная 1,69 эВ. В нижней части спектра, ближе к энергетической оси, наблюдается дополнительный «хвост», который связан с подложкой из кремния. Этот край соответствует ширине запрещенной зоны кремния – примерно 1,1 эВ, что подтверждает корректность интерпретации и выделения вкладов пленки и подложки [16].

По литературным данным, в зависимости от соотношения селена и сурьмы, а также кристаллической структуры ширина запрещенной зоны обычно варьируется от ~1,1 эВ до 1,8 эВ [17].

Таким образом, значение 1,69 эВ представляет собой максимальный диапазон по сравнению с данными для объемных образцов. Это может быть связано со структурой тонкой пленки, дефектами или локализованными состояниями, расширяющими край поглощения, что является распространенным эффектом для тонких пленок Sb_2Se_3 [18].

Заключение

Полученные результаты продемонстрировали, что комбинация магнетронного распыления прекурсора из сурьмы и дальнейший отжиг/селенизация позволяют эффективно синтезировать селенид сурьмы. Данный метод позволил получить структуру с оптимальными характеристиками и имеет перспективу применения в оптоэлектронных устройствах.

Исследования элементного состава пленки показали соотношение селена к сурьме 1,59. Этот показатель очень близок к теоретическому показателю стехиометрии – 1,5, показывая оптимальность примененных параметров синтеза. Морфология пленки продемонстрировала хорошую адгезию к подложке, а также зернистую структуру пленки. Рентгеноструктурный анализ и спектр КРС пленок подтвердил данные, полученные методом ЭДС. Дополнительных, нежелательных фаз в кристаллической фазе пленки не было обнаружено. Было определено, что пленка кристаллизуется в орторомбической сингонии и соответствует данным международной кристаллографии.

Ширина запрещенной зоны, определенная методом Таука, составила 1,69 эВ. Второй «хвост» 1,1 эВ, наблюдаемый в спектре, был определен как ширина запрещенной зоны подложки.

Проведенные исследования показали, что синтезированная пленка селенида сурьмы имеет потенциал для применения в качестве одного из недорогих, сравнительно экологически чистых и энергоэффективных полупроводников в будущих солнечных элементах, оптоэлектронных устройствах и фотоэлектрохимических системах.

Информация о финансировании

Исследование выполнено при финансовой поддержке Комитета науки Министерства образования и науки Республики Казахстан, грант AP19178659 «Исследование динамики решетки селенида сурьмы методом колебательной спектроскопии».

Авторы заявляют об отсутствии конфликта интересов.

ЛИТЕРАТУРА

- 1 Le Marois, J.B., Pales, A.F., Bennett, S. Reaching net zero emissions demands faster innovation, but we've already come a long way. International Energy Agency, 13 (2023).
- 2 Lee, T.D., Ebong, A.U. A review of thin film solar cell technologies and challenges. Renewable and Sustainable Energy Reviews, 70, 1286–1297 (2017).
- 3 Machkih, K., Oubaki, R., Makha, M. A review of CIGS thin film semiconductor deposition via sputtering and thermal evaporation for solar cell applications. Coatings, 14(9), 1088 (2024).
- 4 Scarpulla. M.A. et al. CdTe-based thin film photovoltaics: Recent advances, current challenges and future prospects. Solar Energy Materials and Solar Cells, 255, 112289 (2023).
- 5 Stuckelberger, M., Biron, R., Wyrsh, N., Haug, F.J., Ballif, C. Progress in solar cells from hydrogenated amorphous silicon. Renewable and Sustainable Energy Reviews, 76, 1497–1523 (2017).
- 6 Han, T., Luo, M., Liu, Y., Lu, C., Ge, Y., Xue, X., Xu, X. $\text{Sb}_2\text{S}_3/\text{Sb}_2\text{Se}_3$ heterojunction for high-performance photodetection and hydrogen production. Journal of Colloid and Interface Science, 628, 886–895 (2022).
- 7 Mavlonov, A., Razykov, T., Raziq, F., Gan, J., Chantana, J., Kawano, Y., Qiao, L. A review of Sb_2Se_3 photovoltaic absorber materials and thin-film solar cells. Solar Energy, 201, 227–246 (2020).

- 8 Qin, X., Xu, B., Lin, J., Chen, J., Tong, H., Chen, Y., Sun, L. Above 10% efficient electrodeposited $\text{Cu}_2\text{ZnSn}(\text{S}, \text{Se})_4$ solar cell achieved by modifying precursor. *Solar Energy Materials and Solar Cells*, 242, 111781 (2022).
- 9 Singh, P.K., Rai, S., Lohia, P., Dwivedi, D.K. Comparative study of the CZTS, CuSbS_2 and CuSbSe_2 solar photovoltaic cell with an earth-abundant non-toxic buffer layer. *Solar Energy*, 222, 175–185 (2021).
- 10 Zhang, L., Zheng, J., Liu, C., Xie, Y., Lu, H., Luo, Q., Mai, Y. Over 10% Efficient $\text{Sb}_2(\text{S}, \text{Se})_3$ Solar Cells Enabled by CsI-Doping Strategy. *Small*, 20(27), 2310418 (2024).
- 11 Nowshad, N., Islam, R., Uddin, M.N. Photovoltaic Performance Analysis of $\text{CH}_3\text{NH}_3\text{PbI}_3$ Perovskite Solar Cells Using SCAPS Simulation. *Chemistry of Inorganic Materials*, 100107 (2025).
- 12 Maurya, K.K., Singh, V. NSb_2Se_3 versus Sb_2S_3 solar cell: a numerical simulation. *Solar Energy*, 228, 540–549 (2021).
- 13 Zeng, K., Xue, D. J., Tang, J. Antimony selenide thin-film solar cells. *Semiconductor Science and Technology*, 31(6), 063001 (2016).
- 14 Greene, J.E. Tracing the recorded history of thin-film sputter deposition: From the 1800s to 2017. *Journal of Vacuum Science & Technology A.*, 35(5) (2017).
- 15 Shongalova, A., Correia, M.R., Teixeira, J.P., Leitão, J.P., González, J.C., Ranjbar, S., Fernandes, P.A. Growth of Sb_2Se_3 thin films by selenization of RF sputtered binary precursors. *Solar Energy Materials and Solar Cells*, 187, 219–226 (2018).
- 16 Oliphant, E., Mantena, V., Brod, M., Snyder, G.J., Sun, W. Why does silicon have an indirect band gap? *Materials Horizons*, 12(9), 3073–3083 (2025).
- 17 Nicolás-Marín, M.M., González-Castillo, J.R., Vigil-Galán, O., Courel, M. The state of the art of $\text{Sb}_2(\text{S}, \text{Se})_3$ thin film solar cells: current progress and future prospect. *Journal of Physics D: Applied Physics*, 55(30), 303001(2022).
- 18 Cifuentes, N., Ghosh, S., Shongalova, A., Correia, M.R., Salomé, P.M., Fernandes, P.A., González, J.C. Electronic conduction mechanisms and defects in polycrystalline antimony selenide. *The Journal of Physical Chemistry C*, 124(14), 7677–7682 (2020).

^{1*}Шонгалова А.Қ.,

PhD, ФТИ аға ғылыми қызметкері, ORCID ID: 0000-0002-7352-9007,

*e-mail: shongalova.aigul@gmail.com

²Токмолдин Н.,

PhD, ORCID ID: 0000-0002-0663-0228,

e-mail: tokmoldin@pdi-berlin.de

³Өміртай М.,

магистрант, ORCID ID: 0009-0003-2275-9018,

e-mail: m.omirtay@gmail.com

¹Физика-техникалық институт, Satbayev University, Алматы қ., Қазақстан

²Физика және астрономия институты, Потсдам университеті, Потсдам қ., Германия

³Satbayev University, Алматы қ., Қазақстан

СЕЛЕНДЕУ ӘДІСІМЕН СҮРМЕ СЕЛЕНИДІН АЛУ ЖӘНЕ ЗЕРТТЕУ

Аңдатпа

Осы жұмыста сүрме селениді (Sb_2Se_3) жұқа қабықшаларын синтездеу әдісі, сондай-ақ олардың морфологиялық, құрылымдық және оптикалық қасиеттерін зерттеу нәтижелері ұсынылған. Синтездеу әдісі екі кезеңнен тұрды. Бірінші кезеңде магнетрондық тозаңдату әдісі арқылы сурьмадан прекурсор қабықша алынды. Екінші кезеңде 400°C температурада 10 минут бойы селен буларында селенизация процесі жүргізілді. Алынған қабықшалардың морфологиясы сканерлеуші электрондық микроскопия әдісімен зерттелді. Морфологиялық зерттеу нәтижелері қабықшаның кремнийлі төсемге жақсы адгезиясымен сипатталатын поликристалдық құрылымға ие екенін көрсетті. Қабықшаның элементтік құрамы энергодисперсиялық спектроскопия (ЭДС) әдісімен анықталды. ЭДС нәтижелері бойынша Se/Sb атомдық пайыздық қатынасы 1,59-ға тең болды, бұл алынған қабықшаның стехиометриялық құрамға жақын екенін көрсетеді. ЭДС арқылы

алынған деректер рентгендік фазалық талдау әдісімен жүргізілген фазалық құрамды зерттеу нәтижелерімен расталды. Қабықшаның орторомбиялық сингонияда (Pnma) кристалданатыны анықталды. Құрылымда қосымша фазалар табылған жоқ. Қабықшаның оптоэлектрондық қасиеттерін зерттеу үшін шағылу спектрі өлшенді. Шағылу спектрінен Таука әдісі арқылы тыйым салынған аймақтың ені 1,69 эВ екені анықталды, бұл мән оптоэлектрондық құрылғыларда қолдану үшін оңтайлы болып табылады.

Тірек сөздер: сүрме селениді, селендеу, магнетронды бүрку, оптоэлектроника, жұқа қабықшалар.

¹**Shongalova A.K. ***,

PhD, Senior Researcher at IPT, ORCID ID: 0000-0002-7352-9007,

*e-mail: shongalova.aigul@gmail.com

²**Tokmoldin N.,**

PhD, ORCID ID: 0000-0002-0663-0228,

e-mail: tokmoldin@pdi-berlin.de

³**Omirtay M.,**

Master's student, ORCID ID: 0009-0003-2275-9018,

e-mail: m.omirtay@gmail.com

¹Institute of Physics and Technology, Satbayev University, Almaty, Kazakhstan

²Institute of Physics and Astronomy, University of Potsdam, Potsdam, Germany

³Satbayev University, Almaty, Kazakhstan

PREPARATION AND INVESTIGATION OF ANTIMONY SELENIDE FILM BY SELENIZATION

Abstract

In this work, a method for synthesizing antimony selenide (Sb₂Se₃) thin films is presented, along with investigations of their morphology, structural, and optical properties. The synthesis method consisted of two stages. In the first stage, an antimony precursor film was deposited by magnetron sputtering. In the second stage, selenization was carried out in selenium vapor at a temperature of 400 °C for 10 minutes. The morphology of the obtained films was examined using scanning electron microscopy. The morphological analysis showed that the film has a polycrystalline structure with good adhesion to the silicon substrate. The elemental composition of the film was analyzed by energy-dispersive X-ray spectroscopy (EDS). According to the EDS results, the atomic percentage ratio of Se/Sb was 1.59, indicating that the obtained film is close to stoichiometric composition. The EDS data were confirmed by phase composition analysis performed using X-ray diffraction. It was found that the film crystallizes in the orthorhombic structure (Pnma). No secondary phases were detected in the structure. To investigate the optoelectronic properties of the film, a reflectance spectrum was recorded. From the reflectance spectrum, the band gap energy was determined using the Tauc method and found to be 1.69 eV, which is optimal for applications in optoelectronic devices.

Keywords: antimony selenide, selenization, magnetron sputtering, optoelectronics, films.

Дата поступления статьи в редакцию: 03.09.2025

ӘОЖ 669.2/.8:621.762
FTAXP 29.19.16

<https://doi.org/10.55452/1998-6688-2025-22-4-374-385>

¹**Ескермесов Д.К.,**

PhD, қауымдастырылған профессор, ORCID ID: 0000-0002-2206-8132,
e-mail: dyeskermessov@edu.ektu.kz

²**Лукос С.С.,**

PhD, ғылыми қызметкер, ORCID ID: 0000-0001-9668-0079,
e-mail: c.c.lukose@northumbria.ac.uk

²**Биркетт М.,**

PhD, профессор, ORCID ID: 0000-0002-6135-6639,
e-mail: martin.birkett@northumbria.ac.uk

³**Пазылбек С.А.,**

PhD, қауымдастырылған профессор, ORCID ID: 0000-0002-2318-9155,
e-mail: sapargali.pa@gmail.com

¹**Ескермесова А.С.,**

магистр, кіші ғылыми қызметкер, ORCID ID: 0009-0004-8454-3615,
e-mail: ardakeskermesova@gmail.com

^{1*}**Төлеуханова Ж.Т.**

магистр, аға оқытушы, ORCID ID: 0009-0002-9254-3901,
*e-mail: ztoleukhanova@ektu.kz

¹Д. Серікбаев атындағы Шығыс Қазақстан техникалық университеті,
Өскемен қ., Қазақстан

²Нортумбрия университеті, Ньюкасл қ., Ұлыбритания

³Ж.А. Тәшенев атындағы университет, Шымкент қ., Қазақстан

ФИЗИКА-МЕХАНИКАЛЫҚ СИПАТТАМАЛАРЫ ЖАҚСАРТЫЛҒАН БИОМЕДИЦИНАЛЫҚ ҚОЛДАНУҒА АРНАЛҒАН ТИТАН-АЛТЫН (Ti-Au) НЕГІЗІНДЕГІ ЖҰҚА ЖАБЫНДАРДЫ АЛУ ЖӘНЕ ЗЕРТТЕУ

Аңдатпа

Бұл ғылыми мақалада 450 °C жоғарғы температурада тұрақты (DC) және жоғары жиілікті (RF) ток режимінде вакуумдық ортада PVD (магнетрондық бүрку) әдісімен алынған Ti-Au негізіндегі жұқа жабын бетінің құрылымы мен қасиеттері зерттелді. NanoPVD қондырғысымен алынған жұқа жабындардың қалыңдықтары орташа есеппен 1 мкм құрады. Эксперименттік зерттеулер көрсеткендей, жабындардың химиялық құрамы, микроқұрылымы және қасиеттері магнетрондық бүрку әдісінің технологиялық параметрлеріне тығыз байланысты екені анықталды. Ti-Au жабындарының микроқұрылымы мен механикалық қасиеттеріне Ag және Cu қоспаларының әсері растрлы электронды микроскопия (SEM), энергия-дисперсиялық микроанализ (EDX) және рентгендік дифрактометрия (XRD) әдістері арқылы зерттелді. Жабын беттің кедір-бұдырлық сипаттамалары атомдық-күштік микроскопия (AFM) арқылы талданды, ал жабындардың қасиеттеріндегі өзгерістерді жан-жақты бағалау үшін нанокаттылық пен трибологиялық сынақтар жүргізілді. PVD әдісімен Ti_6Al_4V төсеніш бетінің модификациясы оның тозуға төзімділігін арттырып, үйкеліс коэффициентін төмендететіні көрсетілді. Рентгендік фазалық талдау нәтижелері бойынша, алынған жұқа жабындардың трибологиялық қасиеттерінің жақсаруы, тікелей негізгі Ti_3Au фазасының түзілуімен және оның мөлшерінің артуымен байланысты екені анықталды.

Тірек сөздер: жабын және жұқа қабықшалар, магнетронды бүрку жүйесі, морфология, қаттылық, элементтік құрам.

Кіріспе

Соңғы онжылдықтарда биомедициналық материалдар саласында, әсіресе медициналық имплантаттар мен хирургиялық құралдарға арналған биоүйлесімді және тозуға төзімді инновациялық жабындар әзірлеу технологиялары қарқынды дамып келеді [1–5]. Бұл бағытта жүргізіліп жатқан зерттеулер адамның ағзасымен ұзақ мерзімді өзара әрекеттесу кезінде жабын материалдарының сенімділігі мен тұрақтылығын қамтамасыз етуге бағытталған. Аталған саладағы ғылыми және инженерлік ізденістердің басым бағыты ретінде титан (Ti) және оның қорытпалары негізінде синтезделген жұқа қабықшалы жабындарды қолдануға ерекше назар аударылып отыр [6–10]. Бұл таңдау олардың жоғары коррозияға төзімділігімен, биоинерттілігімен, сондай-ақ жақсы физикалық-механикалық сипаттамаларымен (жоғары қаттылық, тығыздық, адгезия) тікелей байланысты [11, 12]. Сонымен қатар, мұндай жабындардың функционалдық қасиеттерін (мысалы, тозуға қарсы тұрақтылық, үйкеліс коэффициенті, беткі энергия және антибактериалдық белсенділік) мақсатты түрде жақсарту – олардың химиялық құрамын, микро- және нанокұрылымдық ерекшеліктерін, сондай-ақ бүрку технологиясының параметрлерін оңтайландыру арқылы мүмкін болады [13–15]. Бұл тәсілдер жабындардың қолданбалы тиімділігін арттырып, оларды биомедициналық мақсаттарда кеңінен пайдалануға жол ашады.

Осы тұрғыда титан мен алтынның қорытпаларына, атап айтқанда Ti-Au қосылуына ерекше назар аударылады, себебі олар жоғары биосәйкестікпен және қаттылықпен ерекшеленеді. Соңғы зерттеулер мұндай қорытпалардың жоғары механикалық төзімділікке ие болатынын көрсетіп отыр, бұл оларды имплантаттарға арналған жабындар үшін тиімді үміткерлер қатарына қосады. Ti-Au құрылымына Ag және Cu микроэлементтерін енгізу арқылы тек антимикробтық қасиеттерге қол жеткізіп қана қоймай, сонымен қатар тозуға төзімділік сипаттамаларын да жақсартуға болады. Ал бұл тәсіл – тозуға төзімділік пен антимикробтық тиімділікті бір уақытта арттыру – медициналық құрылғылардың қызмет ету мерзімін ұзарту және инфекциялық асқыну қаупін азайту үшін перспективті бағыт.

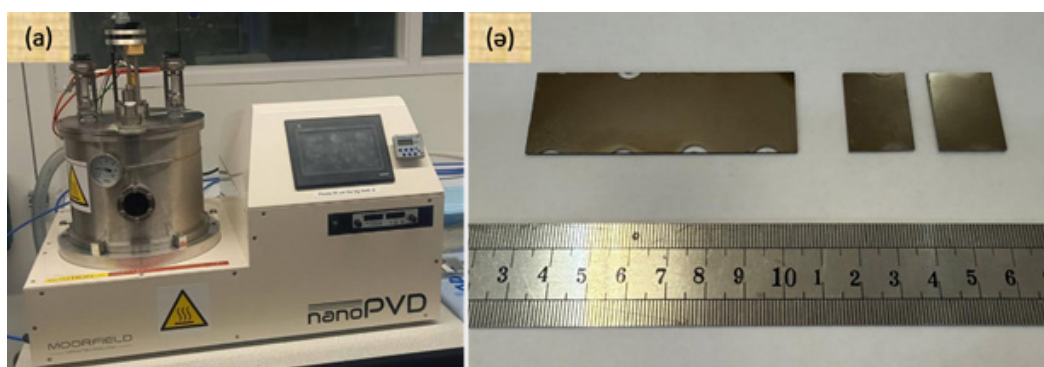
Жоғарыда айтылған жұқа қабықшалы жабындарды синтездеу технологиясына келер болсақ – магнетронды бүрку (PVD) немесе газды фазадан физикалық тұндыру мұндай жабындарды алудың ең перспективті әдістерінің бірі екенін дәлелдеп отыр, себебі ол жабынның құрылымы, құрамы және қалыңдығы үстінен жоғары деңгейде бақылау жүргізуге мүмкіндік береді. Атап айтқанда, көпмақсатты нысаналарды қолдану және тұндыру режимдерін оңтайландыру жабынның морфологиясы мен фазалық құрамын басқаруға мүмкіндік береді, оны биомедициналық қолданудың нақты талаптарына бейімдеуге жағдай жасайды [16–29].

Бұл ғылыми зерттеу жұмысы инновациялық заманауи технология ретінде мойындалған «NanoPVD» қондырғысы арқылы синтезделген Ti-Au негізіндегі жұқа қабықшалы жабындарды және олардың құрамына Ag (күміс) мен Cu (мыс) элементтерін қосымша енгізу арқылы алынған модификацияланған жүйелерді кешенді түрде зерттеуге бағытталған. Жұмыстың басты мақсаты – алынған жабындардың микроқұрылымдық ерекшеліктерін, физикалық-механикалық сипаттамаларын (қаттылық, тозуға төзімділік, үйкеліс коэффициенті және т.б.), сондай-ақ электрохимиялық коррозиялық төзімділігін жүйелі түрде бағалау. Сонымен қатар, зерттеу барысында алынған нәтижелер негізінде осы жабындарды биомедициналық имплантаттардың беттерін модификациялау және функционалдық инженерлік жабындар ретінде қолдану мүмкіндіктерін қарастыру. Бұл бағыттағы жұмыстар биосәйкестілік, коррозиялық тұрақтылық және беткі морфологияны басқару тұрғысынан перспективті шешімдерді айқындауға мүмкіндік береді, әрі биомедицина мен материалтану салаларындағы қолданбалы зерттеулер үшін өзекті.

Материалдар мен әдістер

Бұл зерттеуде химиялық құрамды дәл бақылауға мүмкіндік беретін және үлкен аумаққа біртекті тығыз жұқа жабын түзу қабілеті бар синтездеу әдісі ретінде (NanoPVD) иннова-

циялық магнетрондық бүрку әдісі қолданылды. Магнетронды (PVD) бүрку әдісімен Ti-AuX (X=Ag, Cu) негізіндегі жұқа жабындарды синтездеп алу үшін төсеніш ретінде зертханалық шыны және Ti_6Al_4V жоғары сапалы медициналық қорытпа қолданылды. Аталған жұқа жабындар Ұлыбританияның Moorefield Nanotechnology Ltd компаниясының «NanoPVD-S10A» магнетрондық бүрку жүйесі арқылы синтезделді (1а-суретті қараңыз). $76 \times 25 \times 1$ мм³ өлшемдегі Ti_6Al_4V төсеніштер SiC қағазын (P240-тен P4000-ге дейінгі кедір-бұдырлықпен) пайдалана отырып, жылтырату станогында өңделіп, бетінің кедір-бұдырлығы 40 нм-ден аспайтын тегіс күйге дейін жылтыратылды және кез келген бағытта бетінің кедір-бұдырлығы өлшеу жүйесі Alicona Infinite Focus арқылы тексерілді. Содан кейін жылтыратылған Ti_6Al_4V төсеніштер бірдей өлшемді ұсақ бөліктерге (19×25 мм²) кесіліп, стандартты зертханалық әйнек пластиналарымен ($76 \times 25 \times 1$ мм³) бірге 5:1 қатынасындағы су ерітіндісіндегі Decon 90 тазарту ерітіндісімен, кейін ультрадыбыстық ваннада, сондай-ақ изопропанол (IPA) мен ацетонмен жуу арқылы беткі ластанулардан тазартылды (1ә-сурет). Осыдан кейін төсеніштер құрғақ азотпен кептіріліп, плазма көзіне дейінгі арақашықтығы 100 мм болатын күйде бүрку камерасына орналастырылды және жұқа қабықшалы жабын тұндыру процесі кезінде тұрақты 5 айн/мин жылдамдықпен айналдырылды. Жұқа қабықшалы жабын тұндыру үшін бір-біріне симметриялы орналастырылған екі магнетрон қолданылып, олар бір мезгілде екі бүтін материалды шашырата отырып, интерметаллиттік қосылыстардың түзілуін қамтамасыз ететін бірлескен бүрку режимінде жұмыс істеді. Бір магнетрон диаметрі 50,8 мм және қалыңдығы 6,35 мм болатын мозаикалық титан (Ti) нысанымен жабдықталып, (DC) тұрақты ток көзіне қосылды, ал екінші магнетрон дәл осындай өлшемдегі стандартты алтын (Au) нысанымен жабдықталып, (RF) радиожиілікті қуат көзіне қосылды. Мозаикалық титан (Ti) нысаны легирлеуші элементтер Ag және Cu-ды орналастыру үшін шашырату бетінде диаметрі 3 мм және тереңдігі 4 мм болатын 8 түйік тесігі бар етіп жасалды.



Сурет 1 – «NanoPVD-S10A» қондырғысының сыртқы көрінісі (а) және алынған әртүрлі төсеніштердегі Ti-Au негізіндегі жұқа жабындар (ә)
(зертханалық әйнек және Ti_6Al_4V қорытпасы)

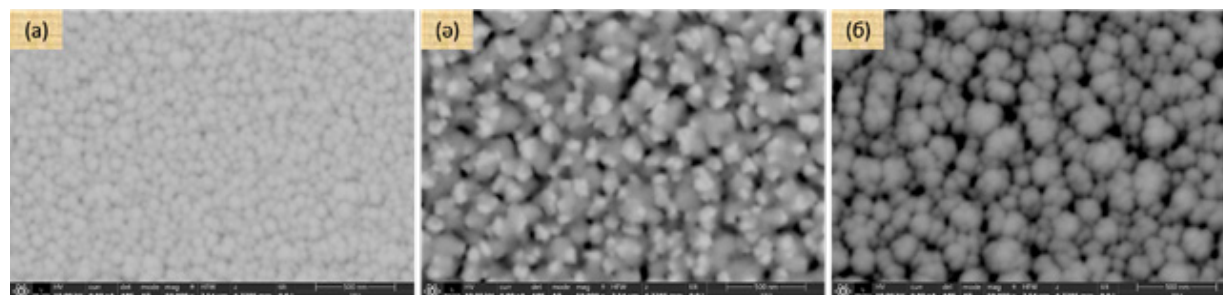
NanoPVD әдісімен алынған жұқа қабықшалы Ti-Au негізіндегі жабынның химиялық құрамын талдау және беткі морфологиясын зерттеу үшін рентгендік энергия-дисперсиялық спектрометрия, сондай-ақ растрлық электронды микроскопия (SEM/EDX) әдістері қолданылды. Зерттеулер «Scios 2, ThermoScientific» және «TESCAN VEGA» құрылғыларында 10–15 кВ кернеулер аралықтарында жүргізілді. Жабындардың кристалдық құрылымын зерттеу үшін рентгендік дифракция (XRD) әдісі қолданылды. Зерттеулер «D8 VENTURE» (BRUKER) негізгі рентгендік дифрактометрі мен қосымша «MiniFlex II» (RIGAKU) үстелдік шағын рентгендік дифрактометр арқылы 40 кВ кернеу мен 40 мА токта, мыс анодымен ($\lambda = 1,5418$ Å) жүргізілді. Сканерлеу қадамы $0,05^\circ$, әр қадамда ұстау уақыты 0,2 с құрады, ал бұрыштар диапазоны 20° -нан 90° -қа дейін болды. Жабындардың беттік морфологиясын мұқият нанодәрежеде зерттеу үшін атомдық-күштік микроскопия (AFM) әдісі қолданылды.

AFM талдауы плазмалық тазартудан өткен бастапқы күйіндегі үлгілерде жүргізілді. Бейнелеу жұмыстарының барлығы тербелмелі режимде орындалды және әрбір үлгі үшін кемінде 3 аймақ сканерленді. Жабындардың наноқаттылығын анықтау «FISCHERSCOPE HM2000 S» жүйесімен жүргізілді.

Ti-Au негізіндегі жабындарының бетінің үйкеліс коэффициентін анықтау үшін трибологиялық зерттеулер $26,14^{\circ}\text{C}$ бөлме температурасымен $71,44\%$ ылғалдылықта ауа атмосферасында «шар-диск» схемасы бойынша «Tribometer TRB³» (ANTON PAAR) құрылғысында жүргізілді. Үлгілердің беті зерттеу алдында $\text{C}_2\text{H}_5\text{OH}$ сұйықтығымен тиянақты тазартылды. Қарсы дене ретінде диаметрі 6 мм болатын, престелген сертификатталған 100Cr6 материалынан жасалған шар қолданылды. Шарға түсірілген жүктеме 1 Н, сырғу жылдамдығы – 1 см/с, тозу ізінің радиусы – 2,5 мм, ал қарсы дененің жалпы жүріп өткен жолы 70 м болды. Талдау шарттары халықаралық ASTM G133-95 және ASTM G99 стандарттарына сәйкес жүргізілді.

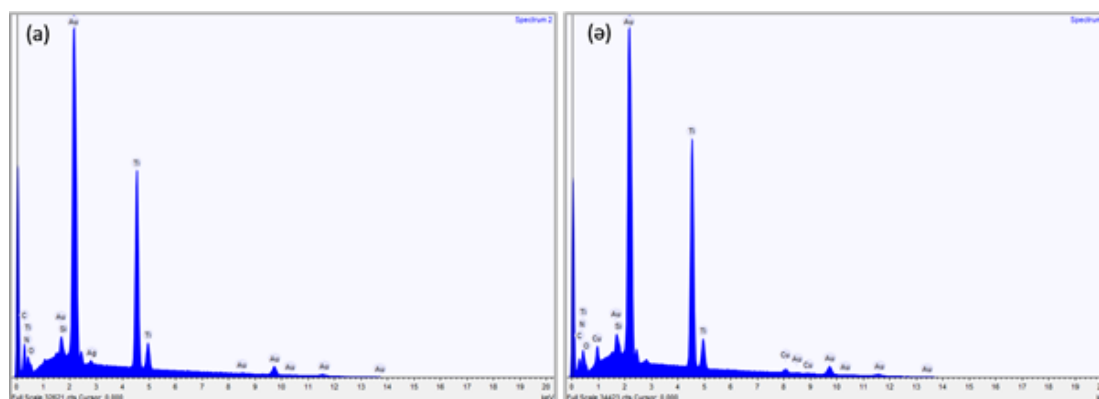
Нәтижелер мен талқылау

Растрлы-сканерлеуші электронды микроскопиядан (SEM) алынған деректерге сәйкес, PVD әдісімен титан-алтын (Ti-Au) негізіндегі жұқа қабықшалы жабындарды синтездеу беттің морфологиясын қалыптастыруда белгілі бір ерекшеліктерге ие екені айқындалды. 2-суретте «NanoPVD-S10A» магнетрондық бүрку жүйесі арқылы синтезделген Ti-Au негізіндегі жұқа қабықшалы жабындар бетінің SEM-суреті көрсетілген. Жабын вакуумды-термиялық бүрку әдістеріне тән типтік құрылымға ие: жоғары тығыздық, біртекті құрамы мен өте жұқа қабат қалыңдығы байқалады. Көріп тұрғанымыздай, беткі қабат анық реттелген құрылым мен құрамға ие. PVD кезінде таза титаннан алынған жұқа қабықша жақсы бағдарланған, үшбұрышты пирамида тәрізді түйіршіктер түзетіні белгілі [16, 27], ал Ti-Au жұқа қабықшалары өте тығыз күмбез тәрізді түйіршіктерге ие болатыны алынған нәтижелерден байқалды (2а-сурет). SEM-көріністерінен Ag және Cu элементтерінің аз мөлшерде қосылуы (2ә, б-суреттер, сәйкесінше) Ti-Au матрицасының сұр түсті фонында біркелкі таралған ашық түсті қырлы көп бұрышты (2ә-сурет) және сфералық (2б-сурет) қосымша түйіршіктердің түзілуіне алып келетінін көрсетеді. Осындай қырлы-дөңгелек түйіршіктер бұған дейін мыс (Cu) титанды (Ti) матрицаға қосылған кезде де байқалған болатын [28, 29]. Азот (N_2) газымен қосымша өңделген Ti-Au негізіндегі үлгілер, негізгі матрица түйіршіктерінің ұсақталуы (нанотүйіршіктерге айналуы), азотпен қосылыс түзілгендіктен SEM зерртеуіне спецификалық (азот әсерінен информативті ақпаратты дұрыс қабылдамау) кедергі келтіруінің себебінен бұл жерде көрсетілмеді. SEM бақылаулары жабындардың морфологиялық құрылымында түбегейлі өзгерістер орын алғанын көрсетті. Сонымен қатар, азотпен әрекеттесу нәтижесінде түзілген қосылыстардың болуы SEM әдісінің ажыратымдылық мүмкіндіктеріне әсер етіп, сурет контрастының төмендеуіне және бейненің бұлдырауына (diffusion effect) әкелгені байқалады. Бұл азоттың жабын құрылымына терең еніп, морфологиялық және электрондық қасиеттерге елеулі әсер ететінін айғақтайды.



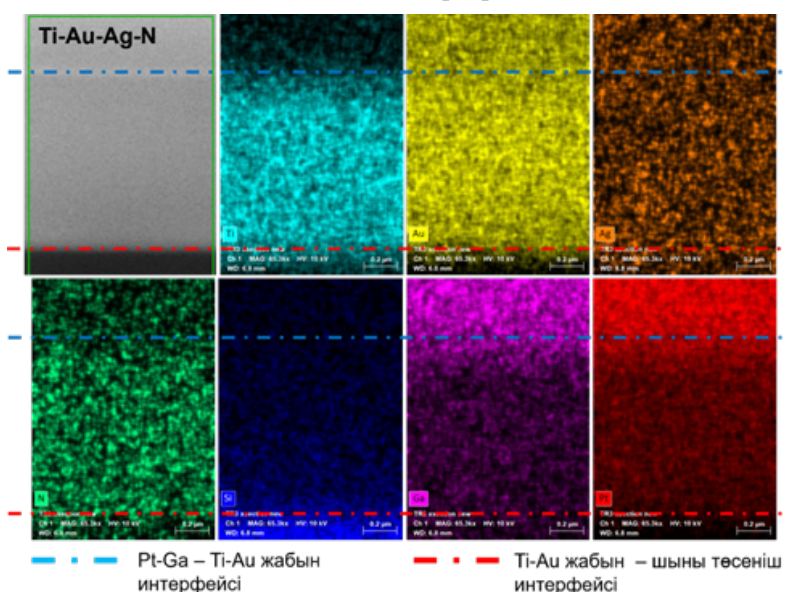
Сурет 2 – Титан-Алтын (Ti-Au) негізіндегі жұқа жабындар бетінің SEM-көріністері:
(а) – таза Ti-Au, (ә) – күміс (Ag) қосылған Ti-Au және (б) – мыс (Cu) қосылған Ti-Au

Ti-Au негізіндегі жабындардың химиялық құрамын сипаттау мақсатында энергия-дисперсиялық рентгендік спектроскопия (EDX) әдісімен элементтік талдау жүргізілді. Бұл талдау жабын бетінің бес түрлі аймағында алынған бес өлшеу нүктесінің орташа мәндеріне негізделіп жүргізіліп, нәтижелері 3-суретте гистограмма және 1-кесте ретінде көрсетілді. Бұдан бөлек, жабындардың көлденең қималарына бағытталған EDX-талдау жасалды, ол кескіндегі жасыл түспен белгіленген ауданның толық көлемін қамтыды. Көлденең қима бойынша жүргізілген талдау нәтижелері барлық үлгілерде титан (Ti) мен алтын (Au) элементтерінің салыстырмалы түрде біркелкі таралғанын көрсетті. Бұдан бөлек, жабындардың көлденең қималарына бағытталған EDX-талдау жасалды, ол кескіндегі жасыл түспен белгіленген ауданның толық көлемін қамтыды. Көлденең қима бойынша жүргізілген талдау нәтижелері барлық үлгілерде титан (Ti) мен алтын (Au) элементтерінің салыстырмалы түрде біркелкі таралғанын көрсетті.



Сурет 3 – Азот (N_2) газымен өңделген Ti-Au жабындарының SEM/EDX-талдау көрінісі:
(a) – күміс (Ag) қосылған және (б) – мыс (Cu) қосылған (элементтік талдау жабынның бетіне жүргізілген)

Азот (N_2) газымен иондық-бұрку процесі кезінде вакуумдық камерада жүретін физикалық-химиялық өзара әрекеттесу нәтижесінде Ag және Cu элементтерінің бүкіл жабын көлеміне біркелкі енуіне ықпал ететінін көрсетеді (4-сурет). Аталған деректер азотпен (N_2) модификациялау жабын құрамындағы легирлеуші элементтердің таралуын оңтайландыруға мүмкіндік беретін тиімді тәсіл екенін дәлелдейді [30].



Сурет 4 – Азот (N_2) газымен өңделген Ag қосылған Ti-Au негізіндегі жабынның көлденең қимасына EDX-талдау нәтижесі

Жүргізілген элементтік талдау нәтижелері бойынша алынған спектрлерде оттегі (O) мен кремний (Si) элементтерінің байқалуы, негізінен, EDX талдауы үшін қолданылған зертханалық төсеніштердің құрамымен түсіндіріледі. Атап айтқанда, бұл элементтердің пайда болуы кремний диоксидінен (SiO_2) жасалған арнайы шыны төсеніштердің әсерінен туындаған. Сонымен қатар, үлгілердің көлденең қималарын алу барысында қолданылған фокусталған иондық шокпен біріктірілген сканерлейтін электрондық микроскопия (FIB-SEM) әдісінің нәтижесінде жабын бетінде галлий (Ga) және платина (Pt) элементтері тіркелді. Бұл элементтер GaPt негізіндегі уақытша өткізгіш қабат ретінде жабын бетіне әдейі енгізілген, ол FIB әдісімен өңделетін аймақты дәл локализациялау және иондық бүрку кезінде жабынның бүлінуін алдын алу мақсатында қолданылатын өте танымал әдіс болып есептеледі [30].

Жабындардың морфологиялық сипаттамаларын жан-жақты зерттеу мақсатында, SEM нәтижелерімен қатар, Ti-Au жабындардың беткі рельефінің кеңістіктік құрылымын, микро-және наномасштаптағы ерекшеліктерін, сондай-ақ олардың орташа кедір-бұдырлық көрсеткіштерін (R_q) нақты әрі сапалы бағалау үшін қосымша зерттеулер атомдық-күштік микроскопия (AFM) әдісі арқылы жүргізілді. Бұл әдіс жабынның үшөлшемді беткі профилін жоғары дәлдікпен визуализациялауға және беткі құрылымдық элементтердің өлшемдері мен таралу сипатын сандық түрде сипаттауға мүмкіндік береді.

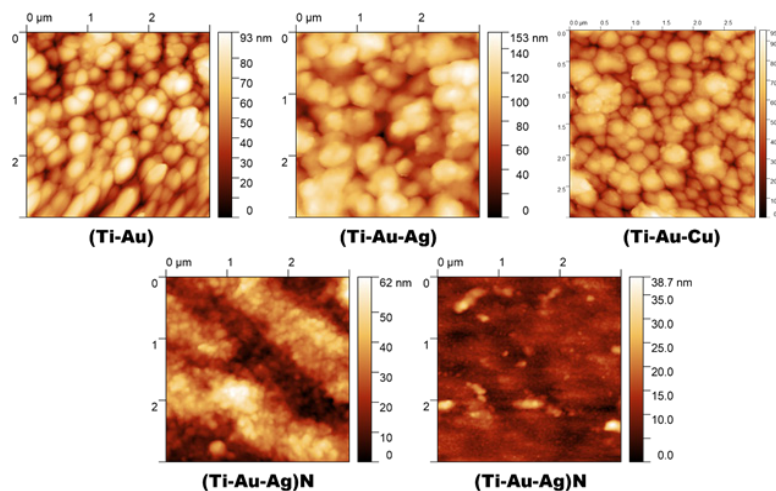
Кесте 1 – Ti-Au негізіндегі жабындарды SEM/EDX әдісі арқылы элементтік талдау кезіндегі алынған нәтижелері

Үлгі №	Орташа қалыңдығы (нм)	EDX (ат. %) Орт. 5 нүкте мен 5 түрлі аймақтағы орташа мәні						Ti/Au қатынасы
		Ti	Au	Ag	Cu	N	Total	
1	1167	73,3	26,7	-	-	-	100	2,7
2	1044	65,2	32,4	2,4	-	-	100	2,0
3	1046	60,7	27,7	0,2	-	11,4	100	2,2
4	1549	64,6	19,8	-	15,6	-	100	3,3
5	1101	63,2	20,1	-	2,9	13,8	100	3,1

5-суретте Ti-Au негізіндегі жұқа қабықшалы жабындардың беткі морфологиясын сипаттайтын AFM бейнелері ұсынылған, олар жабындардың бетінің біртектілігі мен құрылымдық элементтерінің өлшемдік сипаттамалары туралы қосымша мәлімет береді. AFM арқылы алынған талдау нәтижелері күміс (Ag) элементін қосу жабын бетінің морфологиясына айтарлықтай әсер ететінін көрсетті. Атап айтқанда, 4-суретте ұсынылғандай, Ag қоспасының енгізілуі нәтижесінде жабын бетінде ірілеу рельефті және көпқырлы құрылымдардың түзілуі байқалды. Бұл морфологиялық өзгерістер жабын бетінің энергиясының қайта таралуымен және кристалдану үдерістерінің өзгеруімен байланысты болуы мүмкін. AFM әдісімен алынған жоғары ажыратымдылықтағы беткі бейнелер жабын бетінің салыстырмалы түрде жоғары біртектілігін және төмен кедір-бұдырлық мәндерін көрсетті, бұл SEM арқылы алынған морфологиялық бақылауларды растайды. Жабындардың беткі кедір-бұдырлығы бойынша өлшенген орташа мәндері келесідей болды: 1-ші үлгі (Ti-Au) үшін – $13,4 \pm 1,0$ нм; 2-ші үлгі (Ti-Au-Ag) үшін – $17,4 \pm 1,8$ нм; 3-ші үлгі (Ti-Au-Ag-N) үшін – $8,7 \pm 2,4$ нм; 4-ші үлгі (Ti-Au-Cu) үшін – $13,5 \pm 2,7$ нм; және 5-ші үлгі (Ti-Au-Cu-N) үшін – $2,3 \pm 0,8$ нм. Бұл деректер Ag, Cu және N_2 (азот) өңдеудің жабындардың беткі құрылымына және микрогеометриялық сипаттамаларына елеулі әсер ететінін дәлелдейді.

SEM/EDX және AFM әдістерімен келтірілген нәтижелерге сүйене отырып, күміс (Ag), мыс (Cu) бөлшектерінің агломерациясы Ti-Au негізіндегі жұқа қабықшалар мен жабындардың құрылымдық-морфологиялық сипаттамаларына айтарлықтай әсер ететіні анықталды. Атап айтқанда, Ag мен Cu бөлшектерінің бірігуі (агломерациясы) кристаллдық торлардың бағдарлану бағытын өзгертуге, кристаллдық түйін пішіні мен өлшемін қайта қалыптастыруға, қабықшаның көлденең қимасының құрылымын түрлендіруге, оның тығыздығына және беткі

кедір-бұдырлық деңгейіне тікелей ықпал етеді. Бұл құбылыс жабынның жалпы физикалық-механикалық қасиеттерін анықтайтын негізгі факторлардың бірі және легирлеуші элементтердің морфологиялық тұрақтылыққа әсерін түсіну тұрғысынан маңызды ғылыми және практикалық мәнге ие.



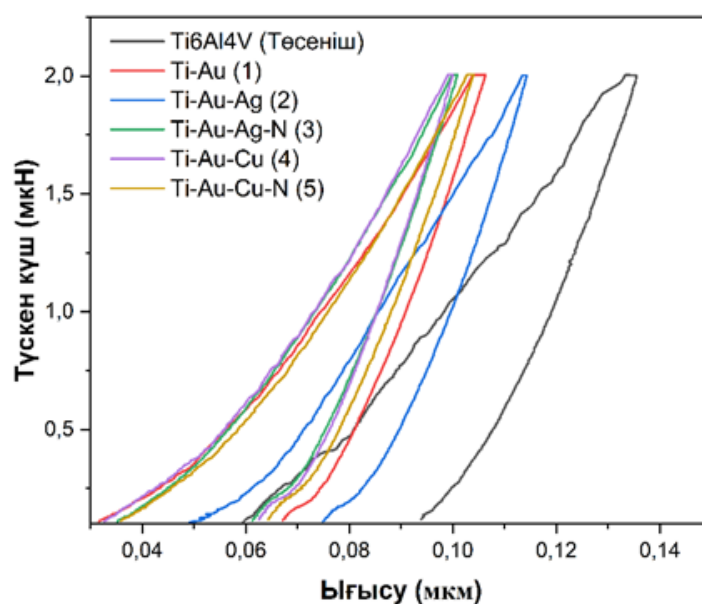
Сурет 5 – AFM-талдау нәтижесінде алынған жабындардың 2D көріністері

Ti-Au негізіндегі алынған жабындардың фазалық құрамын анықтау мақсатында рентгендік дифракциялық талдау (XRD) жүргізілді. Белгілі болғандай, XRD әдісінде қолданылатын рентген сәулесінің материалға ену тереңдігі сәуленің түсу бұрышына тәуелді болып және орта есеппен шамамен 0,3 мкм-ды құрағандықтан. Осыны ескере отырып, аталған зерттеу әдісінің шынайы нәтижелерін алу үшін зертханалық шыны төсеніштеріне тұндырылған жабын үлгілері пайдаланылды. Бұл тәсіл жабынның тек өзіне тән дифракциялық сигналдарды тіркеуге мүмкіндік береді және төсеніш материалының ($\text{Ti}_6\text{Al}_4\text{V}$) ықтимал рентгендік сигналдарының тіркелу ықтималдығын барынша төмендетеді. Сонымен қатар, жабындардың жалпы қалыңдығы орташа есеппен 1 мкм (1-кесте) құрағандықтан, XRD өлшеулерінде шыны төсеніштер қолданылған жағдайда ғана фазалық құрамды бұрмаламай анықтау мүмкіндігі жоғары болуы ескерілді.

Ti-Au жүйесіндегі элементтерді жоғары температурада, атап айтқанда шамамен 450 °C температурада физикалық вакуумдық бүрку (PVD) әдісімен тұндыру барысында, термодинамикалық тұрғыдан тұрақты $\beta\text{-Ti}_3\text{Au}$ интерметаллидтік фазаның түзілуі жүзеге асады. Бұл құбылыс температураның материал құрылымына әсер етуімен, сондай-ақ Ti және Au элементтерінің өзара әрекеттесу кинетикасының жоғарылауымен түсіндіріледі. XRD әдісі арқылы жүргізілген фазалық талдау нәтижелеріне байланысты, Ti-Au, Ti-Au-Ag және Ti-Au-Cu үлгілері үшін құрылымы $\beta\text{-Ti}_3\text{Au}$ (кубты кристалл кеңістіктік тобы: Pm-3n , $a=5,09600 \text{ \AA}$) негізгі фазадан құралатындығын көрсетті. Сонымен қатар, жабын құрамында қосымша аз мөлшерде $\alpha\text{-Ti}_3\text{Au}$ (кубты кристалл кеңістіктік тобы: Pm-3m , $a=4,14700 \text{ \AA}$) фазалық құрылымын бар екенін анықтады. Ti-Au-Ag-N негізіндегі нитридті үлгі үшін дифракциялық спектрлерде айқын өзгерістер құрылымдық орналасуда едәуір ерекшеліктері болатынын көрсетеді. $\beta\text{-Ti}_3\text{Au}$ құрылымына тиесілі (101), (200), (201), (211), (202), (301), (222), (230), (321), (400), (411) және (402) жазықтықтарына сәйкес келетін шыңдар жоғалып, оның орнына TiN (кубты кристалл кеңістіктік тобы: Fm-3m , $a=4,10500 \text{ \AA}$) фазалық құрылымына тиесілі (111), (200), (202), (311), (222) және (400) жазықтықтарына сәйкес келетін жаңа шыңдар пайда болғанын көруге болады. TiN мен AgN (кубты кристалл кеңістіктік тобы: Fm-3m , $a=4,10900 \text{ \AA}$) фазалық құрылымдары бір VN (225) типті кубты кеңістіктік тобына жататындықтан, кристалдық торлардың ұштарында Ti, Cu және Ag элементтерінің атомдары орналасуына негіз болады. Ал Ti-Au-Cu-N үлгісінде $\beta\text{-Ti}_3\text{Au}$ фазалық құрылымы азайып, оның орнына $\alpha\text{-Ti}_3\text{Au}$ құрылымы

артады. Бұл үлгіде α -Ti₃Au фазалық құрылымына тиесілі (101), (200), (201), (211), (300) және (202) жазықтықтарына сәйкес келетін шыңдар арқылы дәлелдейміз.

Ti-Au негізіндегі үлгілердің наноқаттылығын анықтау сынағынан алынған жүктемемен ығысу қисықтық графигі алмаз инденторының ұшы мен оның астындағы жұқа қабықшалы материал арасындағы өзара әрекеттесу туралы ақпаратты 6-суреттен көруімізге болады. Ti-Au үлгілері және жабынсыз төсеніш ретінде қолданылған Ti₆Al₄V қорытпаны индентациялау нәтижелері салыстырылып көрсетілген. Көрсетілген қисықтардың барлығы ең жоғары жүктемесі 2000 мкН болатын тұрақты максимум жүктемеде, жүктеп ұстау- жүктемеден босату кезеңдерінің әрқайсысының ұзақтығы 20 с аралығында орындалған индентация нәтижелерінен алынған [31, 32].



Сурет 6 – Ti-Au негізіндегі жабын мен Ti₆Al₄V төсенішіне жасалған жүктемемен ығысу қисықтық графигі (наноқаттылықты анықтау)

Жабынсыз Ti₆Al₄V төсеніш үшін индентор ұшы ең жоғары 2000 μ N жүктемеде 175 нм-ден астам тереңдікке енеді, содан кейін жүктеме алынғаннан кейін есептелген 134 нм байланыс тереңдігіне дейін серпімді қалпына келеді. Серпімділік модулі 81,4 ГПа және механикалық қаттылықты шамамен 4 ГПа болды. Ti-Au (1) жабыны бар үлгісі үшін индентор ұшы ең жоғары жүктемеде – 102,8 нм тереңдікке енеді және жүктеме алынғаннан кейін 67 нм байланыс тереңдігіне қайта оралады. Бұл қисық механикалық қаттылығы 12,6 ГПа және серпімділік модулі 190,8 ГПа болатын берік жұқа қабатқа сәйкес келеді. Ag қосылған Ti-Au (2) үлгісі индентор ұшының енуі тереңдігі 132 нм-ге жетіп, кейін серпімді түрде 75 нм байланыс тереңдігіне қалпына келеді. Бұл үлгі үшін серпімділік модулі мен механикалық қаттылық 168,1 ГПа және 13,2 ГПа құрайды, сәйкесінше. Ti-Au-Ag-N (3) үлгісі индентор ұшының енуі тереңдігі 91 нм-ге жетіп, кейін серпімді түрде 61 нм байланыс тереңдігіне қалпына келеді. Бұл үлгі үшін серпімділік модулі мен механикалық қаттылық 251,2 ГПа және 16,8 ГПа құрайды, сәйкесінше. Cu қосылған Ti-Au (4) үлгісі индентор ұшының енуі тереңдігі 100 нм-ге жетіп, кейін серпімді түрде 62,5 нм байланыс тереңдігіне қалпына келеді. Бұл үлгі үшін серпімділік модулі мен механикалық қаттылық 282,5 ГПа және 16,2 ГПа құрайды, сәйкесінше. Азотты атмосферада алынған Ti-Au-Cu-N (5) үлгісі индентор ұшының енуі тереңдігі 104 нм-ге жетіп, кейін серпімді түрде 64,4 нм байланыс тереңдігіне қалпына келеді. Бұл үлгі үшін серпімділік модулі мен механикалық қаттылық 219,8 ГПа және 17,3 ГПа құрайды, сәйкесінше.

Қорытынды

Сонымен, зерттеу нәтижелері бойынша қорытындылай келе, NanoPVD әдісімен алынған Ti-Au негізіндегі жұқа жабындар орташа қалыңдығы 1 мкм құрады және олардың қасиеттері технологиялық параметрлерге тікелей тәуелді екені анықталды. Ag және Cu қоспалары жабын морфологиясына айтарлықтай әсер етіп, беттік кедір-бұдырлық мәндерін тиісінше $17,4 \pm 1,8$ нм және $13,5 \pm 2,7$ нм-ге өзгертті, ал азотпен өңдеу бұл көрсеткішті $2,3 \pm 0,8$ нм-ге дейін төмендетті. XRD талдауы негізгі β -Ti₃Au фазасының түзілуін және азотты ортада TiN немесе α -Ti₃Au фазаларының артуын көрсетті. Наноқаттылық сынақтары Ti₆Al₄V төсенішіне қарағанда барлық үлгілерде жоғары нәтижелер берді: ең жоғары мән Ti-Au-Cu-N негізіндегі жабында – 17,3 ГПа, серпімділік модулі 219,8 ГПа болды. Трибологиялық сынақтар жабындардың үйкеліс коэффициентін едәуір төмендетіп, тозуға төзімділігін арттырды. Осылайша, алынған жабындар морфологиясы, механикалық беріктігі және тозуға төзімділігі жағынан биомедициналық имплантаттар үшін перспективті материалдар ретінде қолдануға жасалған зерттеулер толық негіз бола алады.

Қаржыландыру

Бұл зерттеу Қазақстан Республикасы Ғылым және жоғары білім Министрлігінің Ғылым Комитеті тарапынан қаржыландырылды (грант № AP22683206, BR24992854).

ӘДЕБИЕТТЕР

- 1 Wilches, L.V., Uribe, J.A. and Toro, A. Wear of materials used for artificial joints in total hip replacements. *Wear*, 265(1–2), 143–149 (2008). <https://doi.org/10.1016/j.wear.2007.09.010>.
- 2 Rahmati, B., Sarhan, A.A., Zalnezhad, E., Kamiab, Z., Dabbagh, A., Choudhury, D., and Abas W.A. Development of tantalum oxide (Ta-O) thin film coating on biomedical Ti-6Al-4V alloy to enhance mechanical properties and biocompatibility. *Ceramics International*, 42(1), 466–480 (2016). <https://doi.org/10.1016/j.ceramint.2015.08.133>.
- 3 Panjwani, B., Satyanarayana, N., and Sinha, S.K. Tribological characterization of a biocompatible thin film of UHMWPE on Ti6Al4V and the effects of PFPE as top lubricating layer. *Journal of the Mechanical Behavior of Biomedical Materials*, 4(7), 953–960 (2011). <https://doi.org/10.1016/j.jmbbm.2011.02.005>.
- 4 Guan, J., Jiang, X., Xiang, Q., Yang, F., and Liu, J. Corrosion and tribocorrosion behavior of titanium surfaces designed by electromagnetic induction nitriding for biomedical applications. *Surface and Coatings Technology*, 409, 126844 (2021). <https://doi.org/10.1016/j.surfcoat.2021.126844>.
- 5 Wei, M., Yu, H., Song, Z., Yin, Y., Zhou, X., Wang, H., et al. Microstructural evolution, mechanical properties and wear behavior of in-situ TiC-reinforced Ti matrix composite coating by induction cladding. *Surface and Coatings Technology*, 412, 127048 (2021). <https://doi.org/10.1016/j.surfcoat.2021.127048>.
- 6 Mu, J., Wang, H., Qin, B., Zhang, Y., Chen, L., and Zeng, C. Improved wear and corrosion resistance of biological compatible TiZrNb films on biomedical Ti6Al4V substrates by optimizing sputtering power. *Surface and Coatings Technology*, 428, 127866 (2021). <https://doi.org/10.1016/j.surfcoat.2021.127866>.
- 7 Wang, Y.H., Yang, Z.B., Hu, S.Y., Zhao, Y.H., Ren, H., Gong, F., and Xie, Z.W. Tailoring growth structure, wear and corrosion properties of TiN coatings via gradient bias and arc enhanced glow discharge. *Surface and Coatings Technology*, 450, 129015 (2022). <https://doi.org/10.1016/j.surfcoat.2022.129015>.
- 8 Özkan, D., Yılmaz, M.A., Szala, M., Türküz, C., Chocyk, D., Tunç, C., et al. Effects of ceramic-based CrN, TiN, and AlCrN interlayers on wear and friction behaviors of AlTiSiN+TiSiN PVD coatings. *Ceramics International*, 47(14), 20077–20089 (2021). <https://doi.org/10.1016/j.ceramint.2021.04.015>.
- 9 Çomaklı, O., Yazıcı, M., Demir, M., Yetim, A.F., and Celik, A. Effect of bilayer numbers on structural, mechanical, tribological and corrosion properties of TiO₂-SiO₂ multilayer film-coated β -type Ti45Nb alloys. *Ceramics International*, 49(2), 3007–3015 (2023). <https://doi.org/10.1016/j.ceramint.2022.09.285>.
- 10 Bai, H., Zhong, L., Kang, L., Liu, J., Zhuang, W., Lv, Z., and Xu, Y. A review on wear-resistant coating with high hardness and high toughness on the surface of titanium alloy. *Journal of Alloys and Compounds*, 882, 160645 (2021). <https://doi.org/10.1016/j.jallcom.2021.160645>.

- 11 Lopes, C., Gonçalves, C., Borges, J., Polcar, T., Rodrigues, M.S., Barradas, N.P., et al. Evolution of the functional properties of titanium–silver thin films for biomedical applications: Influence of in-vacuum annealing. *Surface and Coatings Technology*, 261, 262–271 (2015). <https://doi.org/10.1016/j.surfcoat.2014.11.020>.
- 12 Łępicka, M., Grądzka-Dahlke, M., Pieniak, D., Pasierbiewicz, K., Kryńska, K., and Niewczas A. Tribological performance of titanium nitride coatings: A comparative study on TiN-coated stainless steel and titanium alloy. *Wear*, 422, 68–80 (2019). <https://doi.org/10.1016/j.wear.2019.01.029>.
- 13 Birkett, M., Penlington, R., Wan, C., and Zoppi, G. Structural and electrical properties of CuAlMo thin films prepared by magnetron sputtering. *Thin Solid Films*, 540, 235–241 (2013). <https://doi.org/10.1016/j.tsf.2013.05.145>.
- 14 Ruggiero, A., D’Amato, R., and Gómez, E. Experimental analysis of tribological behavior of UHMWPE against AISI420C and against TiAl6V4 alloy under dry and lubricated conditions. *Tribology International*, 92, 154–161 (2015). <https://doi.org/10.1016/j.triboint.2015.06.005>.
- 15 Ait-Djafer, A.Z., Saoula, N., Aknouche, H., Guedouar, B., and Madaoui, N. Deposition and characterization of titanium aluminum nitride coatings prepared by RF magnetron sputtering. *Applied Surface Science*, 350, 6–9 (2015). <https://doi.org/10.1016/j.apsusc.2015.02.053>.
- 16 Lukose, C.C., Anastopoulos, I., Mantso, T., Bowen, L., Panayiotidis, M.I., and Birkett, M. Thermal activation of Ti(1-x)Au(x) thin films with enhanced hardness and biocompatibility. *Bioactive Materials*, 15, 426–445 (2022). <https://doi.org/10.1016/j.bioactmat.2022.02.027>.
- 17 Villapún, V.M., Lukose, C.C., Birkett, M., Dover, L.G., and González, S. Tuning the antimicrobial behaviour of Cu₈₅Zr₁₅ thin films in “wet” and “dry” conditions through structural modifications. *Surface and Coatings Technology*, 350, 334–345 (2018). <https://doi.org/10.1016/j.surfcoat.2018.06.094>.
- 18 Lukose, C.C., Chavignon, C., Mantso, T., Panayiotidis, M.I., and Birkett, M. Enhanced mechanical and biocompatibility performance of Ti(1-x)Ag(x) coatings through intermetallic phase modification. *Materials Characterization*, 194, 112401 (2022). <https://doi.org/10.1016/j.matchar.2022.112401>.
- 19 Birkett, M., Dover, L., Lukose, C.C., Zia, A.W., Tambuwala, M.M., and Serrano-Aroca, Á. Recent advances in metal-based antimicrobial coatings for high-touch surfaces. *International Journal of Molecular Sciences*, 23(3), 1162 (2022). <https://doi.org/10.3390/ijms23031162>.
- 20 Lukose, C.C., Anastopoulos, I., Panagiotidis, I.S., Zoppi, G., Black, A.M., Dover, L.G., et al. Biocompatible Ti₃Au–Ag/Cu thin film coatings with enhanced mechanical and antimicrobial functionality. *Biomaterials Research*, 27(1), 93 (2023). <https://doi.org/10.1186/s40824-023-00435-1>.
- 21 Macías, H.A., Yate, L., Coy, L.E., Olaya, J.J., and Aperador, W. Effect of nitrogen flow ratio on microstructure, mechanical and tribological properties of TiWSiN_x thin film deposited by magnetron co-sputtering. *Applied Surface Science*, 456, 445–456 (2018). <https://doi.org/10.1016/j.apsusc.2018.06.129>.
- 22 Baptista, A., Silva, F., Porteiro, J., Míguez, J., and Pinto, G. Sputtering physical vapour deposition (PVD) coatings: A critical review on process improvement and market trend demands. *Coatings*, 8(11), 402 (2018). <https://doi.org/10.3390/coatings8110402>.
- 23 Palmero, A., Alcalá, G., and Alvarez, R. Editorial for Special Issue: Nanostructured Surfaces and Thin Films Synthesis by Physical Vapor Deposition. *Nanomaterials*, 11(1), 148 (2021). <https://doi.org/10.3390/nano11010148>.
- 24 Yumusak, G., Leyland, A., and Matthews, A. The effect of pre-deposited titanium-based PVD metallic thin films on the nitrogen diffusion efficiency and wear behaviour of nitrided Ti alloys. *Surface and Coatings Technology*, 394, 125545 (2020). <https://doi.org/10.1016/j.surfcoat.2020.125545>.
- 25 Yumusak, G., Leyland, A., and Matthews, A. A microabrasion wear study of nitrided α -Ti and β -TiNb PVD metallic thin films, pre-deposited onto titanium alloy substrates. *Surface and Coatings Technology*, 442, 128423 (2022). <https://doi.org/10.1016/j.surfcoat.2022.128423>.
- 26 Thampi, V.A., Bendavid, A., and Subramanian, B. Nanostructured TiCrN thin films by Pulsed Magnetron Sputtering for cutting tool applications. *Ceramics International*, 42(8), 9940–9948 (2016). <https://doi.org/10.1016/j.ceramint.2016.03.095>.
- 27 Wojcieszak, D., Osekowska, M., Kaczmarek, D., Szponar, B., Mazur, M., Mazur, P., and Obstarczyk A. Influence of material composition on structure, surface properties and biological activity of nanocrystalline coatings based on Cu and Ti. *Coatings*, 10(4), 343 (2020). <https://doi.org/10.3390/coatings10040343>.
- 28 Stranak, V., Wulff, H., Ksirova, P., Zietz, C., Drache, S., Cada, M., et al. Ionized vapor deposition of antimicrobial Ti–Cu films with controlled copper release. *Thin Solid Films*, 550, 389–394 (2014). <https://doi.org/10.1016/j.tsf.2013.11.001>.

29 Karimi, A., and Cattin, C. Intermetallic β -Ti₃Au hard thin films prepared by magnetron sputtering. *Thin Solid Films*, 646, 1–11 (2018). <https://doi.org/10.1016/j.tsf.2017.11.027>.

30 Yeskermessov, D., Toleukhanova, Z., Aringozhina, Z., Tabiyeva, Y., & Ahmed, W. Hard and Superhard Multicomponent Nitride Coatings Deposited Using Vacuum-Arc Evaporation: The Effects of Cr and Si on the Structure and Properties of the Nanostructured (Zr-Ti-Nb)N Coatings. In *Innovations in Materials Chemistry, Physics, and Engineering Research* (pp. 136–196). IGI Global (2023). <https://doi.org/10.4018/978-1-6684-6830-2.ch004>.

31 Yeskermessov, D., Rakhadilov, B., Zhurerova, L., Apsezhanova, A., Aringozhina, Z., Booth, M., & Tabiyeva, Y. Surface modification of coatings based on Ni-Cr-Al by pulsed plasma treatment. *AIMS Materials Science*, 10(5), (2023). <https://doi.org/10.3934/matricsci.2023042>.

33 Rakhadilov, B., Bayatanova, L., Kengesbekov, A., Magazov, N., Toleukhanova, Z., & Yeskermessov, D. Study of the Influence of Air Plasma Spraying Parameters on the Structure, Corrosion Resistance, and Tribological Characteristics of Fe-Al-Cr Intermetallic Coatings. *Coatings*, 15(7), 790. (2025). <https://doi.org/10.3390/coatings15070790>.

¹Ескермесов Д.К.,

PhD, ассоциированный профессор, ORCID ID: 0000-0002-2206-8132,
e-mail: dyeskermessov@edu.ektu.kz

²Лукос С.С.,

PhD, научный сотрудник, ORCID ID: 0000-0001-9668-0079,
e-mail: c.c.lukose@northumbria.ac.uk

²Биркетт М.,

PhD, профессор, ORCID ID: 0000-0002-6135-6639,
e-mail: martin.birkett@northumbria.ac.uk

³Пазылбек С.А.,

PhD, ассоциированный профессор, ORCID ID: 0000-0002-2318-9155,
e-mail: sapargali.pa@gmail.com

¹Ескермесова А.С.,

магистр, м.н.с., ORCID ID: 0009-0004-8454-3615,
e-mail: ardakeskermesova@gmail.com

^{1*}Төлеуханова Ж.Т.,

магистр, ст. преподаватель, ORCID ID: 0009-0002-9254-3901,
*e-mail: ztoleukhanova@ektu.kz

¹Восточно-Казахстанский технический университет им. Д. Серикбаева,
г. Усть-Каменогорск, Казахстан

²Университет Нортумбрия, г. Ньюкасл, Великобритания

³Университет им. Ж.А. Ташенева, г. Шымкент, Казахстан

ПОЛУЧЕНИЕ И ИССЛЕДОВАНИЕ ТОНКИХ ПОКРЫТИЙ НА ОСНОВЕ ТИТАНА-ЗОЛОТА (Ti-Au) ДЛЯ БИМЕДИЦИНСКОГО ПРИМЕНЕНИЯ С УЛУЧШЕННЫМИ ФИЗИКО-МЕХАНИЧЕСКИМИ ХАРАКТЕРИСТИКАМИ

Аннотация

В данной научной статье исследованы структура и свойства поверхности тонких покрытий на основе Ti-Au, полученных методом PVD (магнетронного распыления) в вакуумной среде при высокой температуре 450 °C в режимах постоянного (DC) и высокочастотного (RF) тока. Тонкие покрытия, полученные на установке NanoPVD, имели среднюю толщину около 1 мкм. Экспериментальные исследования показали, что химический состав, микроструктура и свойства покрытий находятся в тесной зависимости от технологических параметров метода магнетронного распыления. Влияние добавок Ag и Cu на микроструктуру и механические свойства покрытий Ti-Au исследовалось с использованием растровой электронной микроскопии (SEM), энергодисперсионного микроанализа (EDX) и рентгеновской дифрактометрии (XRD). Характеристики шероховатости поверхности покрытий анализировались методом атомно-силовой микроскопии

(AFM), а для комплексной оценки изменений свойств были проведены испытания на нанотвердость и трибологические тесты. Показано, что модификация поверхности подложки Ti6Al4V методом PVD повышает ее износостойкость и снижает коэффициент трения. По результатам рентгенофазового анализа установлено, что улучшение трибологических свойств полученных тонких покрытий напрямую связано с образованием основной фазы Ti3Au и увеличением ее содержания.

Ключевые слова: покрытие и тонкие пленки, система магнетронного распыления, морфология, твердость, элементный состав.

¹Yeskermessov D.K.,

PhD, Associate Professor, ORCID ID: 0000-0002-2206-8132,

e-mail: dyeskermessov@edu.ektu.kz

²Lukose C.C.,

PhD, Researcher, ORCID ID: 0000-0001-9668-0079,

e-mail: c.c.lukose@northumbria.ac.uk

²Birkett M.,

PhD, Professor, ORCID ID: 0000-0002-6135-6639,

e-mail: martin.birkett@northumbria.ac.uk

³Pazylbek S.A.,

PhD, Associate Professor, ORCID ID: 0000-0002-2318-9155,

e-mail: sapargali.pa@gmail.com

¹Yeskermessova A.S.,

MSc., Junior Researcher, ORCID ID: 0009-0004-8454-3615,

e-mail: ardakeskermesova@gmail.com

^{1*}Toleukhanova Zh.T.

MSc., Sr. Lecturer, ORCID ID: 0009-0002-9254-3901,

*e-mail: ztoleukhanova@ektu.kz

¹D. Serikbayev East Kazakhstan Technical University», Ust-Kamenogorsk, Kazakhstan

²Northumbria University, Newcastle, United Kingdom

³Tashenev University, Shymkent, Kazakhstan

SYNTHESIS AND INVESTIGATION OF TITANIUM-GOLD (TI-AU) BASED THIN COATINGS FOR BIOMEDICAL APPLICATIONS WITH ENHANCED PHYSICAL-MECHANICAL PROPERTIES

Abstract

This scientific article investigates the structure and surface properties of Ti-Au-based thin coatings obtained by the PVD (magnetron sputtering) method in a vacuum environment at a high temperature of 450°C under direct current (DC) and radio frequency (RF) modes. The thin coatings deposited using the NanoPVD system had an average thickness of about 1 μm. Experimental studies showed that the chemical composition, microstructure, and properties of the coatings are closely dependent on the technological parameters of the magnetron sputtering process. The effect of Ag and Cu additions on the microstructure and mechanical properties of the Ti-Au coatings was studied using scanning electron microscopy (SEM), energy-dispersive X-ray spectroscopy (EDX), and X-ray diffraction (XRD). The surface roughness characteristics of the coatings were analyzed by atomic force microscopy (AFM), and nanoindentation and tribological tests were carried out to comprehensively evaluate the changes in their properties. It was shown that the modification of the Ti6Al4V substrate surface by the PVD method increases its wear resistance and reduces the coefficient of friction. X-ray phase analysis results revealed that the improvement in the tribological properties of the obtained thin coatings is directly related to the formation of the primary Ti3Au phase and the increase in its content.

Keywords: coating and thin films, magnetron sputtering system, morphology, hardness, elemental composition.

Мақаланың редакцияға түскен күні: 09.08.2025

УДК 533.95; 533.95:537.84
МРНТИ 29.27.00

<https://doi.org/10.55452/1998-6688-2025-22-4-386-400>

¹Коданова С.К.,

к.ф.-м.н., профессор, ORCID ID: 0000-0001-7098-471X,

e-mail: kodanova@physics.kz

^{1*}Исанова М.К.,

PhD, ассоциированный профессор, ORCID ID: 0000-0002-0264-2694,

*e-mail: issanova@physics.kz

¹Бастыкова Н.Х.,

PhD, ассоциированный профессор, ORCID ID: 0000-0002-8400-7183,

e-mail: bastykova_nuriya@physics.kz

¹Кенжебекова А.И.,

PhD, ст. преподаватель, ORCID ID: 0000-0003-0223-0216,

e-mail: kenzhebekova.a@kaznu.kz

²Хоффманн Д.,

доктор наук, профессор, ORCID ID: 0000-0003-0922-7802

e-mail: hoffmann@physik.tu-darmstadt.de

¹Казахский национальный университет им. аль-Фараби, г. Алматы, Казахстан

²Технический университет Дармштадт, г. Дармштадт, Германия

ИССЛЕДОВАНИЕ ОПТИЧЕСКИХ И ДИНАМИЧЕСКИХ СВОЙСТВ ВОДОРОДНОЙ ПЛАЗМЫ ПРИ ВЫСОКИХ ПЛОТНОСТЯХ

Аннотация

В работе приведены результаты исследования оптических свойств водорода, в частности коэффициентов отражения и преломления электромагнитных волн, при этом диэлектрическая проницаемость вещества описывается обобщенной моделью Друде–Лоренца. В настоящей работе проведено исследование продольных и поперечных спектров колебаний микроскопического ионного тока в водороде при различных значениях температуры и плотности, а также выполнен анализ влияния электронных обменных и корреляционных эффектов. Исследование проводилось на основе эффективного потенциала взаимодействия с учетом коррекции локального поля, полученного из моделирования методом квантового Монте-Карло. Применение точных моделей локального поля, таких как аппроксимации на основе квантово-статистических расчетов, позволяет достоверно воспроизводить транспортные и оптические свойства плотных электронных систем. В частности, учет функции локального поля приводит к существенным поправкам в расчетах оптических и динамических свойств, что критически важно для моделирования горячего плотного вещества, металлизированной плазмы и вырожденных электронных систем. Кроме того, наличие точной формы функции локального поля позволяет корректно описывать оптические и динамические свойства плазмы, включая коэффициенты отражения и поглощения, собственные моды колебаний.

Ключевые слова: плотная плазма, эффективный потенциал, оптические свойства, коэффициент отражения, коэффициент поглощения, собственные моды колебаний.

Введение

В последние десятилетия активно проводятся как экспериментальные, так и теоретические исследования, направленные на изучение поведения водорода, дейтерия и трития в широком диапазоне температур и плотностей. Физические свойства водорода при высоких давлениях и умеренных температурах имеют решающее значение для моделирования процессов, протекающих в недрах планет [1–2], а также представляют практический интерес для исследований

в области инерционного термоядерного синтеза [3–4], поскольку изотопы дейтерий и тритий используются в качестве рабочих материалов (DT-топлива) в подобных экспериментах.

Согласно результатам моделирования методом Монте-Карло, плотный жидкий водород способен испытывать фазовый переход первого рода в металлическое состояние, при котором критическая точка достигается при давлении около 50 ГПа и температуре порядка 10 000 К [5]. В работе [6] экспериментально зафиксировано существование металлической фазы водорода при давлении 140 ГПа и температуре около 3000 К. При экстремальных ударных давлениях также ожидается появление металлической проводимости в жидком состоянии. Для уточнения диапазона давлений, при котором ударно-сжатый водород переходит в проводящее состояние, в исследовании [7] были проведены измерения оптической отражательной способности фронта ударной волны в интервале 20–500 ГПа. Полученные данные показали, что возрастание оптического коэффициента отражения свидетельствует о достижении ударным дейтерием состояния, характерного для металлической жидкости.

С экспериментальной точки зрения измерения кривой Гюгонио и оптической отражательной способности ударно-сжатого водорода [8–9] представляют собой важный источник информации для современных вычислений уравнения состояния и оптических свойств вещества [10–11]. Эти результаты существенно расширяют представления о поведении водорода и дейтерия в условиях теплой плотной материи, характерных для имплозионных процессов в экспериментах по инерционному термоядерному синтезу (ИТС). Особый интерес среди работ, посвященных экспериментальному изучению изменений оптических характеристик в течение нескольких пикосекунд после лазерного облучения мишени, представляет исследование [12], где интерпретация экспериментальных данных была выполнена с использованием формулы Кубо–Гринвуда для описания электропроводности.

Взаимодействие лазерного излучения с веществом лежит в основе множества современных физических и технологических процессов, включая генерацию высокоэнергетических частиц и коротковолнового электромагнитного излучения [13–14]. При облучении твердой мишени интенсивными лазерными импульсами вещество проходит последовательные стадии преобразований: от исходного холодного твердого состояния к горячей плотной плазме, известной как теплая плотная материя (WDM), и далее к слабо связанной плазме, свойства которой динамически изменяются во времени и пространстве.

Материалы и методы

В данной работе рассматривается плотная водородная плазма, представляющая особый интерес для широкого круга направлений – от астрофизики и физики взаимодействия лазерного излучения с твердыми телами до физики высоких давлений. Теоретические и экспериментальные исследования плотного водорода во многом обусловлены интересом к проблематике теплой плотной материи в физике конденсированных сред, а также ее приложениями в области инерционного термоядерного синтеза. В условиях плотного вещества ключевую роль играют межчастичные взаимодействия. Для их описания разработан ряд эффективных моделей. Так, в классической плазме широко используется потенциал Дебая–Хюккеля, тогда как для вырожденных электронных систем применяется потенциал Томаса–Ферми. Возрастающий интерес к исследованиям разогретого плотного вещества и ИТС стимулировал развитие различных аналитических подходов, в частности приближений, описывающих экранирование заряда в частично и слабо вырожденных режимах [15–19]. Многие свойства плотной плазмы были получены с помощью квантово-механических расчетов методом Монте-Карло [20–21], результаты которых впоследствии использовались как исходные данные в ряде прикладных задач [22]. Эти данные легли в основу построения теории функционала плотности (DFT) при описании свойств реальных материалов [23].

В работе рассматривается получение эффективного потенциала взаимодействия в результате использования поправки локального поля, полученного из моделирования методом

квантового Монте-Карло в диэлектрической проницаемости [24]. Далее данный потенциал используется для исследования влияния электронных обменно-корреляционных эффектов на транспортные и оптические свойства водорода при различных параметрах с применением метода молекулярной динамики. Исследование оптических и динамических свойств водородной плазмы проводилось с помощью метода молекулярной динамики. Динамика $N = 10^4$ одинаковых частиц была получена путем решения уравнения движения по алгоритму Бимана (NVT ансамбль). Чтобы устранить граничные эффекты, использовались периодические граничные условия для основной ячейки с длиной стороны $L = (4\pi N/3)^{1/3}a$. Единица времени представлена обратной плазменной частотой ионов ω_p^{-1} .

Согласно теории линейного отклика экранированный потенциал иона зарядом (Ze) может быть вычислен с помощью статической функции диэлектрической функции как

$$\Phi(\vec{r}_i, \vec{r}_j) = \frac{Z^2 e^2}{(\vec{r}_j - \vec{r}_i)} + \int \frac{d^3 k}{(2\pi)^3} |\tilde{\phi}_{ei}(\vec{k})|^2 \chi_e(k) e^{i\vec{k}(\vec{r}_j - \vec{r}_i)} \quad (1)$$

где $\tilde{\phi}(k)$ – микрopotенциал в пространстве Фурье и χ – функция отклика плотности, который связан с диэлектрической функцией плазмы следующим соотношением

$$\varepsilon^{-1}(k, \omega) = 1 + \frac{4\pi e^2}{k^2} \chi_e(k, \omega) \quad (2)$$

Диэлектрическую функцию плазмы возьмем в виде:

$$\varepsilon(k, 0) = 1 - \frac{4\pi e^2}{k^2} \Pi, \quad \Pi(k) = \frac{\Pi_{RPA}(k, \omega=0)}{1 + \frac{4\pi e^2}{k^2} G(k, \omega=0) \cdot \Pi_{RPA}(k, \omega=0)} \quad (3)$$

здесь $\Pi(k)$ – поляризационная функция электронов и $\Pi_{RPA}(k)$ – поляризационная функция квантовых электронов при фиксированной температуре [25]:

$$\Pi_{RPA}(k, \omega) = -\frac{k^2 \chi_0^2}{16\pi e^2 z^3} [g(u+z) - g(u-z)], \quad (4)$$

здесь $u = \frac{\omega}{kv_F}$, $z = \frac{k}{2k_F}$, $\chi_0^2 = \frac{3}{16} \left(\frac{\hbar\omega}{E_F}\right)^2 = \frac{1}{\pi k_F a_B}$, $k_F = (3\pi^2 n_e)^{1/3}$, $\omega_p^2 = 3\pi n_e e^2 / m_e$,

$a_B = \hbar^2 / m_e e^2$ – первый радиус Бора, $\hbar$ – постоянная Планка, v_F – скорость Ферми, и

$$g(x) = -g(-x) = \int \frac{y dy}{\exp\left(\frac{y^2}{\theta} - \eta\right) + 1} \ln \left| \frac{x+y}{x-y} \right| \quad (5)$$

здесь $\theta = \frac{k_B T}{E_F}$ – параметр вырождения, $\eta = \frac{\mu}{k_B T}$ – химический потенциал, зависящий от плотности через отношение $\frac{2}{3} \theta^{-\frac{3}{2}} = I_{\frac{3}{2}}(\eta)$, и $I_{\frac{3}{2}}$ – Ферми интеграл порядком $\frac{3}{2}$.

Для учета обменно-корреляционных эффектов использовано эффективное статическое приближение (ESA), которое представлено в работе [26] в виде аналитической параметризации $G(k, \omega)$ на основе ab initio данных из моделирования методом квантового Монте-Карло (PIMC) для однородного электронного газа в условиях WDM. Если пренебречь электронными корреляциями, т.е. $G(k, \omega) = 0$, получаем приближение случайных фаз (RPA).

Таким образом, с увеличением плотности плазмы необходимо учитывать эффекты неидеальности путем введения функции локальных полей плазмы. Поправка $G(k, \omega)$ содержит полную информацию об обменно-корреляционных эффектах в диапазоне волн и частот.

Обратная диэлектрическая функция $\varepsilon^{-1}(k)$ на основе данных квантового Монте-Карло вычисляется с помощью уравнения (3), с использованием $G(k, \omega)$ и длинноволнового предела G_{CSR} .

Эффективная статическая аппроксимация (ESA) поправки локального поля (LFC) электронного газа, которые получены из квантовых данных Монте-Карло, представлены в работе [27] в следующем виде:

$$\overline{G_{ESA}}(q, r_s, \theta) = G_{CSR}(q, r_s, \theta) \left[\frac{1 + \alpha(r_s, \theta)x + \beta(r_s, \theta)\sqrt{x}}{1 + \gamma(r_s, \theta)x + \delta(r_s, \theta)x^{1.25} + G_{CSR}(q, r_s, \theta)} \right] \quad (6)$$

здесь подгоночные параметры $\alpha, \beta, \gamma, \delta$ и реализация в Python доступны в [28].

Таким образом, определяем ион-ионный потенциал взаимодействия как [29]:

$$\begin{aligned} \Phi(r) &= \int \frac{d^3k}{(2\pi)^3} \frac{4\pi Ze}{k^2} \frac{e^{ikr}}{\varepsilon(k)} = \frac{Ze}{r} + \Phi_{ind}(r) \\ \Phi(r) &= \frac{1}{2\pi^2 r} \int k \cdot |\tilde{\phi}_{ei}(k)|^2 \chi(k) \sin(kr) dk \end{aligned} \quad (7)$$

Выбирая в качестве микропотенциала электрона с ионами в форме

$$\tilde{\phi}_{ei}(k) = \tilde{\phi}_{ei}^C(k) + \delta_{ei}^Q(k), \quad (8)$$

где $\tilde{\phi}_{ei}^C(k)$ – микропотенциала Кулона, $\delta_{ei}^Q(k)$ – разность микропотенциала Кулона и Дойча, $\phi_{ei}^C(k) = \frac{4\pi e^2}{k^2}$ – микропотенциал Кулона, $\phi_{ei}^D(k) = \frac{4\pi e^2}{k^2} \frac{1/\lambda_{ei}^2}{(k^2 + 1/\lambda_{ei}^2)}$ – микропотенциал Дойча [30].

Для определения тепловой длины волны, характеризующей проявление характера электронной волны в парном потенциале взаимодействия λ_{ei} , используем конечное значение при $\lim_{r \rightarrow 0} \phi(r) = \frac{Ze^2}{\lambda_{ei}}$ квантового потенциал Келб-Дойча [31]:

$$\phi_{ei}^{KD}(r) = -\frac{Ze^2}{r} \left\{ 1 - e^{-\frac{r^2}{\lambda_{ei}^2}} + \sqrt{\pi} \frac{r}{\lambda_{ei} \gamma_{ei}} \left(1 - \operatorname{erf} \left[\gamma_{ei} \frac{r}{\lambda_{ei}} \right] \right) \right\}, \quad (9)$$

где γ_{ei} – дополнительный свободный параметр $\gamma_{ei} = \frac{x_1 + x_1^2}{1 + \alpha_{ei}(x_1 + x_1^2)}$.

Таким образом, уравнения (2) – (5) образуют замкнутый набор уравнений, которые можно решить самосогласованно, чтобы получить статическую диэлектрическую функцию.

Оптические свойства плотной водородной плазмы

Оптические характеристики, в частности отражательная способность, напрямую связаны с диэлектрической функцией, которая также определяет проводимость в статическом пределе. Эти параметры служат ключевыми для описания поведения теплого плотного водорода, особенно при высоких давлениях. Важную роль играет отклик среды на внешние возмущения, который в рамках теории линейного отклика описывается диэлектрической проницаемостью $\varepsilon(k, \omega)$.

В данной работе представлены результаты по исследованию оптических свойств водорода, таких как коэффициенты отражения и преломления электромагнитных волн, диэлектрическая проницаемость которого описывается обобщенной моделью Друде-Лоренца:

$$\varepsilon(\omega) = 1 - \frac{\omega_p^2}{\omega[\omega + i\nu(\omega)]}. \quad (10)$$

Используя вещественную и мнимую часть диэлектрической функции, можно получить коэффициент преломления $n(\omega)$ и коэффициент отражения $R(\omega)$ [32–33]:

$$n = \sqrt{\frac{Re\varepsilon(\omega) + \sqrt{(Re\varepsilon(\omega))^2 + (Im\varepsilon(\omega))^2}}{2}}. \quad (11)$$

$$R(\omega) = \left| \frac{1 - \sqrt{\varepsilon(\omega)}}{1 + \sqrt{\varepsilon(\omega)}} \right|^2. \quad (12)$$

В классической модели Друде электропроводимость $\sigma(\omega)$ записывается как

$$\sigma(\omega) = \frac{n e^2}{m} \frac{1}{v(\omega) - i\omega}. \quad (13)$$

В работе использована модель Друде [34] для изучения свойств проводимости теплового плотного водорода для понимания схожести WDM на металл

$$\sigma(\omega) = \frac{\sigma_0}{1 + \omega^2 \tau^2}, \quad (14)$$

где электропроводность $\sigma_0 = \frac{1}{\rho}$ связана с электросопротивлением постоянного тока и временем релаксации $\tau = \frac{m_e}{n_e e^2 \rho}$.

Электрические сопротивления постоянного тока рассчитываются в рамках теории Займана [35], которая описывает рассеяние свободных электронов в веществе ионом. Формула Займана выражает электрическое сопротивление через сдвиги фаз рассеяния электронов и ион-ионный структурный фактор $S(q)$. Ион-ионный структурный фактор предоставляется из молекулярной динамики, а сечения рассеяния – из квантовой теории электрон-ионных столкновений.

Формулировка электрического сопротивления по Займану описывается в рамках теории линейного отклика следующей формулой [35]:

$$\rho = -\frac{\hbar}{3\pi^2 Z^2 n_i} \int_0^\infty \frac{\delta f}{\delta E}(E, \mu) I(E) dE. \quad (15)$$

Распределение Ферми-Дирака и его производная имеют вид:

$$f(E, \mu) = \frac{1}{e^{\beta(E-\mu)} + 1} \quad (16)$$

$$\frac{\delta f}{\delta E}(E, \mu) = -f(E, \mu)(1 - f(E, \mu)), \quad (17)$$

где $n_i = n_e/Z$ – плотность ионов, $E = \frac{\hbar^2 k^2}{2m_e}$, $\beta = \frac{1}{k_B T}$ и μ – химический потенциал.

Для вычисления функции $I(E)$ используем следующую формулу

$$I(E) = \int_0^{2k} q^3 S(q) \sigma(q) dq, \quad (18)$$

где $S(q)$ – ион-ионный структурный фактор; $\sigma(q)$ – транспортное сечение рассеяния.

Для расчета структурного фактора были использованы значения радиальной функции из молекулярной динамики:

$$S(q) = 1 + 4\pi \int [g(r) - 1] \frac{\sin(kr)}{kr} r^2 dr. \quad (19)$$

Исследование характерных собственных мод колебаний ионов в водороде при высоких плотностях

Для анализа спектров продольных и поперечных коллективных возбуждений использовались уравнения, описывающие микроскопические токи в системе. В частности, спектральные плотности $L(k, \omega)$ и $T(k, \omega)$ были получены путем применения временного преобразования Фурье к рассчитанным токам, определяемым по следующим выражениям:

$$L(k, \omega) = \frac{1}{2\pi N t_{meas}} |F_t[\lambda(k, t)]|^2, \quad (20)$$

$$T(k, \omega) = \frac{1}{2\pi N t_{meas}} |F_t[\tau(k, t)]|^2, \quad (21)$$

где t_{meas} – время измерения, F_t – временное преобразование Фурье, $\lambda(k, t)$ и $\tau(k, t)$ – продольная и поперечная компоненты микроскопического тока соответственно.

Последний определяется как:

$$j(k, t) = \sum_{j=1}^N v_j(t) \exp(ikr_j(t)), \quad (22)$$

причем компоненты $\lambda(k, t)$ и $\tau(k, t)$ задаются выражениями:

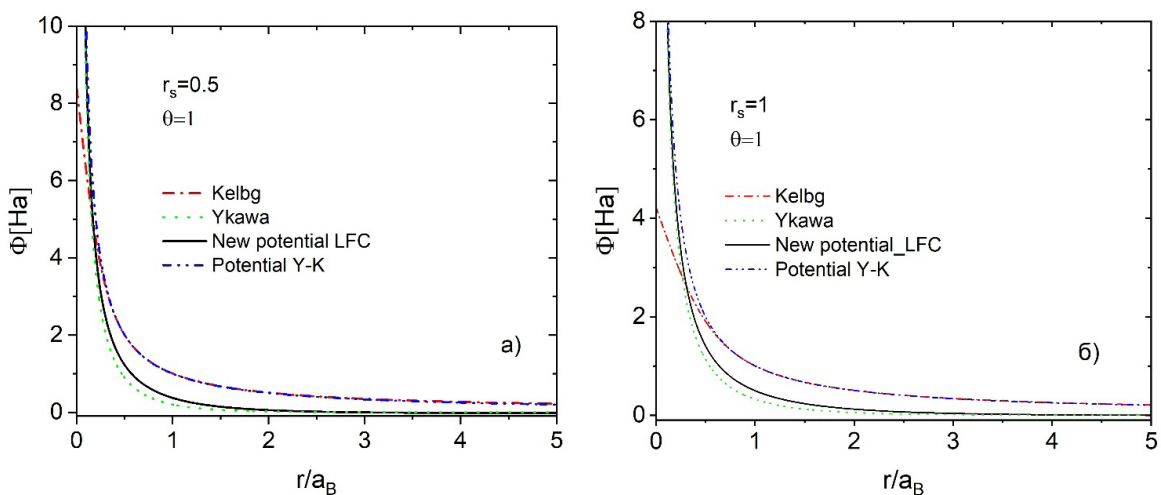
$$\lambda(k, t) = \sum_{j=1}^N v_{jx}(t) \exp(ikx_j(t)), \quad (23)$$

$$\tau(k, t) = \sum_{j=1}^N v_{jy}(t) \exp(iky_j(t)), \quad (24)$$

где в уравнениях (23) и (24) компоненты векторов скорости $v_j(t)$ и положения $r_j(t)$ частицы j в момент времени t получены с помощью МД-моделирования, а k – волновое число, соответствующее волновому вектору, принятому параллельным оси x . Продольная и поперечная скорости звука s_L и s_T определялись в длинноволновом режиме ($k \rightarrow 0$) по соотношению: режим ($k \rightarrow 0$) по уравнению $\lim_{k \rightarrow 0} \frac{\omega}{k} = s_{L/T}$, где s_L и s_T обозначают продольную и поперечную скорость звука соответственно.

Результаты и обсуждение

На рисунке 1 (а-г) и 1.4 (а-г) представлены результаты для экранированного потенциала, вычисленного с использованием уравнения (9). В частности, потенциал на основе LFC сравнивается с потенциалом, вычисленным без учета эффектов обменно-корреляционной связи на отклике плотности (Potential Y-K) и с потенциалами Келбг-Дойча (Kelbg) и Юкавы (Ykawa) для различных значений параметров плазмы. Экранированные потенциалы вычисляются для $Z = 1$, а значения потенциалов нормированы на $[Na]$, а также расстояния на a_B . Из рисунков видно, что эффективный потенциал взаимодействия ион-ион с использованием поправки локального поля также экранируется на больших расстояниях, но, как следствие, из-за волновой природы электронов экранирование немного слабее, чем у потенциала Юкавы.



Продолжение рисунка 1.

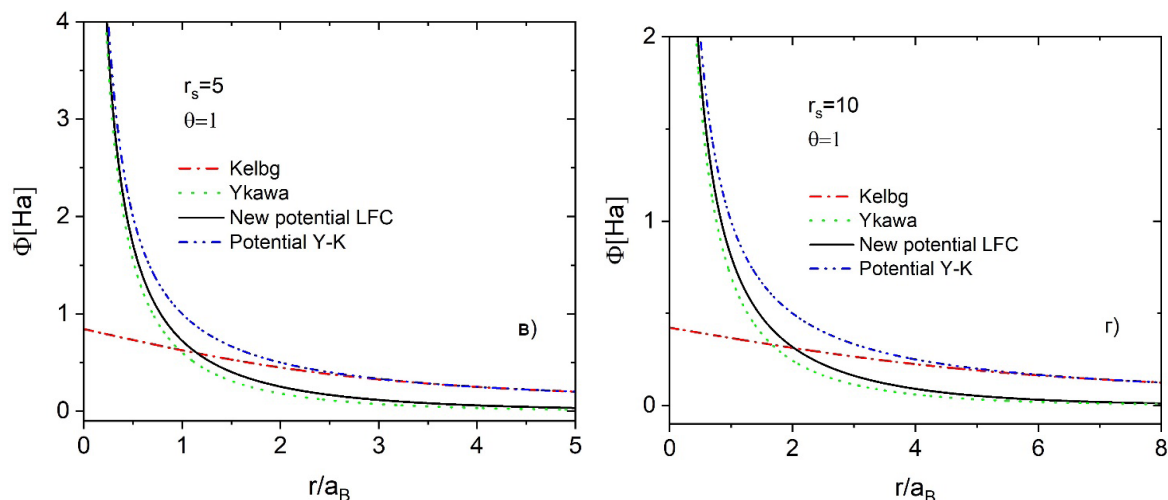


Рисунок 1 – Потенциал ион-ион взаимодействия при параметре вырождения $\theta = 1$ и разных значениях плотности r_s (а-г)

На рисунках 2–3 приведены значения коэффициента отражения при плотности 10^{24} cm^{-3} для различных значений параметра вырождения θ . Как видно из рисунка 2, увеличение параметра вырождения при постоянной плотности приводит к увеличению значения коэффициента отражения. Установлено, с увеличением температуры коэффициент отражения увеличивается при малых значениях длина волны излучения. Таким образом, при постоянной плотности увеличение температуры приводит к увеличению значения коэффициента отражения.

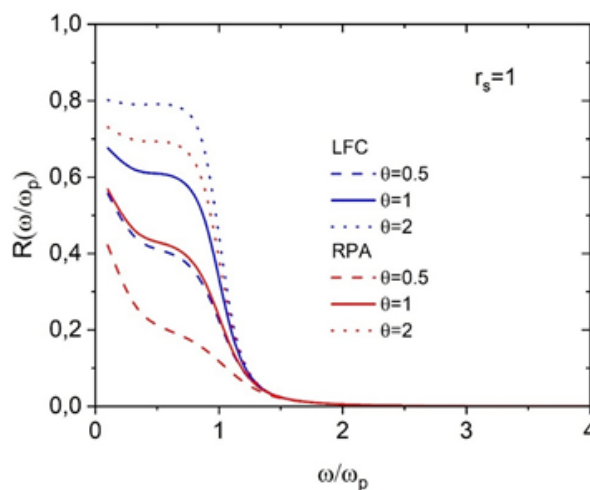


Рисунок 2 – Коэффициент отражения плазмы при $r_s = 1$ для различных параметров вырождения θ

На рисунке 4–5 приведены графики для коэффициента преломления на основе потенциала (1) и структурного фактора при различных значениях r_s, θ . Как видно из рисунков 4–5, с уменьшением параметра вырождения коэффициент преломления увеличивается. Когда длина волны излучения существенно превышает характерную плазменную длину, коэффициент преломления не меняется. Также из рисунков 4–5 можно видеть, что рассчитанные значения

коэффициента преломления имеют корректную асимптотику $n(\omega) = \sqrt{1 - \frac{\epsilon_2}{\epsilon_1^2}}$ и стремятся к единице.

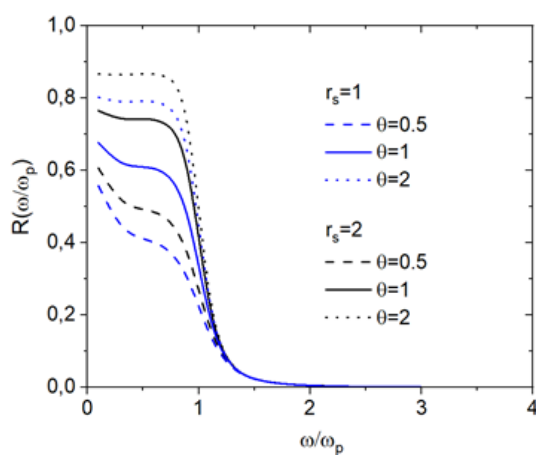


Рисунок 3 – Коэффициент отражения плазмы на основе псевдопотенциала (1) при $r_s = 1$ и $r_s = 2$ для различных параметров вырождения θ

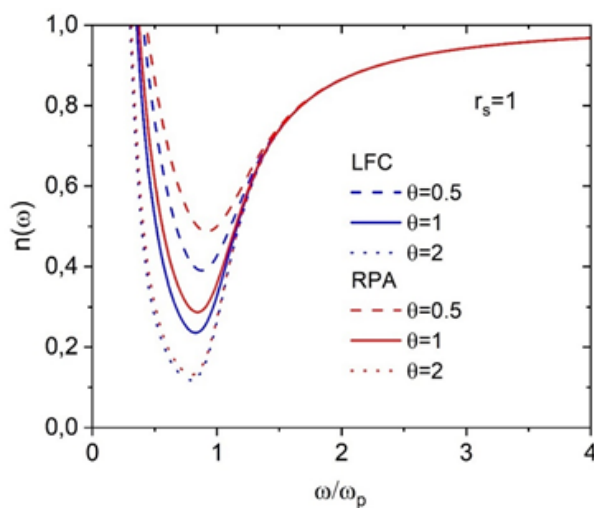


Рисунок 4 – Коэффициент преломления плазмы на основе псевдопотенциала (1) при $r_s = 1$ для различных параметров вырождения θ

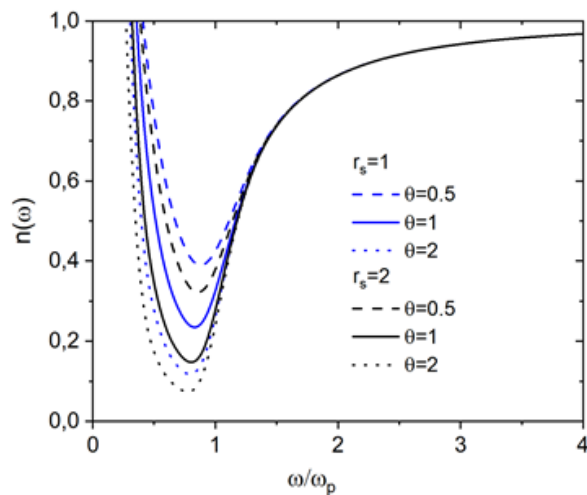


Рисунок 5 – Коэффициент преломления плазмы на основе псевдопотенциала (1) при $r_s = 1$ и $r_s = 2$ для различных параметров вырождения θ

На рисунке 6 представлен коэффициент поглощения $\alpha(\omega)$ плотной электрон-ионной плазмы при $r_s = 1$ в зависимости от нормированной частоты ω/ω_p . Сравниваются расчеты в приближении случайных фаз (RPA, красные линии) и с учетом локальных полей (LFC, синие линии) для разных значений параметра вырождения. Учет обмен-корреляционных эффектов через поправку локальных полей приводит к заметному росту поглощения в окрестности плазменной частоты и лучше отражает усиление межчастичных корреляций в режиме $r_s = 1$, тогда как повышение θ подавляет поглощение, что согласуется с ослаблением квантовой вырожденности.

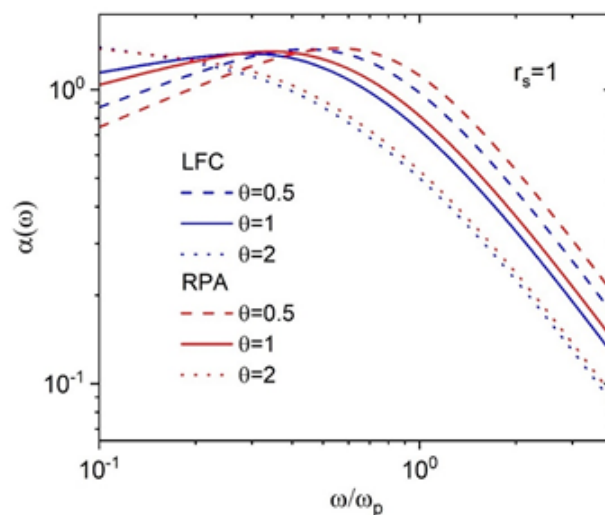


Рисунок 6 – Коэффициент поглощения плазмы при $r_s = 1$ для различных параметров вырождения θ

На рисунке 7 представлен графический анализ отражательной способности плазмы электромагнитного излучения с длинами волн 532 нм и 808 нм соответственно. Как видно из рисунка, значения коэффициента отражения и экспериментальные данные [36] качественно хорошо коррелируют между собой и наблюдается хорошее согласие при меньших значениях длины волны падающего излучения.

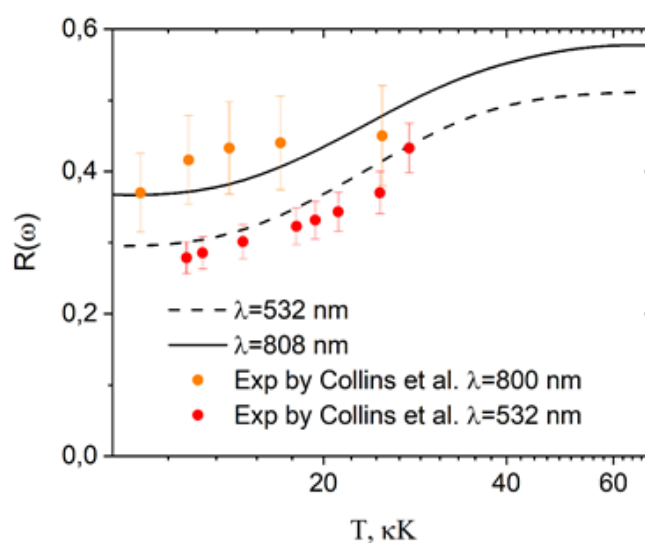


Рисунок 7 – Коэффициент отражения в зависимости от температуры плазмы для различной длины волны излучения

Микроскопические токи были рассчитаны на основе уравнений (23) и (24), после чего к ним было применено преобразование Фурье по времени с целью получения спектральных характеристик продольных $L(k, \omega)$ и поперечных $T(k, \omega)$ флуктуаций тока. Для повышения статистической надежности и уменьшения влияния шумов результаты усреднялись по десяти независимым моделированиям, проведенным при идентичных параметрах системы. Результаты соответствующих дисперсионных соотношений (т.е. зависимость частоты коллективного колебания волны ω от волнового числа k) приведены на рисунках 8 и 9. Для более детального анализа изменений в дисперсионных характеристиках, обусловленных электронными обменно-корреляционными эффектами, из спектров продольных коллективных возбуждений были извлечены положения пиков в зависимости от частоты ω и волнового числа k (см. рисунок 10). Из рисунка 10 можно заметить небольшое увеличение дисперсионной кривой. В результате этого сдвига дисперсионных кривых $L(k, \omega)$ происходит изменение значений скоростей звука. При малых параметрах связи Γ кинетическая энергия частиц сравнима с потенциальной, и частоты продольных коллективных возбуждений ω/ω_p растут почти линейно. С увеличением параметра связи $\Gamma > 10$ доминируют силы отталкивания, и разница в частотах уменьшается. Было выявлено, что с увеличением параметра связи скорость звука уменьшается на 20% при $\Gamma = 10.317$.

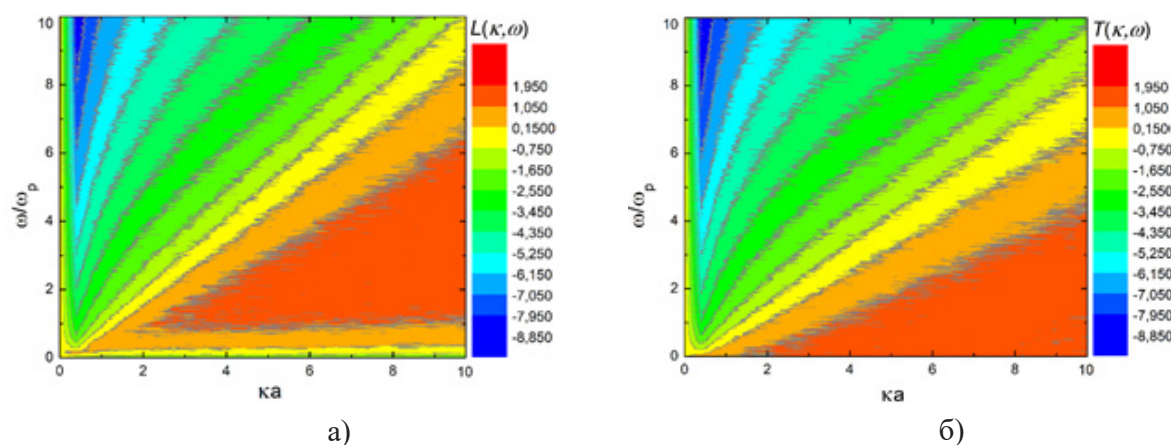


Рисунок 8 – Продольные (а) и поперечные (б) дисперсионные соотношения

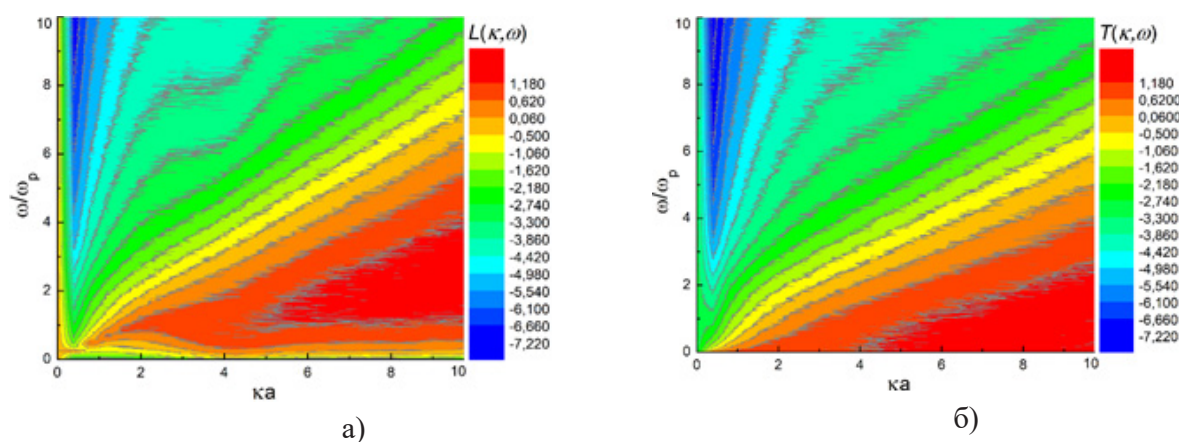


Рисунок 9 – Продольные (а) и поперечные (б) дисперсионные соотношения

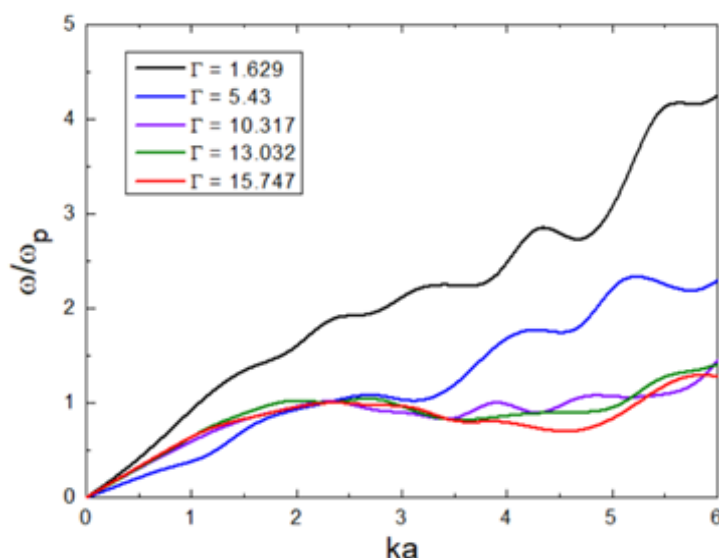


Рисунок 10 – Пики дисперсионного соотношения при различных значениях параметра связи

Заключение

Таким образом было проведено исследование оптических свойств плотной водородной плазмы при высоких температурах на основе эффективного потенциала взаимодействия с учетом коррекции локального поля, полученного из моделирования методом квантового Монте-Карло. Оптическая отражательная способность разогретого плотного водорода (WDM) была сопоставлена с экспериментальными данными для длин волн 532 нм и 808 нм. Полученное хорошее совпадение между рассчитанными значениями коэффициента отражения электромагнитного излучения от плазменного слоя и экспериментальными результатами подтверждает корректность примененного в данной работе математического подхода и адекватность используемой физической модели.

Были проведены исследования продольных и поперечных спектров колебаний микроскопического ионного тока в водороде при различных температурах и плотностях, а также проанализировано влияние электронных обменных и корреляционных эффектов. Для изучения спектров продольных и поперечных коллективных возбуждений использовались уравнения, описывающие микроскопическую динамику токов в системе. В результате были получены спектры продольных и поперечных флуктуаций ионного тока водорода при разных значениях параметра связи. Дополнительно рассчитаны и исследованы продольные и поперечные скорости звука в системе. Установлено, что частоты продольных коллективных возбуждений возрастают почти линейно, что связано с тем, что при малых значениях параметра связи кинетическая энергия частиц становится сопоставимой с потенциальной.

В целом влияние обменно-корреляционных эффектов на свойства плотной плазмы связано с поведением электронов в плотных и высокотемпературных системах. Учет точных моделей локального поля приводит к существенным изменениям в расчетах коэффициентов переноса, что критически важно для моделирования горячего плотного вещества, металлизированной плазмы и вырожденных электронных систем. Кроме того, это имеет фундаментальное значение для теоретического объяснения наблюдаемых спектров излучения, а также для интерпретации экспериментов по сжатию вещества мощными лазерными импульсами и ударами.

Информация о финансировании

Работа выполнена при поддержке Министерства науки и высшего образования Республики Казахстан в рамках гранта AP19678033.

ЛИТЕРАТУРА

- 1 Lindl, J. Development of the indirect- drive approach to inertial confinement fusion and the target physics basis for ignition and gain. *Phys. Plasmas*, 2, 3933–4024 (1995). <https://doi.org/10.1063/1.871025>.
- 2 Haan, S. W. et al. Design and modeling of ignition targets for the National Ignition Facility. *Phys. Plasmas*, 2, 2480–2487 (1995). <https://doi.org/10.1063/1.871209>.
- 3 Magro, W.R., Ceperley, D.M., Pierleoni, C., and Bernu, B. Molecular Dissociation in Hot, Dense Hydrogen. *Phys. Rev. Lett.*, 76, 1240–1243 (1996). <https://doi.org/10.1103/PhysRevLett.76.1240>.
- 4 Weir, S.T., Mitchell, A.C., and Nellis, W.J. Metallization of Fluid Molecular Hydrogen at 140 GPa (1.4 Mbar). *Phys. Rev. Lett.*, 76, 1860–1863 (1996). <https://doi.org/10.1103/PhysRevLett.76.1860>.
- 5 Celliers, P.M. et al. Shock-Induced Transformation of Liquid Deuterium into a Metallic Fluid. *Phys. Rev. Lett.*, 84, 5564 (2000). <https://doi.org/10.1103/PhysRevLett.84.5564>.
- 6 Hu, S. X., Militzer, B., Goncharov, V.N., Skupsky, S. First-principles equation-of-state table of deuterium for inertial confinement fusion applications. *Phys. Rev. B.*, 84, 224109 (2011). <https://doi.org/10.1103/PhysRevB.84.224109>.
- 7 Loubeyre, P., Brygoo, S., et. al. Extended data set for the equation of state of warm dense hydrogen isotopes. *Phys. Rev. B.*, 86, 144115 (2012). <https://doi.org/10.1103/PhysRevB.86.144115>.
- 8 Holst, B., Redmer, R., Desjarlais, M.P. Thermophysical properties of warm dense hydrogen using quantum molecular dynamics simulations. *Phys. Rev. B.*, 77, 184201 (2008). <https://doi.org/10.1103/PhysRevB.77.184201>.
- 9 Lambert, F., Recoules, V., Decoster, A., Cléroutin, J., Desjarlais, M. On the transport coefficients of hydrogen in the inertial confinement fusion regime. *Phys. Plasmas*, 18, 056306 (2011). <https://doi.org/10.1063/1.3574902>.
- 10 Li, D., Ping, Zh., Jun, Y. Quantum molecular dynamics simulations of the thermophysical properties of shocked liquid ammonia for pressures up to 1.3 TPa. *The Journal of Chemical Physics*, 139, 134505 (2013). <https://doi.org/10.1063/1.4823744>.
- 11 Dharma-wardana M.W.C. Electron-ion and ion-ion potentials for modeling warm dense matter: Applications to laser-heated or shock-compressed Al and Si. *Physical Review E*, 86, 036407 (2012). <https://doi.org/10.1103/PhysRevE.86.036407>.
- 12 Hu, S.X., Colins, L.S., et.al. First-principles opacity table of warm dense deuterium for inertial-confinement-fusion applications. *Physical Review E*, 102, 053209 (2014). <https://doi.org/10.1103/PhysRevE.90.033111>.
- 13 Moldabekov, Zh. A., Groth, S., Dornheim, T., Bonitz, M., Ramazanov, T.S. Ion potential in non-ideal dense quantum plasmas. *Contrib. Plasma Phys.*, 57(10), 532 (2017). <https://doi.org/10.1002/ctpp.201700109>.
- 14 Moldabekov, Z., Schoof, T., Ludwig, P., Bonitz, M., Ramazanov, T. Statically screened ion potential and Bohm potential in a quantum plasma. *Phys. Plasmas*, 22(10), 102104 (2015). <https://doi.org/10.1063/1.4932051>.
- 15 Moldabekov, Zh. A., Ludwig, P., Joost, J.-P., Bonitz, M., Ramazanov, T.S. Dynamical Screening and Wake Effects in Classical, Quantum, and Ultrarelativistic Plasmas. *Contrib. Plasma Phys.*, 55(2–3), 186 (2015). <https://doi.org/10.1002/ctpp.201400105>.
- 16 Moldabekov, Zh.A., Ludwig, P., Bonitz, M., Ramazanov, T.S. Theoretical foundations of quantum hydrodynamics for plasmas. *Contrib. Plasma Phys.*, 56(5), 442 (2016). <https://doi.org/10.1063/1.5003910>.
- 17 Moldabekov, Zh.A., Amirov, S.M., Ludwig, P., Bonitz, M., Ramazanov, T.S. Effect of the dynamical collision frequency on quantum wakefields. *Contrib. Plasma Phys.*, 59(4–5), e201800161 (2019). <https://doi.org/10.1002/ctpp.201800161>.
- 18 Bowen, C., Sugiyama, G. and Alder, B.J. Static dielectric response of the electron gas, *Phys. Rev.*, 50, 14838 (1994). <https://doi.org/10.1103/PhysRevB.50.14838>.
- 19 Moroni, S., Ceperley, D.M. and Senatore, G. Static response from quantum Monte Carlo calculations, *Phys. Rev.*, 69, 1837 (1992). <https://doi.org/10.1103/PhysRevLett.69.1837>.
- 20 Perdew, J.P. and Wang, Y. Pair-distribution function and its coupling-constant average for the spin-polarized electron gas, *Phys. Rev.*, 46, 12947 (1992). <https://doi.org/10.1103/PhysRevB.46.12947>.

- 21 Ramazanov, T.S., Moldabekov, Zh.A., Gabdullin, M.T. Effective potentials of interactions and thermodynamic properties of a nonideal two-temperature dense plasma. *Phys. Rev.*, 92, 023104 (2015). <https://doi.org/10.1103/PhysRevE.92.023104>.
- 22 Ramazanov, T.S., Moldabekov, Zh.A., Gabdullin, M.T. Multipole expansion in plasmas: Effective interaction potentials between compound particles. *Phys. Rev.*, 93, 053204 (2016). <https://doi.org/10.1103/PhysRevE.93.053204>.
- 23 Moldabekov, Zh.A., Groth, S., Dornheim, T., Kählert, H., Bonitz, M., Ramazanov, T.S. Structural characteristics of strongly coupled ions in a dense quantum plasma. *Phys. Rev.*, 98, 023207 (2018). <https://doi.org/10.1103/PhysRevE.98.023207>.
- 24 Moldabekov, Z.A., Dornheim, T., Bonitz, M. Screening of a test charge in a free-electron gas at warm dense matter and dense non-ideal plasma conditions. *Contrib. Plasma Phys.*, e202000176 (2021). <https://doi.org/10.1002/ctpp.202000176>.
- 25 Arista, N.R. and Brandt W. Dielectric response of quantum plasmas in thermal equilibrium. *Phys. Rev. A.*, 29, 1471–1480 (1984). <https://doi.org/10.1103/PhysRevA.29.1471>.
- 26 Dornheim, T., Vorberger, J., Groth, S., Hoffmann, N., Moldabekov, Zh.A., Bonitz, M. The static local field correction of the warm dense electron gas: An ab initio path integral Monte Carlo study and machine learning representation. *The Journal of Chemical Physics.*, 151, P. 194104 (2019). <https://doi.org/10.1063/1.5123013>.
- 27 Dornheim, T., Moldabekov, Zh.A., and Tolias, P. Analytical representation of the local field correction of the uniform electron gas within the effective static approximation. *Phys. Rev. B.*, 103, 165102 (2021). <https://doi.org/10.1103/PhysRevB.103.165102>.
- 28 URL: <https://github.com/ToDor90/LFC>.
- 29 Moldabekov, Zh. A., and Dornheim, T., Bonitz, M. Screening of a test charge in a free-electron gas at warm dense matter and dense non-ideal plasma conditions. *Contributions to Plasma Physics*, 62, e202000176 (2022). <https://doi.org/10.1002/ctpp.202000176>.
- 30 Filinov, A.V., Bonitz, M., Ebeling, W.O. Improved Kelbg potential for correlated Coulomb systems. *Journal of Physics A: Mathematical and Theoretical*, 36(22), 5957 (2003). <https://doi.org/10.1088/0305-4470/36/22/317>.
- 31 Nadine Wetta N., Pain, J. Consistent approach for electrical resistivity within Ziman's theory from solid state to hot dense plasma: Application to aluminum. *Physical Review E*, 90, 033111 (2020). <https://doi.org/10.1103/PhysRevE.102.053209>.
- 32 Lorazo, P., Lewis, L.J., Meunier, M. Short-Pulse Laser Ablation of Solids: From Phase Explosion to Fragmentation. *Phys.Rev.Lett.*, 91, 225502 (2003). <https://doi.org/10.1007/s003390000686>.
- 33 Faussurier, G., Blancard, C. Resistivity saturation in warm dense matter. *Phys Rev E. Stat Nonlin Soft Matter Phys.*, 91(1), 013105 (2015). <https://doi.org/10.1103/PhysRevE.91.013105>.
- 34 Pain, J.C., Dejongh, G. Electrical Resistivity in Warm Dense Plasmas Beyond the Average-Atom Model. *Contrib. Plasma Phys.*, 50, 39–45 (2010). <https://doi.org/10.1002/ctpp.201010010>.
- 35 Ziman, J.M. A theory of the electrical properties of liquid metals. I: The monovalent metals. *The Philosophical Magazine: A Journal of Theoretical Experimental and Applied Physics*, 6(68), 1013–1034 (1961). <https://doi.org/10.1080/14786436108243361>.
- 36 Collins, L.A., Bickham, S.R., Kress, J.D., Mazevet, S., Lenosky, T.J., Troullier, N.J., and Windl W. Dynamical and optical properties of warm dense hydrogen. *Phys. Rev. B*, 63, 184110 (2001). <https://doi.org/10.1103/PhysRevB.63.184110>.

¹Коданова С.К.,

ф.-м.ғ.к., профессор, ORCID ID: 0000-0001-7098-471X,

e-mail: kodanova@physics.kz

^{1*}Исанова М.К.,

PhD, қауымдастырылған профессор, ORCID ID: 0000-0002-0264-2694,

*e-mail: issanova@physics.kz

¹Бастыкова Н.Х.,

PhD, қауымдастырылған профессор, ORCID ID: 0000-0002-8400-7183,

e-mail: bastykova_nuriya@physics.kz

¹Кенжебекова А.И.,

PhD, аға оқытушы, ORCID ID: 0000-0003-0223-0216,

e-mail: kenzhebekova.a@kaznu.kz

²Хоффманн Д.,

ғылым докторы, профессор, ORCID ID: 0000-0003-0922-7802,

e-mail: hoffmann@physik.tu-darmstadt.de

¹Әл-Фараби атындағы Қазақ ұлттық университеті, Алматы қ., Қазақстан

²Дармштадт техникалық университеті, Дармштадт қ., Германия

ЖОҒАРЫ ТЫҒЫЗДЫҚТАҒЫ СУТЕГІ ПЛАЗМАСЫНЫҢ ОПТИКАЛЫҚ ЖӘНЕ ДИНАМИКАЛЫҚ ҚАСИЕТТЕРІН ЗЕРТТЕУ

Аңдатпа

Бұл жұмыста сутегінің оптикалық қасиеттерін, атап айтқанда электромагниттік толқындардың шағылысу және сыну коэффициенттерін, жалпыланған Друде–Лоренц үлгісімен сипатталатын заттың өткізгіштігі арқылы зерттеу нәтижелері ұсынылған. Зерттеуде сутегінің әртүрлі температуралар мен тығыздықтағы микроскопиялық иондық ток тербелістерінің бойлық және көлденең спектрлері қарастырылып, электрон алмасу мен корреляциялық әсерлердің ықпалы талданды. Есептеу кванттық Монте-Карло модельдеуі негізінде алынған локальді өріс түзетулерін ескеретін эффективті өзара әрекеттесу потенциалын қолдану арқылы жүргізілді. Кванттық статистикалық есептеулерге негізделген дәл жуықтауларды қамтитын локальді өріс модельдерін қолдану тығыз электрондық жүйелердің көліктік және оптикалық қасиеттерін сенімді сипаттауға мүмкіндік береді. Атап айтқанда, локальді өріс функциясын ескеру оптикалық және динамикалық қасиеттерді есептеуде айтарлықтай түзетулер енгізеді, бұл ыстық тығыз затты, металданған плазманы және азғындалған электрондық жүйелерді модельдеу үшін аса маңызды. Сондай-ақ локальді өріс функциясының нақты формасының болуы плазманың оптикалық және динамикалық қасиеттерін, оның ішінде шағылу және жұтылу коэффициенттерін, сондай-ақ өзіндік тербеліс режимдерін дәл сипаттауға мүмкіндік береді.

Тірек сөздер: тығыз плазма, эффективті потенциал, оптикалық қасиеттер, шағылу коэффициенті, жұтылу коэффициенті, өзіндік тербеліс режимдері.

¹Kodanova S.,Cand. Phys.-Math. Sc., Professor, ORCID ID: 0000-0001-7098-471X,
e-mail: kodanova@physics.kz**^{1*}Issanova M.,**PhD, Associate Professor, ORCID ID: 0000-0002-0264-2694,
*e-mail: issanova@physics.kz**¹Bastykova N.,**PhD, Associate Professor, ORCID ID: 0000-0002-8400-7183,
e-mail: bastykova_nuriya@physics.kz**¹Kenzhebekova A.,**PhD, Senior Lecturer, ORCID ID: 0000-0003-0223-0216,
e-mail: kenzhebekova.a@kaznu.kz**²Hoffmann D.,**D.Sc., Professor, ORCID ID: 0000-0003-0922-7802,
e-mail: hoffmann@physik.tu-darmstadt.de¹Al-Farabi Kazakh National University, Almaty, Kazakhstan²Technical University of Darmstadt, Darmstadt, Germany**STUDY OF OPTICAL AND DYNAMIC PROPERTIES
OF HYDROGEN PLASMA AT HIGH DENSITIES****Abstract**

This paper presents the results of a study of the optical properties of hydrogen, specifically the reflection and refraction coefficients of electromagnetic waves, with the permittivity of the substance described by the generalized Drude–Lorentz model. This paper examines the longitudinal and transverse spectra of microscopic ion current oscillations in hydrogen at various temperatures and densities, and analyzes the influence of electron exchange and correlation effects. The study was conducted using the effective interaction potential, taking into account a local field correction obtained from quantum Monte Carlo simulations. The use of accurate local field models, such as approximations based on quantum statistical calculations, allows for the reliable reproduction of the transport and optical properties of dense electron systems. In particular, taking into account the local field function leads to significant corrections in the calculations of optical and dynamic properties, which is critical for modeling hot dense matter, metallized plasma, and degenerate electron systems. In addition, the presence of an exact form of the local field function allows us to correctly describe the optical and dynamic properties of the plasma, including the reflection and absorption coefficients and natural oscillation modes.

Keywords: dense plasma, effective potential, optical properties, reflection coefficient, absorption coefficient, natural oscillation modes.

Дата поступления статьи в редакцию: 17.11.2025

UDC 620.93
IRSTI 29.31.15

<https://doi.org/10.55452/1998-6688-2025-22-4-401-410>

¹**Zhukeshov A.M.,**

Dr. Phys.-Math. Sc., Professor, ORCID ID: 0000-0001-9153-323X,

e-mail: zhukeshov@physics.kz

^{1*}**Abdybai U.B.,**

PhD student, ORCID ID: 0009-0002-5309-5827,

*e-mail: abdibay_ulan@mail.ru

¹**Salimov E.E.,**

PhD student, ORCID ID: 0009-0005-7712-6263,

e-mail: Salimov_ernar@mail.ru

²**Carlos M.,**

Professor, ORCID ID: 0000-0002-7374-505X,

e-mail: Carlos.meza@ieee.org

¹Al-Farabi Kazakh National University, Almaty, Kazakhstan

²The university of applied Sciences of Anhalt, Anhalt, Germany

COMPARATIVE PERFORMANCE CHARACTERISTICS OF METAL-AIR BATTERIES WITH CAST AND POROUS ELECTRODES

Abstract

Currently, the development of electrochemical energy storage systems plays a key role in meeting the growing demand for electric power. Metal-air batteries (MAB) with high specific energy capacity are considered as promising solutions for use in power plants as backup power sources. One of the main limitations of their widespread implementation is the need to improve the efficiency of anode materials. In this paper, we propose the use of porous aluminum electrodes to improve the performance of MAB. Two types of porous anodes manufactured using different technologies were studied. For the powder aluminum anode, the current density was 20–30 mA/cm², which is comparable to the performance of a monolithic (standard) anode. At the same time, foam aluminum demonstrated higher current density values of 52–64 mA/cm². An additional advantage of porous anodes is their reduced weight (by 10–30%), which helps to improve the weight and size characteristics of MAB and opens up opportunities for creating more efficient energy systems.

Keywords: metal–air battery, aluminum anode, porous electrode, aluminum foam, electrochemical performance, specific surface area, energy storage.

Introduction

At the current stage of energy development, the problem of energy storage during periods of sudden decrease in its generation is of particular relevance, which is necessary to ensure a stable quality of power supply and meet increasing demand [1–3]. Among various energy storage technologies, electrochemical systems demonstrate the best compliance with the key requirements of scalable solutions, such as a long service life, high efficiency, and the ability to adapt to different consumption levels. Of particular interest among electrochemical storage devices are metal-air batteries (MAB), which are distinguished by a significantly higher theoretical energy capacity compared to traditional battery systems [4]. However, their practical application is limited by a number of technical problems associated with the design of the metal anode, air cathode, and electrolyte stability [5]. Considering that MAB have the highest potential energy capacity among known types of energy sources, environmentally friendly and affordable metals, such as iron, zinc, and especially aluminum,

are considered promising materials for electrodes. Aluminum, in particular, has a high degree of recyclability, significant energy density (8.1 kWh/kg) and electrochemical potential (1.7 V), making it commercially attractive for use in MAB.

The study of metal electrodes in metal-air batteries (MAB) faces a number of technical challenges, including dendrite formation, electrode deformation, corrosion processes, side reactions, and surface passivation [6]. Currently, active attention is paid to anode materials based on zinc, aluminum, and iron. For example, Parker et al. proposed a three-dimensional zinc sponge-type structure with interconnected porosity and high conductivity, which made it possible to effectively suppress the formation of large-scale dendrites [7]. Another study presented a Zn anode with a developed specific surface area obtained by bubble synthesis [8]. It was found that reducing the pore size contributes to an increase in the efficiency of zinc-air batteries. Aluminum anodes are also the subject of intensive research. A series of studies by Wen and colleagues [9–11] showed that alloying aluminum with Mg, Ga, Sn, and Mn metals increases its electrochemical activity compared to pure Al. The study describes a simple method for producing a porous aluminum anode by applying Al powder onto aluminum foil [12]. The resulting powder electrode ensured a decrease in local current density and contributed to uniform deposition of aluminum. It should be noted that there are significantly fewer publications devoted to porous anodes for MAW compared to studies of porous cathodes. The latter are actively being developed using nanomaterials such as nanotubes, graphene, and others [13–16]. However, cathode catalysis is not considered in this paper.

Metal-air batteries (MABs) find applications in both automotive and portable electronics, demonstrating potential for widespread implementation in various technological areas [17–19]. In particular, aluminum-air systems are considered as promising solutions for highly efficient energy storage. However, despite active research, a number of fundamental scientific and engineering problems that limit the commercialization and large-scale implementation of MABs remain unresolved [20–23].

Thus, the key tasks in the development of new electrode materials are to minimize corrosion processes, suppress dendrite formation and improve electrochemical characteristics, including specific capacity. One of the effective approaches to improving the anode properties is to increase the specific surface area of the electrode. Expanding the contact zone between the electrode and the electrolyte promotes an increase in the number of active sites, which in turn reduces the local current density and stabilizes electrochemical processes. Porous materials, having a reduced density, allow reducing the mass of the electrode while maintaining its geometric parameters, which can be useful for optimizing energy characteristics. However, questions related to the behavior of exothermic reactions in such structures, possible local heating of the anode body and the magnitude of the total current remain open. The study of these aspects and the assessment of the practical efficiency of porous anodes under real operating conditions of metal-air batteries is the main goal of this study.

There are several ways to create porous materials. The standard method is to obtain metal foam in a gas flow with various manipulations. Depending on the pore structure, namely open or closed, the metal foam can be multifunctional, which makes it suitable for highly specialized applications [24]. The analysis conducted showed that the porosity of aluminum foam can be in the range from 55 to 70%, and the maximum energy absorption capacity is achieved in a metal foam with a porosity of 57% [25]. Recently, studies have been conducted on the preparation of porous intermetallic compounds containing Al using the combustion synthesis (CS) reaction [26]. The reaction mechanism for the M–Al system (M = Ti, Ni, Fe, Nb, and Cu) is closely related to the state of aluminum (e.g., solid or liquid phase) and largely depends on the processing parameters such as particle size, Al content, heating rate, pressing pressure, sintering temperature, content and type of pore-forming agents. It should also be noted that porous materials are successfully used not only as anodes, but also as cathodes in metal-air batteries. 3D catalyst film made of copper and carbon with a double 3D porous structure was manufactured [27]. The finished composites were used as an air cathode in an aluminum-air battery, which demonstrated good water resistance and air permeability.

Materials and methods

In this work, two types of porous materials were used. The surface appearance of the anodes of the first and second types is shown in Figure 1. A cast electrode (hereinafter referred to as a monolith) of the same size, cut from an A0 aluminum sheet, was used as a comparison (Figure 2). All samples had the same area (20x30x4) mm³.

The first type of samples was obtained by self-fabrication from metal powders. Pressed powder materials used in this process also have a porous structure, which makes them promising for use as anodes. Chemically pure aluminum powder with a particle size of about 30 μm was used to form porous aluminum anodes. The process included the following main stages: preliminary mixing of aluminum powder with 10% copper powder, loading the resulting mixture into a collapsible press mold and subsequent pressing at a force of 12 tons until a dense blank was obtained. Then thermal annealing was carried out in a vacuum electric furnace at a temperature corresponding to the formation of an equilibrium eutectic (about 0.6 of the melting point), followed by cooling until it solidified. As a rule, pressure in the range of 250–400 MPa is used to press aluminum powders. If it is necessary to calculate the force required to press a platform with a given area S at a pressure of 2500–4000 kgf/cm², the following formula is used:

$$F = P \times S, \quad (1)$$

Thus, to press disks with a diameter of 2–5 cm at a pressure of 2500–4000 kgf/cm², a force of ~8 to ~79 thousand kg/f is required, which corresponds to approximately 8–79 tons. In the experiment, a force of about 12 tons was used, which corresponds to pressing a disk with a diameter of about 2.5–3 cm at medium pressure.

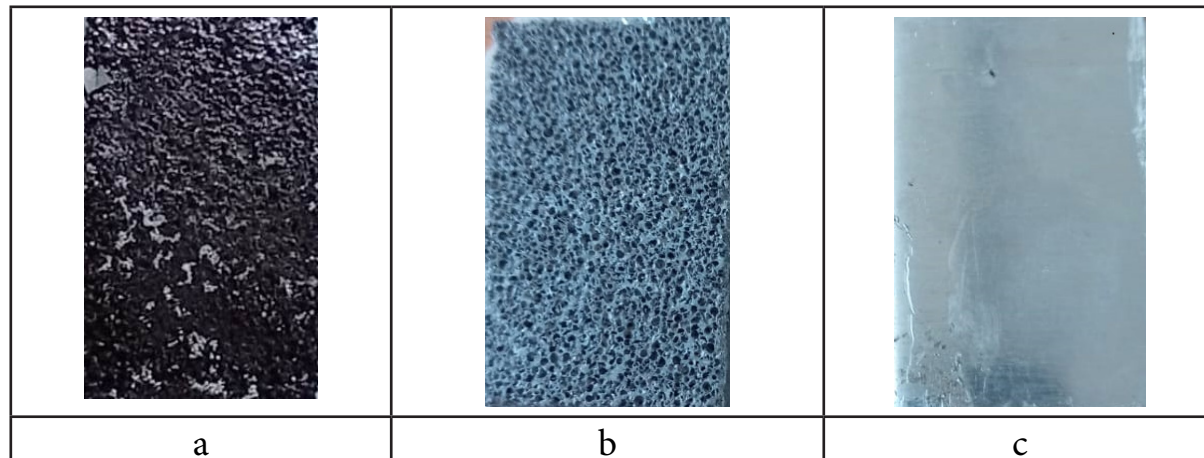


Figure 1 – Samples of porous anodes of type 1 (a) and type 2 (b) and Cast aluminum anode (c). Sample dimensions 20x30 mm.

Samples of porous anodes of the second type are aluminum foam (hereinafter referred to as foam) obtained by industrial gas spraying (manufacturer - Kunshan New Materials Corp., China). Foam density characteristics are given in Table 1. The key parameters determining the physical and mechanical properties of the foam are its density and porosity, i.e. the volume of pore space. Porosity is calculated using the following formula:

$$\Pi = \left(1 - \frac{\rho_0}{\rho}\right) 100, \quad (2)$$

where Π is the porosity in %, ρ_0 is the density of the finished (sintered) material, ρ is the density of the cast material. The density of the foam metal (g/cm^3) was determined in a standard way using the formula: $\rho_0 = m/V$ where m is the mass of the sample, V is the volume of the sample. Traditionally, in the production of foam materials, the main characteristic of porosity is designated PPI - the average number of Pores Per Inch. The data for our material are as follows: $\text{PPI} = 45$, the average pore size is 0.5 mm. Figure 1 b shows the foam surface at an optical magnification of 10X, from which it is clear that the pore sizes are indeed about 500 μm . The free surface area is determined by the pore size and the bulk density of the foam material [28]. The nature of the dependence of these parameters is linear and is presented below by the formula

$$S = k \times (\text{PPI}) \left[\frac{m^2}{m^3} \right]. \quad (3)$$

For the given PPI value, the specific area is in the range of 2500–3500 m^2/m^3 . Then, according to formula (1), for our foam material, the effective surface area will be $S_{\text{ef}} = (32000 \pm 3000) \text{ cm}^2$.

The surface area of the powder material was determined based on the following considerations. Unlike 3D foam, here by the surface area we mean the area of the microscopic relief on the surface formed by the natural roughness due to the structure of the material. If we assume that the powder particles are ideally sintered and form spherical structures, we can assume that the surface is covered with microspheres (more precisely, hemispheres). If we discuss about the effective surface area S_{ef} , we understand that its assessment should be conducted in this study. This parameter cannot be measured directly, but we can model the relative increase in the original (flat) surface area S . The surface relief of aluminum powder compound showed in Figure 2 (a). At the surface level, white particles of sintered aluminum powder with sizes from 60 to 80 μm are visible, located in a matrix of black oxide. The average size of spherical aluminum particles is $70 \pm 10 \mu\text{m}$. 5-10 particles fit on the 1000 μm line. Then the degree of filling of the surface with aluminum particles will be 60%. Let us assume that the roughness of the surface is conventionally sphere-shaped and distributed uniformly over the surface. Then, using the square of the semi-sphere, we can determine the coefficient of increase in the surface area using the following formula:

$$K = \frac{S_{\text{ef}}}{S} = \frac{NS_{\text{sph}}}{2S_{\text{flat}}} \quad (4)$$

Where S_{sph} and S_{flat} are the areas of the lateral surface of the spheres and flat base, respectively; N is the number of spheres pleased on the flat square of sample $S = ab$; If $N \sim (ab)/d^2$

$$K = \frac{ab \pi d^2}{d^2 2ab} = \frac{\pi}{2} \quad (5)$$

According to equation (4), the coefficient K is 1.57. Thus, according to the calculation results, the effective surface area S of the aluminum sample exceeds the nominal area S by 57%. In this case, the calculations assumed that the spherical particles uniformly cover the entire surface. The image shown in Figure 2a visually creates the impression of dark gaps between the spheres. However, this is not true: similar particles are present in these areas, but they are located at a lower level. The uniformity of the surface relief is confirmed by the profilometry data shown in Figure 3a. As can be seen, the vertical profile is an alternation of protrusions and depressions, while the maximum depth of the depressions reaches $R_z = 67$, which is consistent with the previously obtained microscopic data.

An experimental cell of a metal-air battery (MAB) was developed to test the anodes. A schematic cross-sectional view of the device is shown in Figure 4. The electrolyte (position 1) fills the reaction chamber (2) located between the air cathode (3) and the anode (4). The anode under test was immersed in the electrolyte and fixed at a distance of 2 cm from the cathode surface. An aqueous solution of

potassium hydroxide with a concentration in the range of 0.1–1.4 mol/l was used as the electrolyte. The air cathode was made of cardboard based on porous carbon fiber. Current collection was carried out using a metal grid fixed to the cathode using a metal plate.

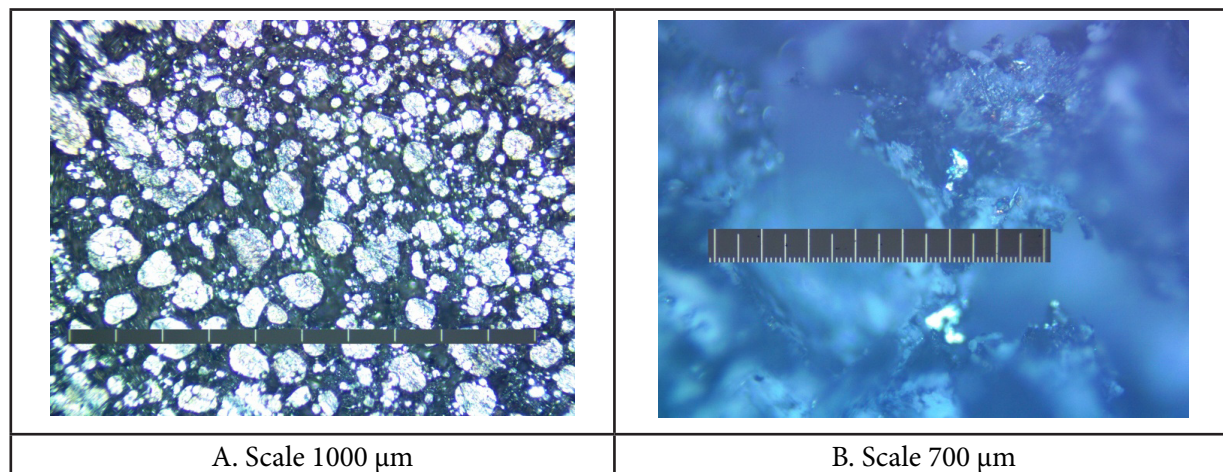


Figure 2 – Particles of aluminum powder (a) and Pores in aluminum foam (b) at 10X optical magnification.

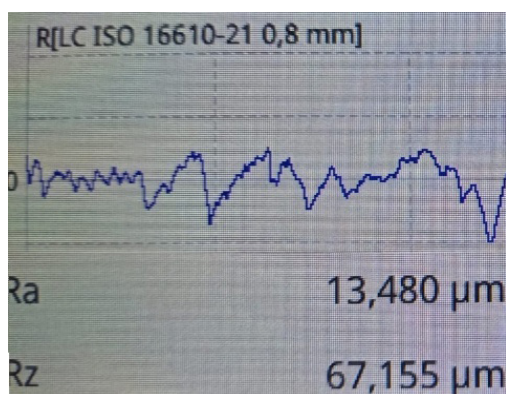


Figure 3 – Roughness of powder alumina surface

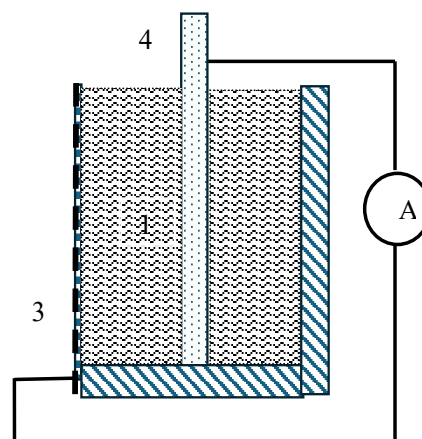


Figure 4 – Elements of the design of the experimental cell MAB for testing anodes

The current was measured using a Fluke 8655 multimeter. Optical images of the surface were taken using a digital camera in macrofocus mode and an AM Scope 580 T microscope. Roughness profile measurements were performed using a device MahrSurf 10S. The device has an induction reference sensor, tip radius 2 μm . The measurement ranges of the roughness parameter are: Ra from 0.02 to 10 μm , Rz, Rmax — from 0.1 to 50 μm . The length of the probing path is 0.56 - 16 mm.

Results and discussion

When powdered aluminum (see Figure 1) interacted with a 1-mol KOH electrolyte, an intense exothermic reaction was observed. During the experiment, the sample significantly lost weight, and white spots characteristic of oxide compounds formed on its surface. Foamed aluminum also demonstrated an active reaction when in contact with the same electrolyte, accompanied by intense gas evolution. In order to preserve the integrity of the three-dimensional structure of the material,

it was decided to reduce the alkali concentration to 0.1 mol for subsequent tests. Unlike the above-mentioned samples, the anode made of cast aluminum demonstrated more stable behavior when interacting with a 1-mol electrolyte solution — the oxidation process occurred uniformly and without pronounced turbulence.

Table 1 shows the mass losses of different electrodes after 30 minutes of operation in 1M KOH electrolyte.

Table 1 – Mass loss of different electrodes within 30 min at a concentration of 1 M

Material	Cast aluminum	Aluminum powder	Foam aluminum
Initial mass, g	5.645	6.158	1.640
Final Mass, g	5.592	4.709	1.0145
Density, g/cm ³	2.437	1.102	0.683
Specific area, S _{ef} , cm ²	12	18.75	33000
Weight loss in 10 min, mg	13	1450	622

The high mass loss coefficient observed in porous materials, especially aluminum foam, is due to their developed specific surface area. This leads to an increase in the area of active reaction, which in turn helps to reduce the current density per unit surface area. The current density was calculated using the following formula:

$$j = \frac{I}{S_{ef}},$$

where I is the short-circuit current measured by the device in Figure 3, and S_{ef} is the effective surface area of the anode, i.e. the area that is in contact with the electrolyte. To calculate the effective area, we take both sides of the surface of the samples in contact with the electrolyte, both from the cathode side and from the back side. Further in the calculations, we took $S = 2ab = 12 \text{ cm}^2$.

Taking these considerations into account, the porous anode samples were tested in the MAB electrochemical cell for all types of anodes in electrolytes with a concentration of 0.1 and 4 mol, respectively. Figure 5a shows the time dependences of the short-circuit current I for anodes made of aluminum powder (p), aluminum foam (f), and monolith (s) during testing for 10 min at 0.1 mol. A characteristic feature of the current behavior for all types of samples is an increase in current at the beginning of the experiment, then its saturation and decline over time. This is due to the activation of the entire surface of the sample after the removal of the oxide layer on the surface.

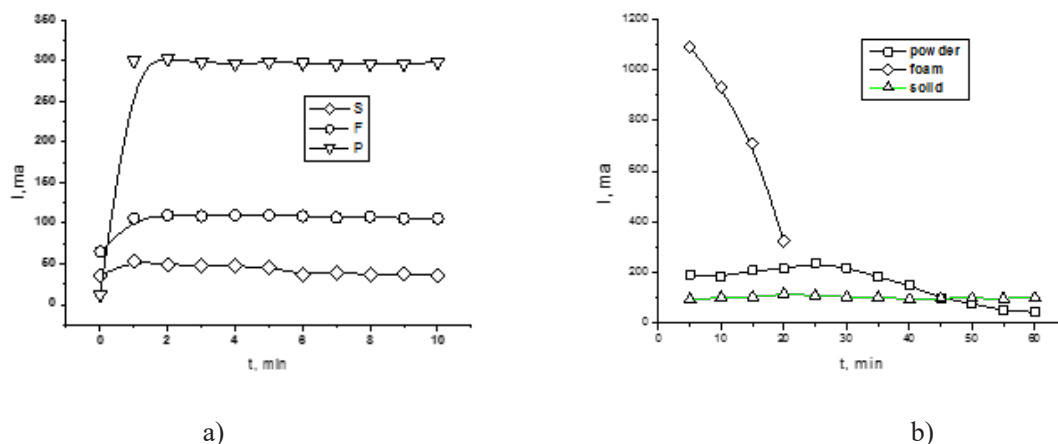


Figure 5 – a) Dependence of the anode current on time for monolithic (s) and powder (p) at a concentration of 1 mol/l, b) aluminum foam (f) at a concentration of 0.1 mol/l)

For the anodes of the first type, the highest current density was recorded for the aluminum powder sample, reaching values of $j = (25-30) \text{ mA/cm}^2$. Let us move on to the analysis of the anodes of the second type. The key observation here is that foam aluminum demonstrates a more than twofold excess of current density compared to monolithic aluminum. This is explained by the capillary effect: the electrolyte actively penetrates the porous structure of the material, which significantly increases the effective area of the electrochemical reaction.

According to the measured values of current density:

- ♦ for foam aluminum – $j = (9-11) \text{ mA/cm}^2$
- ♦ for the monolith – $j = (3-5) \text{ mA/cm}^2$

Considering that the electrolyte concentration in this case is only 0.1 mol, it can be assumed that at a standard concentration of 5–6 mol, the current density will increase by an order of magnitude. For comparison, according to literature data, cast aluminum anodes at a concentration of 6 mol demonstrate an average current density of about 50 mA/cm². However, for the first two types of materials, the effective areas of the samples cannot be considered equal, since the foam material, unlike the powder, interacts with the electrolyte not only on the surface, but also in the bulk structure. Taking into account the above formulas (1–4), the current density for the powder anode decreases by 1.57 times, and for foam aluminum – by 32,000 times. As a result, we obtain the following values: 39 mA/cm² for the powder anode and 0.003 mA/cm² for the foam. Thus, an important conclusion is that the powder anode demonstrates a high current density even at an extremely low concentration of the working electrolyte – 0.1 mol/l.

Next, we will consider the data at an electrolyte concentration of 4 mol/l (Figure 5b). In this case, the highest short-circuit current is demonstrated by foam aluminum. The current values exceed 1 A and are not the limit, since in some experiments currents of up to 5 A were recorded at a cell voltage of up to 4 V. However, at a high electrolyte concentration (4 mol/l and higher), the foam quickly loses mass, and after 20 minutes the sample is completely destroyed. Unlike foam, the powder anode shows more stable behavior: in the first 30 minutes, the current increases slightly within 190–220 mA, and then gradually decreases. This is due to a decrease in the anode mass, since the active area is reduced when the lateral surface becomes thinner. The behavior of the monolith remains unchanged, similar to that observed at a concentration of 0.1 mol/l. Taking into account the specific area, the current density for the three materials is: 0.03 mA/cm² for the foam, 26 mA/cm² for the powder, and 8.3 mA/cm² for the monolith. At the same time, the calculations did not take into account the decrease in the surface area and volume of the samples during etching, so the given values reflect the initial stage of the experiment. As a result, at a concentration of 4 mol/l, the current density of the powder anode is again higher than that of the monolith.

Conclusion

Thus, porous materials demonstrate different current densities depending on their manufacturing method. For powder anodes, this value is close to the values of monolithic samples (26–39 mA/cm²), but the high etching rate leads to structural changes, which requires additional study. At the same time, the mass of powder anodes is 10–30% lower, which is an advantage in the context of the mass-dimensional characteristics for MABs.

The foam material showed significantly higher current characteristics compared to powder samples. However, it is important to consider that intense current leads to accelerated destruction of the aluminum foam electrode. Despite this, when developing MABs as energy storage devices, where a sharp increase in current output is required while maintaining the overall mass of the device, this parameter may be critically important.

A notable effect recorded during the experiment was the ability of aluminum foam to draw in electrolyte due to the capillary effect. In the future, this phenomenon can be used to create high-speed MABs that operate without the need for forced injection of electrolyte.

In practical terms, the materials studied here can be considered for use in systems that operate in intermittent or emergency modes, where a fast current response is required. Such metal–air modules

may become relevant for small backup power units, portable field electronics and communication sensors that work away from centralized infrastructure. At the same time, the reduced mass of porous anodes opens up the possibility of using them in compact autonomous energy systems, where every additional gram is critical. These aspects outline a realistic direction for further development of the proposed solutions.

The experiments conducted confirm the potential of using porous anodes in the creation of highly efficient autonomous power microsystems (APM).

Information about funding. The work was carried out with the support of grants from the Ministry of Education and Science of the Republic of Kazakhstan. BR21882187 and IPH AP19676182.

REFERENCES

- 1 Narayanan, S.R., Prakash, G.K.S., Manohar, A., Yang, B., Malkhandi, S., & Kindler, A. Materials challenges and technical approaches for realizing inexpensive and robust iron–air batteries for large-scale energy storage. *Solid State Ionics*, 216, 105–109 (2012). <https://doi.org/10.1016/j.ssi.2012.05.016>.
- 2 Han, X., Li, X., White, J., Zhong, C., Deng, Y., Hu, W., & Ma, T. Metal–air batteries: From static to flow system. *Advanced Energy Materials*, 8, 1801396 (2018). <https://doi.org/10.1002/aenm.201801396>.
- 3 Fan, X., Liu, B., Liu, J., Ding, J., Han, X., Deng, Y., Lv, X., Xie, Y., Chen, B., Hu, W., et al. Battery technologies for grid-level large-scale electrical energy storage. *Transactions of Tianjin University*, 26, 92–103 (2020). <https://doi.org/10.1007/s12209-019-00236-w>.
- 4 Salameh, T., Sayed, E.T., Abdelkareem, M.A., Olabi, A.G., & Rezk, H. Optimal selection and management of hybrid renewable energy system: Neom city as a case study. *Energy Conversion and Management*, 244, 114434 (2021). <https://doi.org/10.1016/j.enconman.2021.114434>.
- 5 Li, Y., & Lu, J. Metal–air batteries: Will they be the future electrochemical energy storage device of choice? *ACS Energy Letters*, 2(6), 1370–1377 (2017). <https://doi.org/10.1021/acsenergylett.7b00119>.
- 6 Wang, H.F., & Xu, Q. Materials design for rechargeable metal–air batteries. *Matter*, 1(3), 565–595 (2019). <https://doi.org/10.1016/j.matt.2019.05.018>.
- 7 Parker, J.F., Chervin, C.N., Nelson, E.S., Rolison, D.R., & Long, J.W. Wiring zinc in three dimensions rewrites battery performance – dendrite-free cycling. *Energy & Environmental Science*, 7, 1117–1124 (2014). <https://doi.org/10.1039/C3EE43674H>.
- 8 Liu, P., Ling, X., Zhong, C., Deng, Y., Han, X., & Hu, W. Porous zinc anode design for Zn–air chemistry. *Frontiers in Chemistry*, 7, 656 (2019). <https://doi.org/10.3389/fchem.2019.00656>.
- 9 Pino, M., Cuadrado, C., Chacón, J., et al. The electrochemical characteristics of commercial aluminium alloy electrodes for Al–air batteries. *Journal of Applied Electrochemistry*, 44, 1371–1380 (2014). <https://doi.org/10.1007/s10800-014-0739-4>.
- 10 Mutlu, R.N., & Yazici, B. Copper-deposited aluminum anode for aluminum–air battery. *Journal of Solid State Electrochemistry*, 23, 529–541 (2019). <https://doi.org/10.1007/s10008-018-4147-7>.
- 11 Ma, J., Wen, J., Gao, J., et al. Performance of Al–0.5Mg–0.02Ga–0.1Sn–0.5Mn as anode for Al–air battery in NaCl solutions. *Journal of Power Sources*, 253, 419–423 (2014). <https://doi.org/10.1016/j.jpowsour.2013.12.088>.
- 12 Li, D., Liu, Y., Xie, Y., et al. Porous powder anode for high performance rechargeable aluminum batteries. *Journal of Power Sources*, 641, 236860 (2025). <https://doi.org/10.1016/j.jpowsour.2025.236860>.
- 13 Mori, R. Recent developments for aluminum–air batteries. *Electrochemical Energy Reviews*, 3, 344–369 (2020). <https://doi.org/10.1007/s41918-020-00065-4>.
- 14 Ma, Y., Sumboja, A., Zang, W., et al. Flexible and wearable all-solid-state Al–air battery based on iron carbide encapsulated in electrospun porous carbon nanofibers. *ACS Applied Materials & Interfaces*, 11, 1988–1995 (2019). <https://doi.org/10.1021/acsami.8b15877>.
- 15 Mohamed, S.G., Tsai, Y.Q., Chen, C.J., et al. Ternary spinel MCo₂O₄ (M = Mn, Fe, Ni, and Zn) porous nanorods as bifunctional cathode materials for lithium–O₂ batteries. *ACS Applied Materials & Interfaces*, 7, 12038–12046 (2015). <https://doi.org/10.1021/acsami.5b02601>.
- 16 Niu, Q., Chen, B., Guo, J., et al. Flexible, porous, and metal–heteroatom-doped carbon nanofibers as efficient ORR electrocatalysts for Zn–air battery. *Nano-Micro Letters*, 11, 8 (2019). <https://doi.org/10.1007/s40820-018-0231-0>.

- 17 Liu, W., Placke, T., & Chau, K.T. Overview of batteries and battery management for electric vehicles. *Energy Reports*, 8, 4058–4084 (2022). <https://doi.org/10.1016/j.egyr.2022.03.016>.
- 18 Sun, Q., Dai, L., Luo, T., Wang, L., Liang, F., & Liu, S. Recent advances in solid-state metal–air batteries. *Carbon Energy*, 2022, 1–23. <https://doi.org/10.1002/cey2.276>.
- 19 Wang, Y., Sun, Y., Ren, W., Zhang, D., Yang, Y., Yang, J., Wang, J., Zeng, X., & NuLi, Y. Challenges and prospects of Mg–air batteries: A review. *Energy Materials*, 2, 200024 (2022). <https://doi.org/10.20517/energymater.2022.20>.
- 20 Wang, L., Snihirova, D., Deng, M., Vaghefinazari, B., Xu, W., Höche, D., Lamaka, S.V., & Zheludkevich, M.L. Sustainable aqueous metal–air batteries: An insight into electrolyte system. *Energy Storage Materials*, 52, 573–597 (2022). <https://doi.org/10.1016/j.ensm.2022.08.032>.
- 21 Li, J., Zhang, K., Wang, B., & Peng, H. Light-assisted metal–air batteries: Progress, challenges, and perspectives. *Angewandte Chemie International Edition*, 61, e202213026 (2022). <https://doi.org/10.1002/anie.202213026>.
- 22 Chantavas, A. Global market outlook for solar power. *SolarPower Europe*, 2022, 1–25. <https://www.solarpowereurope.org/press-releases/world-installs-a-record-168-gw-of-solar-power-in-2021-enters-solar-terawatt-age>.
- 23 Liu, Y., Lu, X., Lai, F., Liu, T., Shearing, P.R., Parkin, I.P., He, G., & Brett, D.J.L. Rechargeable aqueous Zn-based energy storage devices. *Joule*, 5, 2845–2903 (2021). <https://doi.org/10.1016/j.joule.2021.10.011>.
- 24 Murali, A.P., Duraisamy, S., Samuthiram, S., et al. Current and emerging methods for manufacturing of closed pore metal foams and its characteristics: A review. *Journal of Materials Science*, 60, 1187–1227 (2025). <https://doi.org/10.1007/s10853-024-10318-y>.
- 25 Mahto, R.P., Bhaduria, A., Bandhu, D., et al. A study on porosity and mechanical properties of the open aluminum metal foam through spark plasma sintering SDP technique. *International Journal of Advanced Manufacturing Technology*, 136, 4407–4417 (2025). <https://doi.org/10.1007/s00170-025-15077-x>.
- 26 Sun, S., Zhang, Z., Yan, L., et al. A novel superimposed porous copper/carbon film derived from polymer matrix as catalyst support for metal–air battery. *Journal of Porous Materials*, 29, 249–255 (2022). <https://doi.org/10.1007/s10934-021-01163-4>.
- 27 Jiao, X., Liu, Y., Cai, X., et al. Progress of porous Al-containing intermetallics fabricated by combustion synthesis reactions: A review. *Journal of Materials Science*, 56, 11605–11630 (2021). <https://doi.org/10.1007/s10853-021-06035-5>.

¹Жукешов А.М.,

ф.-м.ғ.д., профессор, ORCID ID: 0000-0001-9153-323X,

e-mail: zhukeshov@physics.kz

^{1*}Абдыбай Ұ.Б.,

докторант, ORCID ID: 0009-0002-5309-5827,

*e-mail: abdibay_ulan@mail.ru

¹Салимов Е.Е.,

докторант, ORCID ID: 0009-0005-7712-6263,

e-mail: Salimov_ernar@mail.ru

²Карлос М.,

профессор, ORCID ID: 0000-0002-7374-505X,

e-mail: Carlos.meza@ieee.org

¹Әл-Фараби атындағы Қазақ ұлттық университеті, Алматы қ., Қазақстан

²Анхальт қолданбалы ғылымдар университеті, Анхальт, Германия

ҚҰЙМА ЖӘНЕ КЕУЕКТІ ЭЛЕКТРОДТАРҒА НЕГІЗДЕЛГЕН МЕТАЛЛ-АУА БАТАРЕЯЛАРЫНЫҢ ӨНІМДІЛІК СИПАТТАМАЛАРЫНЫҢ САЛЫСТЫРМАЛЫ ТАЛДАУЫ

Аңдатпа

Қазіргі уақытта электр энергиясына деген өсіп келе жатқан сұранысты қанағаттандыруда электрохимиялық энергия сақтау жүйелерінің дамуы негізгі рөл атқарады. Жоғары меншікті энергия сыйым-

дылығына ие металл-ауа батареялары (МAB) электр станцияларында резервтік қуат көздері ретінде пайдалануға перспективті шешімдердің бірі болып саналады. Оларды кеңінен енгізудегі негізгі шектеулердің бірі – анодтық материалдардың тиімділігін арттыру қажеттілігі. Бұл мақалада МAB-тың жұмыс тиімділігін арттыру мақсатында кеуекті алюминий электродтарын қолдану ұсынылады. Өртүрлі технологиялар арқылы жасалған екі түрдегі кеуекті анодтар зерттелді. Ұнтақ алюминий анодында ток тығыздығы 20–30 мА/см² деңгейінде болып, монолитті (стандартты) анодтың көрсеткіштеріне сәйкес келді. Ал алюминий көбігі ток тығыздығының едәуір жоғары мәндерін – 52–64 мА/см² көрсетті. Кеуекті анодтардың қосымша артықшылығы – олардың салмағының 10–30%-ға аз болуы, бұл өз кезегінде МAB-тың массалық-габариттік сипаттамаларын жақсартуға мүмкіндік береді және неғұрлым тиімді энергия жүйелерін құруға жол ашады.

Тірек сөздер: металл-ауа батареясы, алюминий аноды, кеуекті электрод, алюминий құймасы, электрохимиялық сипаттамалар, меншікті беткі аудан, энергия сақтау.

¹Жукешов А.М.,

д.ф.-м.н., профессор, ORCID ID: 0000-0001-9153-323X,

e-mail: zhukeshov@physics.kz

^{1*}Абдыбай У.Б.,

докторант, ORCID: 0009-0002-5309-5827,

*e-mail: abdibay_ulan@mail.ru

¹Салимов Е.Е.,

докторант, ORCID: 0009-0005-7712-6263,

e-mail: Salimov_ernar@mail.ru

²Карлос М.,

профессор, ORCID ID: 0000-0002-7374-505X,

e-mail: Carlos.meza@ieee.org

¹Казахский национальный университет им. аль-Фараби, г. Алматы, Казахстан

²Университет прикладных наук Анхальта, Анхальт, Германия

СРАВНИТЕЛЬНЫЕ ХАРАКТЕРИСТИКИ РАБОТЫ МЕТАЛЛ-ВОЗДУШНЫХ АККУМУЛЯТОРОВ С ЛИТЫМИ И ПОРИСТЫМИ ЭЛЕКТРОДАМИ

Аннотация

На сегодняшний день развитие электрохимических систем хранения энергии имеет важное значение для удовлетворения растущего спроса на электроэнергию. Металл-воздушные батареи (МAB) из-за присущей им высокой энергоёмкости являются перспективными устройствами для использования в энергетических установках как резервных источников. Проблема полноценного внедрения МAB в эти системы упирается в решение вопроса об эффективных анодах. В настоящей работе предлагается использовать пористые материалы для увеличения эффективности работы МAB. Пористые алюминиевые электроды двух типов с разной технологией изготовления были испытаны в качестве анодов для МAB. Для анода из порошкового алюминия плотность тока составила (20–30) мА/см², что сравнимо с током для монолитного (стандартного) анода, но для пеноалюминия были получены более высокие значения плотности тока (52–64) мА/см². Так как масса пористых анодов будет на (10–30)% ниже, это является преимуществом в разработке новых типов МAB с улучшенными массогабаритными показателями.

Ключевые слова: металл-воздушная батарея, алюминиевый анод, пористый электрод, алюминиевая пена, электрохимические характеристики, удельная поверхность, накопление энергии.

Article submission date: 12.09.2025

МҰНАЙ ГАЗ ИНЖЕНЕРИЯСЫ ЖӘНЕ ГЕОЛОГИЯ OIL AND GAS ENGINEERING, GEOLOGY НЕФТЕГАЗОВАЯ ИНЖЕНЕРИЯ И ГЕОЛОГИЯ

УДК 550.5574
МРНТИ 26.29.13

<https://doi.org/10.55452/1998-6688-2025-22-4-411-418>

^{1*}Базарбай Д.Қ.,
магистрант, ORCID ID: 0009-0001-2964-3523,
*e-mail: d_bazarbai@mail.ru

¹Казахстанско-Британский технический университет, г. Алматы, Казахстан

ИЗУЧЕНИЕ ПЛАСТОВ-КОЛЛЕКТОРОВ ВЕРХНЕКУМКОЛЬСКОЙ ПОДСВИТЫ МЕСТОРОЖДЕНИЯ ЮЖНЫЙ КУМКОЛЬ

Аннотация

Своевременность проведения данного исследования обусловлена потребностью в детальном анализе извлекаемых запасов нефти в продуктивных горизонтах, что связано с условиями осадконакопления и процессами образования залежей в Южно-Тургайском осадочном бассейне. Для точного определения условий образования пород требуется комплексный анализ, включающий гранулометрические исследования терригенных отложений. Такой подход позволяет получить достоверные данные о среде осадконакопления и провести более детальный фациальный анализ. Целью работы является определение фациальной обстановки осадконакопления продуктивного горизонта Ю-II месторождения Южный Кумколь с использованием гранулометрического анализа. Методология исследования включает использование различных методов гранулометрического анализа. Базовым методом для выявления условий была акцентирована по методике Фюхтбауэра и Мюллера. Генетическая интерпретация осадков проводится с использованием диаграммы К. Бьёрликке, которая исследует соотношение отсортированности и асимметрии распространения частиц. Для более достоверного уточнения кондиции седиментации применяется динамогенетическая диаграмма Г.Ф. Рожкова, учитывающая относительность асимметрии и эксцесса. Результаты исследования позволяют более объективно охарактеризовать условия формирования пород-коллекторов и уточнить параметры осадконакопления. Комплексный подход к анализу гранулометрических данных способствует повышению точности фациальной диагностики нефтеносных горизонтов.

Ключевые слова: фациальный анализ, гранулометрический анализ, генетические диаграммы, месторождение Южный Кумколь, верхнекумкольская подсвита.

Введение

Оценка условий формирования продуктивных горизонтов нефтяных месторождений является важной частью их литологического и фациального изучения. В пределах Южно-Тургайского осадочного бассейна гранулометрия терригенных отложений играет ключевую роль в понимании динамических условий накопления и генезиса залежей.

Гранулометрический анализ осадков позволяет установить размерную структуру зернистой части, понять условия транспортировки и накопления обломочного материала. При этом применение комплексного подхода, включающего изучение классификаций Фюхтбауэра и Мюллера [4, 5], а также диаграмм К. Бьёрликке и Г.Ф. Рожкова [8, 10], способствует более точной фациальной диагностике.

Цель настоящей работы – уточнение фациальных условий осадконакопления верхнекумольской подсистемы на месторождении Южный Кумколь путем анализа гранулометрических характеристик.

Материалы и методы

Методика исследования и ее апробация

Для имплементации цели применялся комплекс методов гранулометрического анализа, включающий:

- ♦ визуальное изучение прозрачных петрографических шлифов;
- ♦ механическое ситовое разделение фракций крупнее 0,05 мм;
- ♦ пипеточное разделение мелких частиц.

На основе полученных данных были построены гистограммы и кумулятивные кривые распределения, позволившие оценить степень сортировки пород, характер асимметрии распределения фракций, а также определить статистические параметры: средний размер зерен, стандартное отклонение, асимметрию и эксцесс.

С целью интерпретации условий осадконакопления использовался ряд апробированных в литологической практике графоаналитических моделей. В частности, применялись:

- ♦ обобщенная схема К. Фюхтбауэра и Д. Мюллера, позволяющая предварительно классифицировать осадочные обстановки как морские, континентальные или переходные;
- ♦ диаграмма К. Бьёрликке, показывающая зависимость между степенью сортировки, асимметрией и генезисом осадков;
- ♦ динамогенетическая диаграмма Г.Ф. Рожкова, учитывающая степень механического разделения обломочного материала в зависимости от фациальных условий.

Комплексное применение указанных методик позволяет значительно повысить точность интерпретации условий седиментации, обеспечивая более надежную реконструкцию палеогеографической обстановки и фациальной принадлежности коллекторов.

Значимую роль в модернизированной литологии играют подходы, базирующиеся на гранулометрических особенностях, разработанные Р. Пассегой, Г. Фридманом, Л.Б. Рухиным, Г.Ф. Рожковым и др. Применение таких методов, как диаграмма Бьёрликке, показало, что отрицательная асимметрия характерна для маринных фаций океанического типа, континентальные типы имеют положительные асимметрию с выдержанностью распространения [8, 9].

Высокую эффективность также демонстрирует динамогенетическая диаграмма Р. Пассеги и Г.Ф. Рожкова, построенная на основе анализа степени механической дифференциации частиц. Основной фактор дифференциации – энергетический потенциал среды осадконакопления, определяющий транспортировку и осаждение частиц различного гранулометрического состава [10].

Таким образом, анализ существующих методик показывает, что при наличии сходных гранулометрических характеристик осадки могут быть сформированы в различных фациальных обстановках. Это требует использования комплексного подхода, объединяющего различные методы анализа. В рамках настоящей работы предложена следующая схема:

1. Предварительная диагностика условий осадконакопления по схеме Фюхтбауэра–Мюллера [4, 5];

2. Определение генезиса отложений на основе степени сортировки и асимметрии (диаграмма К. Бьёрликке);

3. Уточнение условий седиментации по асимметрии и эксцессу (диаграмма Г.Ф. Рожкова).

Данная последовательность была испытана на разрезах верхне- и среднеюрских отложений Ем-Еговской площади Красноленинского месторождения (Западная Сибирь) [13]. Продуктивные пласты ассоциированы с арыскупским горизонтом и верхнекумкольской свитой. Коллекторы верхнекумкольской подсвиты представлены мелко- и крупнозернистыми песчанистыми алевролитами, реже песчаниками. Гранулометрические методы показали, что размер обломочных частиц колеблется от 0,01 до 0,20 мм, в отдельных случаях достигая 0,38 мм. Наиболее распространенная фракция – зерна размером 0,04–0,08 мм, составляющие до 50% от общей обломочной массы [14].

Результаты и обсуждение

В процессе были изучены 22 образца керна, отобранных из слоев верхнекумкольской подсвиты двух скважин на месторождении Южный Кумколь. Однако были показаны только полученные по восьми образцам, извлеченным из скважин № 1 и № 25.

Для определения гранулометрического состава коллекторских пород использовались два подхода. Это обусловлено высокой степенью уплотненности и прочным цементным связыванием песчано-алевритовых пород указанного возраста, что затрудняет их механическое разделение на отдельные зерна [15].

В дальнейшем были построены кумулятивные кривые, которые не только представляют аналитическую информацию в наглядной форме, но и служат инструментом для определения количественных параметров, описывающих структуру породы (рисунок 1).

По каждому из образцов были рассчитаны основные гранулометрические характеристики: средний размер частиц (М), медиана (Md), коэффициент асимметрии (В), мера эксцесса (R) и стандартное отклонение (коэффициент сортировки).

Средний размер определялся по формуле (1) [10]:

$$M = \frac{(\phi_{16} + \phi_{50} + \phi_{84})}{3}. \quad (1)$$

Коэффициент сортировки зерен рассчитывался по формуле (2) [15]:

$$F = \sqrt{\frac{Q_3}{Q_1}}. \quad (2)$$

Асимметрия вычислялся по формуле (3) [10]:

$$B = \frac{\phi_{16} + \phi_{84} - 2 * \phi_{50}}{2 * (\phi_{84} - \phi_{16})} + \frac{\phi_{5} + \phi_{95} - 2 * \phi_{50}}{2 * (\phi_{95} - \phi_{5})}. \quad (3)$$

Эксцесс определялся по формуле (4) [10]:

$$R = \frac{\phi_{95} - \phi_{5}}{2,44 * (\phi_{75} - \phi_{25})}. \quad (4)$$

Примеры, полученные для образцов из скважин 1 и 25 месторождения Южный Кумколь (таблица 1).

Таблица 1 – Описание атрибутов (скважины № 1 и № 25 месторождения Южный Кумколь)

№Скв.	№ обр. Sample	Глубина, м Depth, m	Порода Rock	M	F	Md, мм (mm)	B	R
1	813	1315,5	Песч. глин.-карб.	0,187	0,479	0,016	0,962	3,288
	815	1321	Алевр. глин.-песч.	0,035	0,679	0,009	0,872	1,695
	817	1325	Алевр. глин.-песч.	0,031	0,695	0,009	0,854	1,579
	819	1325	Алевролит глин.	0,025	0,658	0,003	0,937	2,493
	821	1325	Песч. сильно алевр.	0,083	0,619	0,023	0,862	1,629
25	823	1272,5	Песчаник	0,077	0,643	0,025	0,828	1,451
	825	1272,5	Песчаник	0,146	0,483	0,005	0,987	5,942
	827	1277,5	Алевролит	0,080	0,676	0,037	0,744	1,186
	829	1280,45	Алевролит	0,067	0,689	0,031	0,747	1,193
	831	1280,45	Алевролит	0,051	0,660	0,014	0,861	1,622
	833	1286,5	Гравелит	0,437	0,369	0,212	0,725	1,145

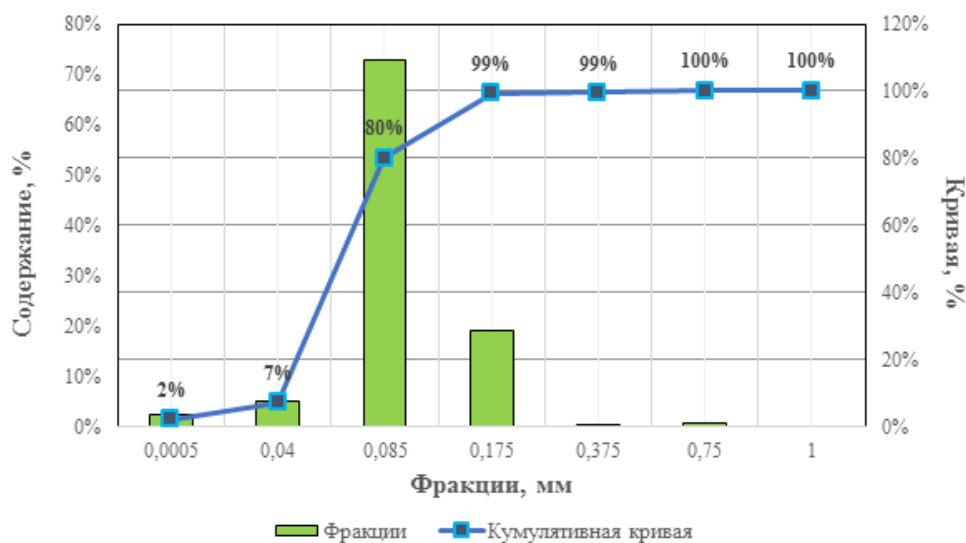


Диаграмма 1 – Накопительная кривая зерен по фракциям для скважины № 25 месторождения Южный Кумколь.

Этот вывод подтверждается низкой степенью сортировки (коэффициент сортировки менее 1) и отрицательной асимметрией (показатель асимметрии ниже 1), что является характерным для морских осадков. В то же время сопоставление показателей отсортированности и асимметрии по диаграмме К. Бьёрликке свидетельствует о принадлежности большинства исследуемых отложений к речному генезису (рисунок 1). Это указывает на возможность наложения различных фациальных признаков, что требует дальнейшего уточнения условий осадконакопления с применением дополнительных диагностических критериев.

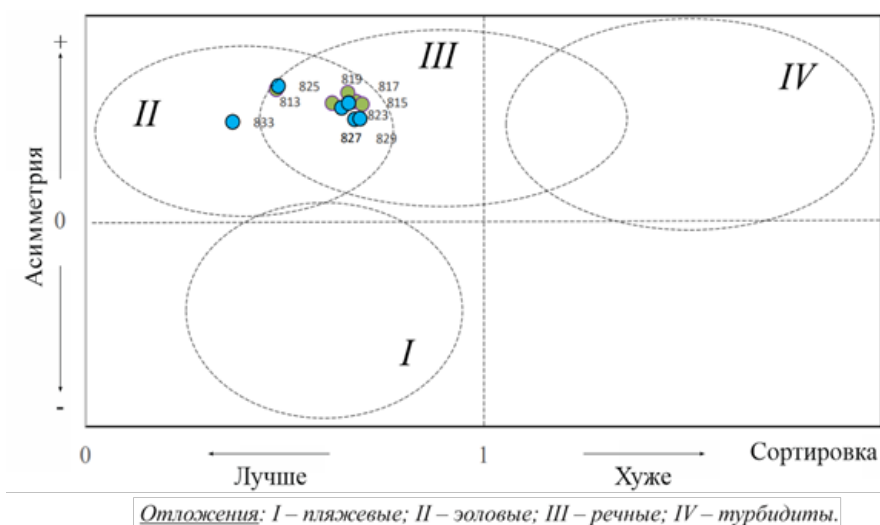
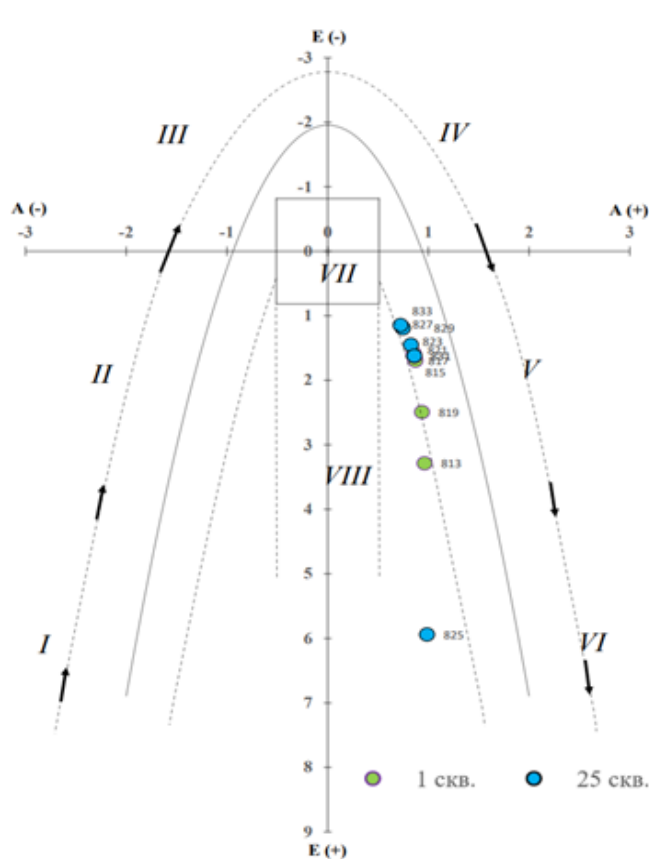


Рисунок 1 – Относительность сортировки
по скважинам 1 и 25 месторождения Южный Кумколь
(по К. Бьёрликке, 1989 г.)



Поля: I – дно водоемов, II – зоны сноса, III – реки со слабым приливом IV – реки с сильным приливом, V – приливно-отливная зона, VI – береговые дюны, VII – золотые отложения, VIII – сильный прибой волн с интенсивной переработкой осадков – прибрежные фации открытых водоемов.

Рисунок 2 – Динамогенетическая диаграмма Г.Ф. Рожкова
для скважин № 1 и № 25 месторождения Южный Кумколь

Заключение

Результаты исследования показали, что осадконакопление продуктивного горизонта Ю-II месторождения Южный Кумколь происходило в дельтовой обстановке прибрежной морской зоны под воздействием приливных течений. Гранулометрический анализ выявил, что породы-коллекторы представлены мелко- и среднезернистыми песчаниками и алевролитами, отражающими переменную гидродинамику осадочной среды.

Анализ по диаграмме К. Бьёрликке указал на сочетание признаков речных и прибрежно-морских отложений, что свидетельствует о сложной фациальной природе осадков. Динамогенетическая диаграмма Г.Ф. Рожкова подтвердила формирование в условиях активной гидродинамики дельтовой системы с влиянием морских приливов. Это проявляется в положительной асимметрии распределения частиц и умеренной сортировке, характерных для осадков смешанного генезиса.

Таким образом, коллекторы верхнекумкольской подсвиты сформировались в дельтовой системе с участием речных и приливных процессов, обеспечивших перераспределение осадочного материала. Комплексный гранулометрический анализ позволил уточнить фациальную принадлежность пород и повысить точность прогноза их коллекторских свойств.

ЛИТЕРАТУРА

- 1 Романовский С.И. Седиментологические основы литологии. – Л.: Недра, 1977. – 408 с.
- 2 Рухин Л.Б. Гранулометрический анализ песков. – Л.: ЛГУ, 1947. – 213 с.
- 3 Логвиненко Н.В., Орлова Л.В. Образование и изменение осадочных пород на континенте и в океане. – Л.: Недра, 1987. – 237 с.
- 4 Крашенинников Г.Ф., Волкова А.Н., Иванова Н.В. Учение о фациях с основами литологии. – М.: Изд-во МГУ, 1988. – 214 с.
- 5 Рейнек Г.-Э., Сингх И.Б. Обстановки терригенного осадконакопления (с рассмотрением терригенных кластических осадков). – М.: Недра, 1981. – 439 с.
- 6 Реддинг Х.Г. Обстановки осадконакопления и фации. – М.: Мир, 1990. – 352 с.
- 7 Вакуленко Л.Г., Предтеченская Е.А., Чернова Л.С. Опыт применения гранулометрического анализа для реконструкции условий формирования песчаников продуктивных пластов васюганского горизонта (Западная Сибирь) // Литосфера. – 2003. – № 3. – С. 99–108.
- 8 Кузнецов В.Г. Литология. Осадочные горные породы и их изучение. – М.: ООО «Недра-Бизнес-центр», 2007. – 511 с.
- 9 Селли Р.Ч. Древние обстановки осадконакопления. – М.: Недра, 1989. – 294 с.
- 10 Гроссгейм В.А., Бескровная О.В., Геращенко И.Л. и др. Методы палеогеографических реконструкций (при поисках залежей нефти и газа). – Л.: Недра, 1984. – 271 с.
- 11 Bull W.B. Relation of textural (CM) patterns to depositional environment of alluvial-fan deposits // Journal of Sedimentary Petrology. – 1962. – V. 32. – No. 2. – P. 211–216.
- 12 Петтиджон Ф., Поттер П., Сивер Р. Пески и песчаники. – М.: Мир, 1976. – 536 с.
- 13 Кудряшова Л.К. Изучение литолого-фациальной модели для увеличения нефтеотдачи залежи на примере песчаных пластов Тюменской свиты Красноленинского месторождения», Развитие минерально-сырьевой базы Сибири: от В.А. Обручева, М.А. Усова, Н.Н. Урванцева до наших дней: Матер. I Всеросс. геологической молодежной школы. – Томск: ТПУ, 2013. – С. 88–91.
- 14 Кудряшова Л.К. Гранулометрический анализ как основной метод обоснования условий формирования пластов-коллекторов ЮК (2-5) Ем-Еговской площади (Западная Сибирь) // Известия Томского политехнического университета. Инжиниринг георесурсов. – 2015. – Т. 326. – № 10. – 143 с.
- 15 Япаскурт О.В. Литология. – М.: ИЦ «Академия», 2008. – 336 с.

REFERENCES

- 1 Romanovskij, S.I. Sedimentologicheskie osnovy litologii (L.: Nedra, 1977), 408 p.
- 2 Ruhin, L.B. Granulometricheskij analiz peskov (L.: LGU, 1947), 213 p.
- 3 ëLogvinenko, N.V., Orlova, L.V. Obrazovanie i izmenenie osadochnyh po-rod na kontinente i v okeane (L.: Nedra, 1987), 237 p.
- 4 Krashennnikov, G.F., Volkova, A.N., Ivanova, N.V. Uchenie o facijah s osnovami litologii (M.: Izd-vo MGU, 1988), 214 p.
- 5 ëRejnek, G.-Je., Singh, I.B. Obstanovki terrigenного osadkonakoplenija (s rassmotreniem terrigennyh klasticheskikh osadkov) (M.: Nedra, 1981), 439 p.
- 6 Reding, H.G. Obstanovki osadkonakoplenija i facii (M.: Mir, 1990), 352 p.
- 7 Vakulenko, L.G., Predtechenskaja, E.A., Chernova, L.S. Opyt primenenija granulometricheskogo analiza dlja rekonstrukcii uslovij formirovanija peschanikov produktivnyh plastov vasjuganskogo gorizonta (Zapadnaja Si-bir'). Litosfera, 3, 99–108 (2003).
- 8 Kuznecov, V.G. Litologija. Osadochnye gornye porody i ih izuchenie (M.: ООО «Nedra-Biznescentr», 2007), 511 p.
- 9 Selli, R.Ch. Drevnie obstanovki osadkonakoplenija (M.: Nedra, 1989), 294 p.
- 10 Grossgejm, V.A., Beskrovnaja, O.V., Gerashhenko, I.L. i dr. Metody paleogeograficheskikh rekonstrukcij (pri poiskah zalezhej nefi i gaza) (L.: Nedra, 1984), 271 p.
- 11 Bull, W.B. Relation of textural (CM) patterns to depositional environment of alluvial-fan deposits. J. Sediment. Petrol, 32, 211–216 (1962).
- 12 Pettidzhon, F., Potter, P., Siver, R. Peski i peschaniki (M.: Mir, 1976), 536 p.
- 13 Kudrjashova, L.K. Izuchenie litologo-facial'noj modeli dlja uveli-chenija nefteotdachi zalezhi na primere peschanyh plastov Tjumenskoj svity Krasnoleninskogo mestorozhdenija», Razvitie mineral'nosyr'evoj bazy Sibiri: ot V.A. Obrucheva, M.A. Usova, N.N. Urvanceva do nashih dnei: mater. I Vseross. geologicheskoi molodezhnoj shkoly (Tomsk: TPU, 2013), pp. 88–91.
- 14 Kudrjashova, L.K. Granulometricheskij analiz kak osnovnoj metod obosnovanija uslovij formirovanija plastov-kollektorov JuK (2-5) Em-Egovskoj ploshhadi (Zapadnaja Sibir'). Izvestija Tomskogo politehnicheskogo universiteta. Inzhiniring georesursov, 326 (10), 143 (2015).
- 15 Japaskurt O.V. Litologija (M.: Publishing Center «Akademija», 2008), 336 p.

¹*Базарбай Д.Қ.,

магистрант, ORCID ID: 0009-0001-2964-3523,

*e-mail: d_bazarbai@mail.ru

¹Қазақстан-Британ техникалық университеті, Алматы қ., Қазақстан

ОҢТҮСТІК ҚҰМКӨЛ КЕН ОРНЫНЫҢ ЖОҒАРҒЫ ҚҰМКӨЛ КІШІ ШӨГІНДІ ҚАБАТЫНЫҢ ЖЫНЫС-КОЛЛЕКТОРЛАРЫН ЗЕРТТЕУ

Аңдатпа

Бұл зерттеудің өзектілігі өнімді қабаттардағы мұнай қорының шығу ерекшеліктерін егжей-тегжейлі талдауға деген қажеттілікпен түсіндіріледі. Бұл, өз кезегінде, шөгінді жиналу жағдайларымен және Оңтүстік Торғай шөгінді алабындағы кен орындарының түзілу процестерімен тығыз байланысты. Жыныстардың түзілу жағдайларын дәл анықтау үшін терригендік шөгінділердің гранулометриялық талдауын қамтитын кешенді зерттеу қажет. Мұндай тәсіл шөгінді жиналу ортасы туралы шынайы деректер алуға және фациялық талдауды неғұрлым дәл жүргізуге мүмкіндік береді. Зерттеудің мақсаты – гранулометриялық талдау әдісін пайдалана отырып, Оңтүстік Құмкөл кен орнының Ю-II өнімді горизонтының шөгінді жиналу фациялық жағдайларын айқындау. Зерттеу әдістемесі түрлі гранулометриялық талдау тәсілдерін қолдануға негізделеді. Негізгі әдіс ретінде Фюхтбауэр мен Мюллер әдістемесіне сүйену қарастырылған. Шөгінділердің генетикалық интерпретациясы Бьёрликке диаграммасы арқылы жүзеге асырылады, ол бөлшектердің ірілігі мен асимметриясының өзара байланысын талдайды. Седиментация жағдайларын неғұрлым нақтылау мақсатында асимметрия мен эксцесс арасындағы қатынасты есепке алатын Г.Ф. Рожковтың динамикалық-генетикалық диаграммасы қолданылады. Зерттеу нәтижелері жыныс-коллекторлардың түзілу жағдайларын неғұрлым объективті сипаттауға және шөгінді жиналу параметрлерін нақтылауға мүмкіндік береді.

Гранулометриялық деректерді кешенді талдау тәсілі мұнайлы қабаттардың фациялық диагностикасының дәлдігін арттыруға ықпал етеді.

Тірек сөздер: фациялық талдау, гранулометриялық талдау, генетикалық диаграммалар, Оңтүстік Құмкөл кен орны, жоғарғы Құмкөл кіші шөгінді қабаты.

¹*Bazarbay D.K.,

Master's student, ORCID ID: 0009-0001-2964-3523,

*e-mail: d_bazarbai@mail.ru

¹Kazakh-British Technical University, Almaty, Kazakhstan

STUDY OF RESERVOIR FORMATIONS OF THE UPPER KUMKOL SUBFORMATION OF THE SOUTH KUMKOL FIELD

Abstract

The relevance of this study is due to the need for a detailed study of recoverable oil reserves in reservoirs of productive horizons, which is associated with the conditions of sedimentation and deposit formation processes within the South Turgai sedimentary basin. A comprehensive analysis, including granulometric studies of terrigenous sediments, is required to accurately determine the conditions of rock formation. This approach allows to obtain reliable data on the depositional environment and to perform a more detailed facies analysis. The aim of the study is to determine the facies of sedimentation of the South Kumkol productive horizon Y-II using granulometric analysis. The methodology of the study includes application of various methods of granulometric analysis. The generalised Fuchtbauer and Muller definition of depositional setting was used as the main approach. Genetic interpretation of sediments is based on the K. Björlikke diagram analysing the ratio of sorting and asymmetry of particle distribution. The G.F. Rozhkov dynamogenetic diagram, which takes into account the ratio of asymmetry and excess, was used to clarify sedimentation conditions. The results of the study make it possible to characterise more objectively the conditions of reservoir rock formation and specify sedimentation parameters. An integrated approach to the analysis of granulometric data helps to improve the accuracy of facies diagnostics of oil-bearing horizons.

Keywords: facies analysis, granulometric analysis, genetic diagrams, South Kumkol field, Upper Kumkol Subformation.

Дата поступления статьи в редакцию: 28.05.2025

UDC 622.276.1
IRSTI 52.47.19

<https://doi.org/10.55452/1998-6688-2025-22-4-419-426>

¹**Liang Peng,**

Master's student, e-mail: 1539552876@qq.com

^{1,2*}**Hongbin Yang,**

Associate Professor, *e-mail: hongbinyang@upc.edu.cn

¹**Zixiang Yi,**

Undergraduate student, e-mail: 3297353308@qq.com

¹**Shuhe Zhang,**

Master's student, e-mail: 1009383847@qq.com

Xin Chen,

Master student, e-mail: xin_baobei77@163.com

¹**Haobin Shi,**

Master's student, e-mail: 2842117353@qq.com

¹**Yubin Zhang,**

Master's student, e-mail: 1940390250@qq.com

³**Wanli Kang,**

Professor, e-mail: kangwanli@126.com

¹Faculty of Petroleum Engineering, China University
of Petroleum (East China), Qingdao, PR China

²Shandong Key Laboratory of Oil and Gas Field Chemistry,
China University of Petroleum (East China), Qingdao, PR China

³Xin Jiang Keli New Technology Development Co. Ltd, Karamay, PR China

RESEARCH PROGRESS ON THRESHOLD PRESSURE GRADIENT IN HEAVY OIL RESERVOIRS

Abstract

Polymer flooding is one of the key technologies for enhancing oil recovery. Partially Hydrolyzed Polyacrylamide (HPAM) is widely used due to its excellent viscosity-increasing properties. However, the adsorption and retention behavior of HPAM in reservoir porous media presents a dual effect: on one hand, it improves sweep efficiency by increasing flow resistance; on the other hand, it leads to a loss in effective polymer concentration and viscosity, reducing displacement efficiency and increasing costs. Therefore, a systematic understanding and control of HPAM adsorption behavior are crucial for improving the effectiveness of polymer flooding. This work systematically reviews seven main measurement methods for HPAM adsorption quantity, comparing their applicable conditions and limitations. It summarizes the key factors influencing HPAM adsorption and retention behavior from three aspects: polymer properties, rock mineral characteristics, and reservoir environmental conditions. Furthermore, it outlines chemical anti-adsorption methods, represented by competitive adsorption and nanofilm protection, along with their mechanisms. Finally, future research directions are proposed, focusing on building adsorption prediction models, deepening the understanding of adsorption mechanisms under multi-field coupling conditions, and developing novel functional polymers with anti-adsorption capabilities.

Keywords: partially hydrolyzed polyacrylamide, adsorption retention, factors influencing adsorption, anti-adsorption, enhanced oil recovery.

Introduction

Low-permeability heavy oil reservoirs are challenging to develop due to high crude oil viscosity and poor flow capacity under reservoir conditions. A threshold pressure gradient may exist, meaning

heavy oil only begins to flow when the driving pressure gradient exceeds this threshold. During the development of heavy oil reservoirs, the threshold pressure gradient typically influences the process by increasing flow resistance, leading to distinct patterns of remaining oil distribution and ultimate recovery compared to conventional reservoirs [1]. This is one of the main reasons for low productivity and development difficulties in such reservoirs [2]. The low permeability of the reservoir matrix and high fluid flow resistance often result in excessively high injection pressures during water flooding, as well as phenomena such as inadequate water injection in wells [3–5]. The presence of a threshold pressure gradient accelerates the formation of dominant flow channels, reduces the sweep efficiency of water flooding, exacerbates the mobility difference between water and oil, and promotes fingering effects, significantly impacting the pressure and saturation fields [6]. Therefore, accurately determining the threshold pressure gradient is crucial for the rational development of heavy oil reservoirs.

1. Physical Significance of Threshold Pressure Gradient

The threshold pressure gradient is a key parameter in the development of low-permeability oil and gas reservoirs, influenced by both microscopic capillary forces and macroscopic stress sensitivity [7]. As early as 1951, B.A. Florin first proposed the concept of the threshold pressure gradient while studying issues related to dense mudstone and hard clay [8]. Since then, researchers both domestically and internationally have continued to investigate this phenomenon. The threshold pressure gradient refers to the additional pressure gradient required to overcome the resistance caused by adsorbed or hydration films on rock surfaces during fluid flow in low-permeability reservoirs. It plays a critical role in simulating fluid flow, particularly under conditions where fluid begins to move only when pressure accumulation is sufficient and shear stress exceeds the yield stress. Within the framework of classical seepage theory, the threshold pressure gradient depends on fluid properties, surface interactions with the medium, and pore structure. It becomes significant when the thickness of the boundary fluid in rock pores is comparable to the pore radius, resulting in the presence of a threshold pressure gradient during fluid flow through the porous media. The existence of the threshold pressure gradient serves as the theoretical basis for the effective development of low-permeability reservoirs and helps describe low-velocity non-Darcy seepage phenomena.

Ji et al. investigated the influence of the threshold pressure gradient on pressure propagation behavior around a single well [9]. The presence of the threshold pressure gradient substantially affects pressure distribution within the reservoir. Compared with conventional seepage flow, the threshold pressure gradient intensifies the pressure drop near the wellbore, resulting in a smaller and sharper “pressure depression funnel.” Moreover, at any given location within the drainage area, a higher threshold pressure gradient corresponds to a greater pressure gradient.

2. Measurement Methods for Threshold Pressure Gradient

Currently, methods for determining the threshold pressure gradient can be categorized into three major types: laboratory physical experiments, numerical experiments, and well-test interpretation [10–12]. Laboratory physical experiments include steady-state methods, unsteady-state methods, capillary equilibrium methods, bubble methods, among others. In laboratory settings, the threshold pressure gradient is typically measured by determining the relationship between pressure difference and flow rate under stable flow conditions, while in oilfield applications, it is commonly obtained through well-test analysis.

2.1 Laboratory Physical Experimental Methods

(1) Steady-State Method

The steady-state method primarily determines the threshold pressure gradient by directly regressing the linear segment of the seepage curve. For a given core sample, the pressure difference-flow rate method is employed: the flow rate through the core under different pressure differences is measured, with equilibrium pressure plotted on the horizontal axis and displacement velocity on the vertical axis in a Cartesian coordinate system. Using mathematical fitting, if the resulting curve is a straight line passing through the origin, it indicates the absence of a threshold pressure gradient in the core. If the fitted line has an intercept, it confirms the presence of a threshold pressure gradient, and the value of the intercept corresponds to the magnitude of the threshold pressure gradient [13].

(2) Unsteady-State Method

The unsteady-state method defines the threshold pressure gradient as the critical pressure gradient at which fluid transitions from flowing to non-flowing. The experimental procedure is designed as follows: initially, the core sample is saturated with crude oil under high pressure, with one end sealed and connected to a pressure gauge. After the system pressure reaches equilibrium and stabilizes, the other end of the core is depressurized to a specific pressure value. The pressure change at the sealed end is continuously monitored until the system re-stabilizes. This method offers advantages such as operational simplicity, ease of controlling experimental conditions, and short experimental duration.

(3) Capillary Equilibrium Method

The capillary equilibrium method is based on the principle of connected vessels. In this method, capillary tubes are connected to both ends of the core sample, and a height difference is created through gravitational effects. The measured height difference corresponds to the threshold pressure. The purpose of connecting capillaries to both ends of the holder is twofold: first, to sensitively and accurately reflect changes in liquid level, and second, to reduce the total seepage volume and shorten the measurement period. This approach not only confirms the existence of a threshold pressure gradient in low-permeability cores but also allows direct determination of its minimum value. During the measurement, both the capillary tubes and the core are filled with the experimental fluid, with the liquid level at the outlet end maintained lower than that at the inlet end to establish a constant pressure gradient. Driven by this pressure gradient, fluid flows through the core, causing the liquid level at the outlet to rise and that at the inlet to fall. If no threshold pressure gradient exists in the core, the liquid levels at both ends will equalize after sufficient equilibration. If a threshold pressure gradient is present, a height difference will remain, which represents the actual threshold gradient of the core [14, 15]

(4) Bubble Method

The bubble method involves immersing the outlet pipeline of the core sample into water. When liquid displaces the core, an additional pressure is applied to overcome both the resistance within the core and the interfacial tension between fluids, enabling the displacing fluid to enter the pore channels. Once the fluid begins to move, bubbles emerge at the outlet end of the core. The displacement pressure at this moment corresponds to the threshold pressure.

2.2 Numerical Experimental Methods

Mathematical models are used to indirectly determine the minimum threshold pressure gradient by incorporating various influencing factors and establishing mathematical equations. Based on capillary pressure curves or pore-throat distribution curves of core samples, a numerical structure of the porous medium is constructed using the Lattice Boltzmann (LB) model. The LB method is then employed for simulation to generate a corresponding curve. By extending the linear segment of this simulated curve until it intersects the pressure gradient axis, the value at the intersection point represents the dimensionless threshold pressure gradient. A scaling factor between dimensionless and dimensional quantities is derived by comparing simulation results with actual field pressure gradient data, enabling the determination of the threshold pressure gradient for the target formation [16]. Yun proposed an irregular model for the threshold pressure gradient of Bingham fluids in porous media, based on pore properties and capillary pressure effects [17]. Each parameter in this model has a clear physical meaning, and it establishes relationships between the structural parameters of the porous medium, yield stress, capillary pressure parameters, the fractal dimension of the porous medium, and the threshold pressure gradient of Bingham fluids.

2.3 Well-Test Interpretation Methods

Well-test interpretation methods analyze actual field data to dynamically and in situ determine the threshold pressure gradient [18]. Based on previous research experience, Liu et al. proposed a well-test interpretation approach to address the threshold pressure gradient issue by developing models incorporating control equations, flow velocity equations, and inner/outer boundary conditions distinct from those used in conventional reservoirs [19]. A comparison reveals the key difference between conventional and low-permeability reservoir well-test models: in conventional reservoirs,

fluid flow occurs whenever a pressure gradient exists, whereas in low-permeability reservoirs, flow only initiates when the pressure gradient exceeds the threshold pressure gradient. This distinction is reflected in theoretical curves through an upward deflection in the pressure derivative curve.

2.4 Comparative Analysis of the Three Methods

Laboratory simulation methods provide an intuitive and direct approach for studying seepage mechanisms and are currently widely recognized for determining the threshold pressure gradient. However, this approach faces two main challenges: extended experimental durations and the difficulty of maintaining cores in their natural underground state due to stress release. Although the underlying seepage mechanisms remain valid, these factors may introduce significant data errors.

Numerical experimental methods offer advantages such as simplicity, speed, and the ability to explore a wide range of parameters. However, they rely heavily on the pore-throat distribution characteristics of the formation and require comparative experimental data for validation. When combined with laboratory simulation methods, numerical approaches can yield results that align closely with comparative experiments. This makes numerical experimentation a promising auxiliary tool for enhancing laboratory-based physical simulation studies of reservoirs.

Well-test interpretation methods dynamically reflect reservoir behavior, making the threshold pressure gradient determined through this approach highly relevant for practical applications. However, a drawback of this method is that field operations are time-consuming and relatively costly.

3. Factors Influencing Threshold Pressure Gradient in Heavy Oil

Heavy oil only begins to flow when the driving pressure gradient exceeds the threshold pressure gradient. While accurately measuring the threshold pressure gradient is essential, its influencing factors cannot be overlooked. Current research indicates that crude oil viscosity, permeability, and water saturation significantly affect the threshold pressure gradient. Experimental studies have also demonstrated the influence of other factors such as wettability and confining pressure.

3.1 Crude Oil Viscosity

Viscosity is one of the key characteristic parameters of heavy oil. The variable nature of heavy oil viscosity is a primary reason for its classification as a non-Newtonian fluid, and as such, it exhibits specific flow behaviors typical of non-Newtonian fluids. Sun suggested that the viscosity of heavy oil is mainly influenced by the content of non-hydrocarbon components (e.g., resins) and asphaltenes [20]. Asphaltenes, being the most polar components in heavy oil, determine the molecular polarity of heavy oil and affect both liquid-liquid and solid-liquid interfacial tension, thereby influencing the shear stress of heavy oil. During seepage through porous media, the shear stress of heavy oil impacts its threshold pressure gradient. Under porous media conditions, heavy oil exhibits shear-thinning behavior, and the transition point can be used to determine the critical threshold pressure gradient. As the content of resins and asphaltenes increases, the threshold pressure gradient also increases. Cao et al. demonstrated that crude oil viscosity significantly affects the threshold pressure gradient [21]. The threshold pressure gradient is negatively correlated with crude oil viscosity, meaning it decreases as the viscosity increases.

3.2 Permeability

Ke et al. demonstrated that permeability significantly influences the threshold pressure gradient [22]. The threshold pressure gradient is positively correlated with permeability, meaning it decreases as permeability increases. Bai et al. reported similar findings in their research [23]. As permeability increases, the threshold pressure gradient of the core decreases accordingly. This is attributed to the narrow pores and throats in tight reservoirs, which result in significantly higher seepage resistance during fluid flow compared to conventional reservoirs, leading to a larger threshold pressure gradient. When permeability falls below a certain critical value, the threshold pressure gradient decreases sharply.

3.3 Water Saturation

For low-permeability cores, the oil phase remains relatively continuous under irreducible water conditions. After water flooding for a period, water saturation increases, and some oil is displaced. The oil phase in the core becomes disconnected by the water phase, existing as oil droplets or

columns within the pores. As a result, additional resistance is generated when oil begins to flow. Higher water saturation leads to more dispersed oil droplets, increasing this additional resistance. Therefore, the threshold pressure gradient in low-permeability cores increases with rising water saturation. In contrast, for medium-permeability cores, increasing water saturation causes the water phase to gradually become continuous, forming dominant flow channels. Due to the relatively larger pore radii, the additional resistance caused by oil droplet mobilization has less impact. Thus, the threshold pressure gradient decreases with increasing water saturation [24, 25].

Zou et al. studied the two-phase threshold pressure gradient in cores and found that, depending on the strength of capillary effects, the two-phase threshold pressure gradient first increases and then decreases with increasing average water saturation [26].

3.4 Other Factors

Zhang et al. verified that the threshold pressure gradient of core samples decreases as the relative wettability index increases [27]. The stronger the hydrophilicity of the core, the lower the threshold pressure gradient. Additionally, the threshold pressure gradient exhibits a strong linear correlation with the thickness of the boundary layer liquid: as the boundary layer thickness increases, the threshold pressure gradient gradually rises.

Wang et al. suggested that enhanced oil-wettability of rock samples strengthens the molecular interactions at the oil-solid interface, ultimately leading to an increase in the thickness of the crude oil boundary layer and a higher threshold pressure gradient [28]. In other words, greater oil-wettability of the rock sample corresponds to an increase in the threshold pressure gradient.

Tian et al. investigated the effects of confining pressure and clay properties on the threshold pressure gradient [29]. They found that the threshold pressure gradient increases with rising net confining pressure, though the rate of increase gradually diminishes. The presence of microfractures reduces the threshold pressure gradient. An increase in the volume fraction of expansive clay worsens seepage efficiency. Additionally, an increase in water saturation causes the threshold pressure gradient to first rise and then decline.

Xu et al. studied the influence of irreducible water on the threshold pressure gradient [30]. Experimental results showed that the measured threshold pressure gradient in rock samples containing irreducible water is higher than in those without irreducible water.

Xiong et al. investigated the influence of dominant throat radius and movable fluid saturation on the threshold pressure gradient [31]. They found that the threshold pressure gradient gradually decreases as the dominant throat radius increases. Similarly, a higher movable fluid saturation corresponds to a lower threshold pressure gradient. This is because lower movable fluid saturation leads to a thicker interfacial layer, significantly reducing the effective seepage space in low-permeability reservoirs and resulting in a higher threshold pressure.

Conclusion

This study summarizes the measurement methods for the threshold pressure gradient, including laboratory physical experiments, numerical simulations, and well-test interpretation. It compares the advantages and disadvantages of these different approaches to reduce measurement errors and improve accuracy. A systematic summary of the influencing factors on the threshold pressure gradient—such as crude oil viscosity, permeability, and water saturation—is provided. The research findings offer theoretical support and a foundation for experiments on the threshold pressure gradient in heavy oil.

REFERENCES

- 1 Boming, J., Langfeng, M., Yiqun, Y., et al. Calculation method of shale oil fractured horizontal well productivity based on non-Darcy and pressure sensitive features [J]. *Petroleum Geology and Development of Daqing*, 42(02), 152–159 (2023).

- 2 Ziqiang, W., Daiyan, Z., Jun, Y., et al. Experimentals study on percolation characteristics of ordinary heavy oil reservoirs: a case study from the Qigu Formation reservoir in Nine-4 Block of Xinjiang oilfield [J]. *Oil & Gas Geology*, 33(02), 302–306 (2012).
- 3 Zhihao, Q., Lingrong, K. A study of characteristics of low permeability formation with water flooding using micromodel experiment [J]. *Journal of Northwest University (Natural Science Edition)*, (04), 329–334 (2002).
- 4 Tongluo, Q. How to Develop the Low Permeable Reservoir [J]. *Fault-Block Oil & Gas Fields*, (03), 21–23 (1994).
- 5 Yongming, L., Wenwu, Z., Jinzhou, Z., et al. Semi-analytical productivity calculation and sensitive factors for the multi-stage fractured horizontal well in low permeability reservoirs [J]. *Reservoir Evaluation and Development*, 8(2), 52–57 (2018).
- 6 Weiyang, W., Shan, T., Wentao, T., et al. Nonlinear Percolation of Heavy Oil and Its Impact on Oil Recovery Factors [J]. *Journal of Oil and Gas Technology*, 32(05), 115–404 (2010).
- 7 Peng, Y., Jun, W., Zhijun, Z., et al. Analysis of low permeability gas reservoir fracturing well productivity equation considering multi-factors [J]. *Reservoir Evaluation and Development*, 8(1), 34–37 (2018).
- 8 Gorbunov, A.T. Development of Abnormal Oilfields [M]. Translated by Shubao Z. Beijing: Petroleum Industry Press, 1987.
- 9 Bingyu, J., Yingfu, H. Formation pressure distribution of a single well based on low-velocity non-Darcy flow [J]. *Acta Petrolei Sinica*, 32(03), 466–469 (2011).
- 10 Yizhong, Z., Yuanfang, C., Yuchuan, L., et al. Study on influence of start-up pressure gradient to micro-seepage in low permeability reservoirs and development trends [J]. *Petroleum Geology and Recovery Efficiency*, 20(01), 67–73, 115 (2013).
- 11 Zhengming, Y. Porous flow mechanics for low permeability reservoir sand its application [D]. Graduate School of Chinese Academy of Sciences (Institute of Porous Flow and Fluid Mechanics), 2005.
- 12 Quanlin, W., Hai, T., Dongliang, L., et al. An experimental study on threshold pressure gradient in low permeability reservoir [J]. *Petroleum Geology and Recovery Efficiency*, 18(01), 97–118 (2011).
- 13 Bin, L., Zhiwei, X., Mingming, G., et al. Experimental Study on Start-up Pressure Gradient in Low Permeability Reservoirs [J]. *Guangdong Chemical Industry*, 51(15), 21–25 (2024).
- 14 Yang, W., Shenglai, Y., Bin, W., et al. Experimental research of single phase water on threshold pressure gradient of ultra-low permeability sandstone cores of Daqing Oilfield [J]. *Complex Hydrocarbon Reservoirs*, 3(01), 62–65 (2010).
- 15 Yongshou, L., Hua, Z., Yujie, Y., et al. An Experimental study on startiong pressure gradient in ultra-low permeability reservoirs [J]. *Xinjiang Oil & Gas*, 8(04), 81–84+1 (2012).
- 16 Hua, Z., Weiqiang, L., Yunjun, Z., et al. The Advance of Research on Starting Pressure Gradient in Low-Permeability Sandstone Flow [J]. *Inner Mongolia Petrochemical Industry*, 36(14), 100–102 (2010).
- 17 Yun, M., Yu, B., Cai, J. A fractal model for the starting pressure gradient for Bingham fluids in porous media [J]. *International Journal of Heat and Mass Transfer*, 51(5), 1402–1408 (2007).
- 18 Huiling, G., Yong, T., Guoquan, Q. Productivity calculation of fractured horizontal wells in low permeability gas reservoir by the consideration of threshold pressure gradients [J]. *Reservoir Evaluation and Development*, 5(4), 29–32 (2015).
- 19 Yuewu, L., Zhenhua, D., Fengzhen, H. Three Kinds of Methods for Determining the Start-up Pressure Gradients in Low Permeability Reservoir [J]. *Well Testing*, (04), 1–4+74 (2002).
- 20 Jianfang, S. Research and Application on the Heavy Oil Flow Mode [D]. China University of Geosciences (Beijing), 2012.
- 21 Liying, C., Chuanxi, L., Jianfang, S., et al. The Starting Pressure Gradient Research of Low Permeability Carbonate Heavy Oil Reservoir [J]. *Science Technology and Engineering*, 15(01), 215–219 (2015).
- 22 Wenli, K., Gaoming, Y., Wencheng, Z., et al. Experimental study on start-up pressure gradient for nonlinear flow of heavy oil [J]. *Oil Drilling & Production Technology*, 38(03), 341–346 (2016).
- 23 Huifang, B., Liyu, S., Lei, Z., et al. The Actuating Pressure Gradient Experimental Study of Tighe Sandstone Gas Reservoir in Ordos Basin [J]. *Unconventional Oil & Gas*, 7(03), 60–64+54 (2020).
- 24 Ying, L., Aifen, L., Min, M., et al. Experimental study on gas starting pressure and movable water saturation in low permeability water bearing gas [C]. *Proceedings of the 2021 International Conference on Oil and Gas Field Exploration and Development*, 2–9 (2021).

- 25 Ga, Y., Hai, T., Dongliang, L. The study and analysis of starting pressure gradient in low permeability gas reservoirs [J]. Offshore Oil, (03), 51–54 (2006).
- 26 Jian, Z., Fayuan, Z., Mingda, D., et al. Change Law of Two-phase Threshold Pressure Gradient in Tight-low Permeability Reservoir [J]. Science Technology and Engineering, 20(28), 11562–11567 (2020).
- 27 Guohui, Z., Xinrong, L., Futang, L., et al. Experiment Research on Starting Pressure Gradient of Oil-Water Two Phase Flow in Cores [J]. Science Technology and Engineering, 24(13), 5372–5380 (2024).
- 28 Youfu, W., Gang, W., Shukai, A., et al. Experimental Study on Influence Factors of Start-up Pressure Gradient of Permeability Reservoir Rocks [J]. Journal of Oil and Gas Technology (Journal of Jiangnan Petroleum Institute), (03), 112–113+446 (2006).
- 29 Wei, T., Weiyao, Z., Huayin, Z., et al. Influencing factors of threshold pressure gradient value for tight sandstone [J]. Fault-Block Oil & Gas Fields, 21(05), 611–614 (2014).
- 30 Kun, X., Jing, L., Chunlei, T., et al. Study on Influencing Factors of Threshold Pressure Gradient [J]. China Petroleum and Chemical Standard and Quality, 34(12), 252 (2014).
- 31 Wei, X., Qun, L., Xiangui, L., et al. Pseudo threshold pressure gradient to flow for low permeability reservoirs [J]. Petroleum Exploration and Development, 36(02), 232–236 (2009).

¹Лян Пэн,

магистрант, e-mail: 1539552876@qq.com

^{1,2*}Хунбин Ян,

доцент, *e-mail: hongbinyang@upc.edu.cn

¹Цзисян И,

бакалавриат студенті, e-mail: 3297353308@qq.com

¹Шухэ Чжан,

магистрант, e-mail: 1009383847@qq.com

¹Синь Чен,

магистрант, e-mail: xin_baobei77@163.com

¹Хаобин Ши,

магистрант, e-mail: 2842117353@qq.com

¹Юйбинь Чжан,

магистрант, e-mail: 1940390250@qq.com

³Ваньли Кан,

профессор, e-mail: kangwanli@126.com

¹Мұнай инженериясы факультеті, Қытай мұнай университеті (Шығыс Қытай),
Циндао қ., ҚХР

²Шаньдун провинциясының мұнай және газ кен орындарының химиясы жөніндегі негізгі
зертханасы, Қытай мұнай университеті (Шығыс Қытай), Циндао қ., ҚХР

³Xin Jiang Keli New Technology Development Co. Ltd, Қарамай қ., ҚХР

АУЫР МҰНАЙ КОЛЛЕКТОРЛАРЫНДАҒЫ ҚЫСЫМНЫҢ ШЕКТІ ГРАДИЕНТІ ЖӨНІНДЕГІ ЗЕРТТЕУЛЕРДІҢ ПРОГРЕСІ

Аңдатпа

Ауыр мұнай құрамындағы шайырлар мен асфальттер сияқты жоғары молекулалық қосындылар оның бірегей құрылымдық қасиеттерін, яғни сызықты емес сүзілу сипаттамаларын айқындайды. Қысым градиентінің табалдырық мәні (threshold pressure gradient) – Дарси емес ағынды зерттеудегі маңызды параметр. Бұл мақалада ауыр мұнай коллекторларындағы қысым градиентінің табалдырық мәнін зерттеу саласындағы жетістіктерге жан-жақты шолу жасалған. Онда қысым градиентінің табалдырық мәнін анықтаудың негізгі әдістері – зертханалық физикалық эксперименттер, сандық модельдеу және ұңғымаларды сынау нәтижелерін талдау – қарастырылып, әр тәсілдің артықшылықтары мен шектеулері салыстырылған. Сонымен қатар, шикі мұнайдың тұтқырлығы, өткізгіштігі және суға қанығуы сияқты негізгі әсер етуші факторлар талданады. Зерттеу нәтижелері қысым градиентінің табалдырық мәні, әдетте, шикі мұнайдың тұтқырлығы артқан сайын өсетінін, төмен өткізгіштігі бар керндерде суға қанығу артқан сайын көбейетінін, ал өткізгіштік артқан сайын азаятындығын көрсетеді. Бұл зерттеу әсер етуші факторлар мен өлшеу әдістеріне қатысты жүйелі

қорытындылар ұсына отырып, өлшеу қателіктерін барынша азайтуға және ауыр мұнай коллекторларындағы айдау мен өндіру процестерін оңтайландыруға теориялық негіз береді.

Тірек сөздер: ауыр мұнай, Дарси емес ағыс, қысым градиентінің табалдырық мәні, сынау әдістері, әсер етуші факторлар.

¹Лян Пэн,
магистрант, e-mail: 1539552876@qq.com
^{1,2*}Хунбин Ян,
доцент,*e-mail: hongbinyang@upc.edu.cn
¹Цзисян И,
бакалавр, e-mail: 3297353308@qq.com
¹Шухэ Чжан,
магистрант, e-mail: 1009383847@qq.com
¹Синь Чен,
магистрант, e-mail: xin_baobei77@163.com
¹Хаобин Ши,
магистрант, e-mail: 2842117353@qq.com
¹Юйбинь Чжан,
магистрант, e-mail: 1940390250@qq.com
³Ваньли Кан,
профессор, e-mail: kangwanli@126.com

¹Факультет нефтяной инженерии, Китайский нефтяной университет
(Восточный Китай), г. Циндао, КНР

²Ключевая лаборатория химии нефтяных и газовых месторождений провинции Шаньдун,
Китайский нефтяной университет (Восточный Китай), г. Циндао, КНР

³Xin Jiang Keli New Technology Development Co. Ltd, г. Карамай, КНР

ПРОГРЕСС ИССЛЕДОВАНИЙ ПО ПОРОГОВОМУ ГРАДИЕНТУ ДАВЛЕНИЯ В КОЛЛЕКТОРАХ ТЯЖЕЛОЙ НЕФТИ

Аннотация

Высокомолекулярные смеси, такие как смолы и асфальтены в тяжелой нефти, определяют ее уникальные структурные свойства, проявляющиеся в характеристиках нелинейного просачивания. Пороговый градиент давления является критическим параметром при изучении недарсиевского течения. В данной статье представлен всесторонний обзор прогресса исследований погрешности градиента давления в коллекторах тяжелой нефти. В ней обобщены основные методы определения порогового градиента давления, включая лабораторные физические эксперименты, численное моделирование и интерпретацию результатов испытаний скважин, а также сравниваются преимущества и ограничения каждого подхода. Кроме того, в ней анализируются ключевые факторы влияния, такие как вязкость сырой нефти, проницаемость и водонасыщенность. Результаты показывают, что пороговый градиент давления, как правило, увеличивается с ростом вязкости сырой нефти, увеличивается с ростом водонасыщенности в низкопроницаемых ядрах, но уменьшается с ростом проницаемости. В данном исследовании представлены систематические выводы относительно этих факторов влияния и методов измерения, направленные на минимизацию погрешностей измерений и предоставление теоретической основы для оптимизации процессов закачки и добычи в коллекторах тяжелой нефти.

Ключевые слова: тяжелая нефть, недарсиевский поток, пороговый градиент давления, методы испытаний, влияющие факторы.

Article submission date: 10.10.2025

УДК 550.8.05
МРНТИ 38.49.15

<https://doi.org/10.55452/1998-6688-2025-22-4-427-431>

¹*Болтанов А.К.,
магистр, ORCID ID: 0009-0003-7908-8008,
*e-mail: a_boltanov@kbtu.kz

¹Тулемисова Ж.С.,
PhD, ассоциированный профессор, ORCID ID: 0000-0003-1803-4535,
e-mail: z.tulemissova@kbtu.kz

¹Казахстанско-Британский технический университет, г. Алматы, Казахстан

ПЕРСПЕКТИВЫ ВЫЯВЛЕНИЯ НОВЫХ МЕСТОРОЖДЕНИЙ ПОЛЕЗНЫХ ИСКОПАЕМЫХ НА ЛИСАКОВСКОЙ ПЛОЩАДИ ПО РЕЗУЛЬТАТАМ ГЕОЛОГИЧЕСКОГО ДОИЗУЧЕНИЯ

Аннотация

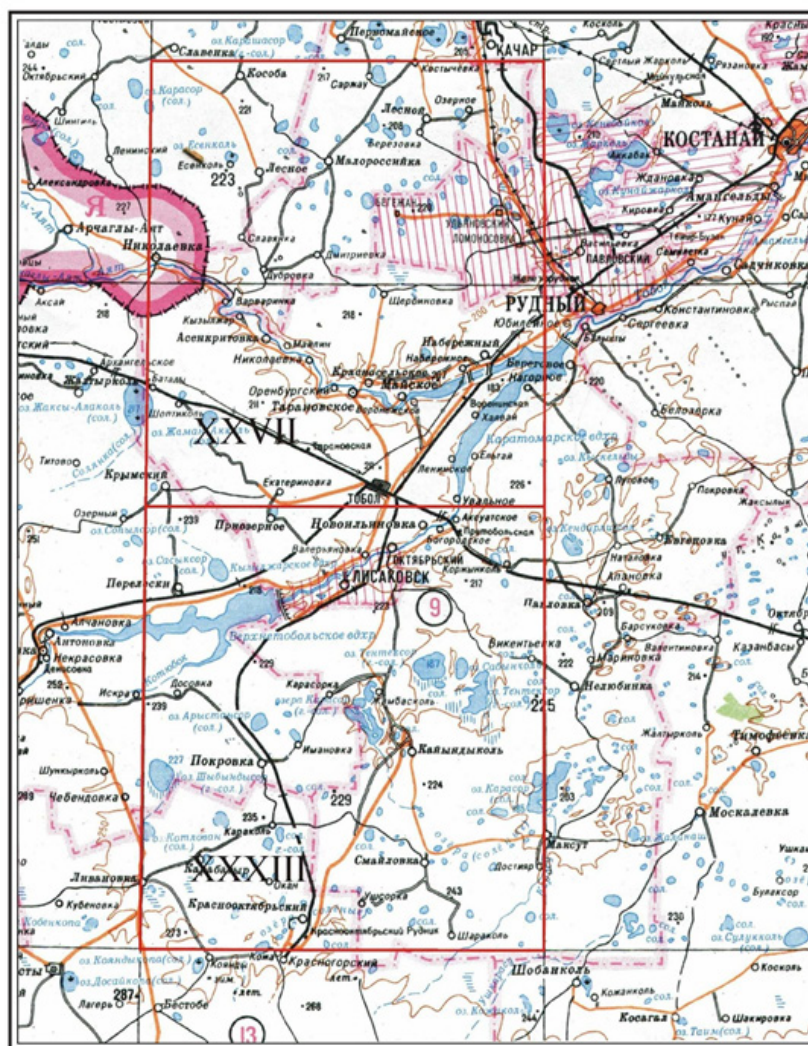
В статье представлены результаты геологического доизучения Лисаковской площади (листы N-41-XXVII и N-41-XXXIII), позволившего существенно уточнить геологическое строение и перспективы рудоносности этой важной в металлогеническом отношении территории. На основе анализа и переинтерпретации большого объема фондовых и архивных материалов, а также результатов собственных полевых исследований уточнена стратиграфическая схема района. Впервые в Денисовской структурно-формационной зоне выделен комплекс метаморфических пород докембрия – нижнего палеозоя. Уточнено стратиграфическое положение лисаковской толщи оолитовых железных руд. Создана цифровая модель покровного комплекса, позволившая составить погоризонтные карты всех стратиграфических подразделений. Получены новые данные о тектоническом строении зоны сочленения Денисовской и Валерьяновской структурно-формационных зон. Предложена новая геодинамическая модель формирования этой структуры. Проведено минерагеническое районирование территории с выделением перспективных площадей на различные виды полезных ископаемых. Выделены минерагенические зоны, районы, рудные узлы и рудные поля, перспективные на железо, медь, полиметаллы, золото, серебро, бокситы. Рекомендованы участки для постановки поисковых работ на золото-медно-порфировый, колчеданный полиметаллический, золото-кварцево-жильный типы оруденения. К первоочередным отнесены Северный и Алакольский-1 участки. Полученные результаты создают надежную основу для дальнейшего изучения территории и проведения целенаправленных геолого-разведочных работ на выявленных перспективных участках.

Ключевые слова: Лисаковская площадь, геологическое доизучение, стратиграфия, тектоника, минерагеническое районирование, перспективные участки, прогноз полезных ископаемых.

Введение

Лисаковская площадь располагается в зоне сочленения складчатых структур Урала и Казахстана и характеризуется исключительно сложным геологическим строением и многообразной историей развития. Территория имеет важное значение с точки зрения минерально-сырьевой базы республики, здесь расположены крупные месторождения железных руд и бокситов. В то же время геологическое строение района изучено недостаточно детально, что ограничивает возможности прогнозирования новых месторождений.

Целью проведенных работ являлось геологическое доизучение Лисаковской площади на основе комплексного анализа всех имеющихся геолого-геофизических материалов и результатов собственных полевых исследований. Основными задачами были уточнение стратиграфической схемы, тектонического строения, выделение перспективных площадей на различные виды полезных ископаемых.



Границы исследуемой территории

Рисунок 1 – Обзорная карта исследуемой территории.
Масштаб 1: 1 000 000

Материалы и методы

В основу работы положены результаты анализа и переинтерпретации большого объема фондовых и архивных материалов по геологической съемке, поисковым и разведочным работам, накопленных за последние 40 лет. Кроме того, были использованы данные собственных полевых исследований – маршрутные наблюдения, детальное изучение обнажений и карьеров, передокументация керна глубоких скважин.

Для создания цифровой модели покровного комплекса использовались данные бурения. Проведена переинтерпретация геофизических материалов – составлены детальные карты магнитного и гравитационного полей. Построены новые геологические карты масштаба 1:200 000.

Для прогнозирования перспективных участков применялись методы минерагенического анализа и математического моделирования.

Результаты и обсуждение

В результате проведенных работ существенно уточнена стратиграфическая схема района. Впервые в Денисовской структурно-формационной зоне выделен комплекс метаморфических пород докембрия –нижнего палеозоя, ранее относившийся к эйфельскому ярусу среднего девона. В Денисовской СФЗ выделены отложения нижнего ордовика, эмсского яруса нижнего девона. Уточнено стратиграфическое положение лисаковской толщи оолитовых железных руд.

Создана цифровая модель покровного комплекса, позволившая составить погоризонтные карты всех стратиграфических подразделений. Составлены детальные карты геофизических полей, что позволило существенно уточнить геологическое строение территории.

Получены принципиально новые данные о тектоническом строении и истории геологического развития зоны сочленения Денисовской и Валерьяновской структурно-формационных зон. Предложена новая геодинамическая модель формирования этой структуры, учитывающая процессы рифтогенеза, субдукции и коллизии.

Проведено минерагеническое районирование территории с выделением перспективных площадей на различные виды полезных ископаемых. Выделены минерагенические зоны, районы, рудные узлы и рудные поля, перспективные на железо, медь, полиметаллы, золото, серебро, бокситы. Определены закономерности размещения оруденения и факторы, контролирующие локализацию рудных тел.

Рекомендованы перспективные участки для постановки поисковых работ на золото-медно-порфировый, колчеданный полиметаллический, золото-кварцево-жильный типы оруденения. Участки ранжированы по очередности проведения поисковых работ. К первоочередным отнесены Северный (на золото и медь) и Алакольский-1 (на золото).

На Северном участке выявлена зона метасоматически измененных пород с прожилково-вкрапленной минерализацией. Содержания золота в отдельных пробах достигают 5 г/т, меди – 0.5%. Размеры перспективной зоны – 2х1 км. Рекомендуется проведение детальных поисковых работ с бурением скважин.

На участке Алакольский-1 установлена серия кварцевых жил с видимым золотом. Содержания золота в штуфных пробах до 15 г/т. Протяженность рудоносной зоны около 1 км при ширине 100–150 м. Рекомендуется проведение канавных работ и бурение поисковых скважин.

Заключение

Проведенные исследования позволили существенно уточнить геологическое строение и перспективы рудоносности Лисаковской площади – важного горнорудного района Казахстана, являющегося крупной минерально-сырьевой базой железных руд и бокситов.

Полученные результаты создают надежную основу для дальнейшего изучения территории и проведения целенаправленных геолого-разведочных работ на выявленных перспективных участках. Это будет способствовать укреплению и расширению минерально-сырьевой базы горнодобывающей промышленности Казахстана.

Рекомендуется продолжить геолого-разведочные работы на выделенных перспективных участках с применением комплекса современных геолого-геофизических методов, включая геохимические поиски по вторичным ореолам рассеяния, наземные геофизические исследования методами магниторазведки и электроразведки, а также поисковое и оценочное бурение.

ЛИТЕРАТУРА

1 Безнедельный Б.Д. Обобщение геолого-геофизических материалов в масштабе 1:50000 с врезками м-ба 1:25000 по северной части Западно-Тургайского бокситорудного района // Отчет. – Кустанай, 1980. – 245 с.

- 2 Бelyh M.P. Отчет о результатах доразведки Белинского месторождения бокситов в 1974–76 гг. // Дополнение к отчету о доразведке месторождения за период 1965–69 гг. – Темир, 1976. – 180 с.
- 3 Внукова Н.Т. Изучение закономерностей размещения геолого-технологических сортов бокситов по рудным телам 53, 57, 58, 59 Белинского месторождения и рудному телу 46 Аятского месторождения // Отчет. – Октябрьский, 1985. – 210 с.
- 4 Дербенко П.Т. Изучение закономерностей размещения типов бокситов на Белинском месторождении // Отчет по теме № 22. – Алма-Ата, 1978. – 156 с.
- 5 Заседателев И.Н. Отчет о результатах инженерно-геологической съемки масштаба 1:200000 на массиве орошения в пределах листов: N-41-XV11, N-41-XX11, N-41-XX111, N-41-XXV111, N-41-XXX111, N-41-XXX1V, N41-XXXV за 1977-1978 гг. // Отчет. – Затобольск, 1981. – 1957 с.
- 6 Зубрицкий В.В. Оперативный подсчет запасов бокситов Аятской группы месторождений по состоянию на 01.01.1979 г. // Отчет. – Павловский, 1978. – 120 с.

REFERENCES

- 1 Beznedelnyj, B.D. Generalization of geological and geophysical materials on a scale of 1:50000 with inserts of 1:25000 scale for the northern part of the West Turgai bauxite ore region. Report. Kostanay, 1980. 245 p. [in Russian].
- 2 Belyh, M.P. Report on the results of additional exploration of the Belinsky bauxite deposit in 1974-76. Addition to the report on additional exploration of the deposit for the period 1965-69. Temir, 1976. 180 p. [in Russian].
- 3 Vnukova, N.T. Study of the patterns of distribution of geological and technological grades of bauxite in ore bodies 53, 57, 58, 59 of the Belinsky deposit and 46 of the Ayat deposit. Report. Oktyabrsky, 1985. 210 p. [in Russian].
- 4 Derbenko, P.T. Study of the patterns of distribution of bauxite types at the Belinsky deposit. Report on topic No. 22. Alma-Ata, 1978. 156 p. [in Russian].
- 5 Zasedatelev, I.N. Report on the results of engineering-geological survey on a scale of 1:200000 on the irrigation massif within the sheets: N-41-XV11, N-41-XX11, N-41-XX111, N-41-XXV111, N-41-XXX111, N-41-XXX1V, N41-XXXV for 1977-1978. Report. Zatulbolsk, 1981. 1957 p. [in Russian].
- 6 Zubrickij, V.V. Operational calculation of bauxite reserves of the Ayat group of deposits as of 01.01.1979. Report. Pavlovsky, 1978. 120 p. [in Russian].

¹*Болтанов А.К.,

магистр, ORCID ID: 0009-0003-7908-8008,

*e-mail: a_boltanov@kbtu.kz

¹Түлемисова Ж.С.,

PhD, қауымдастырылған профессор, ORCID ID: 0000-0003-1803-4535,

e-mail: z.tulemissova@kbtu.kz

¹Қазақстан-Британ техникалық университеті, Алматы қ., Қазақстан

ГЕОЛОГИЯЛЫҚ ҚОСЫМША ЗЕРТТЕУ НӘТИЖЕЛЕРІ БОЙЫНША ЛИСАКОВ АЛАҢЫНДА ЖАҢА ПАЙДАЛЫ ҚАЗБА КЕН ОРЫНДАРЫН АНЫҚТАУ ПЕРСПЕКТИВАЛАРЫ

Аңдатпа

Мақалада Лисаков алаңында (N-41-XXVII және N-41-XXXIII парақтары) жүргізілген геологиялық қосымша зерттеу нәтижелері ұсынылған. Бұл зерттеулер металлогендік тұрғыдан маңызды аумақтың геологиялық құрылымы мен кенділік перспективаларын нақтылауға мүмкіндік берді. Қор және мұрағаттық материалдардың үлкен көлемін талдау мен қайта интерпретациялау, сондай-ақ автордың далалық зерттеу нәтижелері негізінде ауданның стратиграфиялық схемасы нақтыланды. Денисов құрылымдық-формациялық аймағында алғаш рет докембрий мен төменгі палеозойдың метаморфтық жыныстар кешені анықталды.

Оолитті темір кендерінің Лисаков қабатының стратиграфиялық жағдайы нақтыланды. Жабын кешеннің сандық моделі жасалып, барлық стратиграфиялық бөлімшелер үшін горизонт карталарын құрастыруға мүмкіндік берілді. Денисов және Валерьянов құрылымдық-формациялық аймақтарының түйісу аймағының тектоникалық құрылымы туралы жаңа деректер алынды. Құрылымның қалыптасуына арналған жаңа геодинамикалық модель ұсынылды. Пайдалы қазбалардың әртүрлі түрлеріне перспективалы аудандарды бөліп, аумақтың минерагендік аудандастырылуы жүргізілді. Темір, мыс, полиметалдар, алтын, күміс және боксит бойынша перспективалы минерагендік аймақтар, аудандар, кен түйіндері мен кен алаңдары анықталды. Алтын-мыс-порфирлі, колчеданды-полиметалды, алтын-кварц-желілік кендену түрлерін іздестіру жұмыстарына арналған учаскелер ұсынылды. Солтүстік және Алакөл-1 учаскелері басым бағыттар ретінде белгіленді. Алынған нәтижелер болашақта аумақты кешенді зерттеу мен перспективалы учаскелерде барлау жұмыстарын тиімді ұйымдастыруға негіз болады.

Тірек сөздер: Лисаков алаңы, геологиялық қосымша зерттеу, стратиграфия, тектоника, минерагендік аудандастыру, перспективалы учаскелер, пайдалы қазбаларды болжау.

^{1*}**Boltanov A.K.,**

MSc, ORCID ID: 0009-0003-7908-8008,

*e-mail: a_boltanov@kbtu.kz

¹**Tulemisova Zh. S.,**

PhD, Associate Professor, ORCID ID: 0000-0003-1803-4535,

e-mail: z.tulemisova@kbtu.kz

¹Kazakh-British Technical University, Almaty, Kazakhstan

POTENTIAL FOR NEW DEPOSITS IN THE LISAKOVSK AREA BASED ON THE RESULTS OF EXTENDED GEOLOGICAL EXPLORATION

Abstract

The article presents the results of geological additional study of the Lisakovsk area (sheets N-41-XXVII and N-41-XXXIII), which made it possible to significantly clarify the geological structure and prospects of ore content of this important metallogenic territory. Based on the analysis and reinterpretation of a large volume of stock and archival materials, as well as the results of our own field studies, the stratigraphic scheme of the area has been refined. For the first time in the Denisov structural-formational zone, a complex of metamorphic rocks of the Precambrian-Lower Paleozoic has been identified. The stratigraphic position of the Lisakovsk stratum of oolitic iron ores has been clarified. A digital model of the cover complex has been created, allowing for the compilation of horizon maps for all stratigraphic units. New data on the tectonic structure of the junction zone of the Denisov and Valeryanovsk structural-formational zones have been obtained. A new geodynamic model for the formation of this structure has been proposed. Minerogenic zoning of the territory was carried out with the identification of promising areas for various types of minerals. Minerogenic zones, districts, ore nodes and ore fields prospective for iron, copper, polymetals, gold, silver, and bauxite have been identified. Areas are recommended for prospecting for gold-copper-porphyry, pyrite-polymetallic, gold-quartz-vein types of mineralization. The Northern and Alakolsky-1 areas are designated as priority sites. The obtained results provide a reliable basis for further study of the territory and targeted exploration work in the identified promising areas.

Keywords: Lisakovsk area, geological additional study, stratigraphy, tectonics, minerogenic zoning, prospective areas, mineral forecast.

Дата поступления: 28.04.2025

UDC 622.276.6
IRSTI 52.47.27

<https://doi.org/10.55452/1998-6688-2025-22-4-432-443>

¹**Haobin Shi,**

Master's student, e-mail: 2842117353@qq.com

^{1,2*}**Hongbin Yang,**

Associate Professor, *e-mail: hongbinyang@upc.edu.cn

¹**Xin Chen,**

Master's student, e-mail: xin_baobei77@163.com

¹**Liang Peng,**

Master's student, e-mail: 1539552876@qq.com

¹**Shuhe Zhang,**

Master's student, e-mail: 1009383847@qq.com

³**Wanli Kang,**

Professor, e-mail: kangwanli@126.com

⁴**Abdusamig Kadirov,**

Professor, e-mail: abdusamig@rambler.ru

¹School of Petroleum Engineering, China University of Petroleum (East China),
Qingdao, PR China

²Shandong Key Laboratory of Oil and Gas Field Chemistry, China University of Petroleum
(East China), Qingdao, PR China

³XinJiang Keli New Technology Development Co.Ltd, Karamay, PR China

⁴Department of Physical Chemistry, National University of Uzbekistan, Tashkent, Uzbekistan

A REVIEW ON ANTI-ADSORPTION BEHAVIORS OF PARTIALLY HYDROLYZED POLYACRYLAMIDE IN RESERVOIR FORMATIONS

Abstract

Polymer flooding is one of the key technologies for enhancing oil recovery. Partially Hydrolyzed Polyacrylamide (HPAM) is widely used due to its excellent viscosity-increasing properties. However, the adsorption and retention behavior of HPAM in reservoir porous media presents a dual effect: on one hand, it improves sweep efficiency by increasing flow resistance; on the other hand, it leads to a loss in effective polymer concentration and viscosity, reducing displacement efficiency and increasing costs. Therefore, a systematic understanding and control of HPAM adsorption behavior are crucial for improving the effectiveness of polymer flooding. This work systematically reviews seven main measurement methods for HPAM adsorption quantity, comparing their applicable conditions and limitations. It summarizes the key factors influencing HPAM adsorption and retention behavior from three aspects: polymer properties, rock mineral characteristics, and reservoir environmental conditions. Furthermore, it outlines chemical anti-adsorption methods, represented by competitive adsorption and nanofilm protection, along with their mechanisms. Finally, future research directions are proposed, focusing on building adsorption prediction models, deepening the understanding of adsorption mechanisms under multi-field coupling conditions, and developing novel functional polymers with anti-adsorption capabilities.

Keywords: partially hydrolyzed polyacrylamide, adsorption retention, factors influencing adsorption, anti-adsorption, enhanced oil recovery.

1. Introduction

In the process of Enhanced Oil Recovery (EOR), specifically in chemical flooding, polymer flooding has become a vital technique. Among the polymers used, Partially Hydrolyzed Polyacrylamide (HPAM) is extensively applied due to its excellent water solubility and thickening capability [1]. When high-molecular-weight polymers enter the formation, retention occurs through mechanisms

including adsorption on mineral/rock surfaces, mechanical entrapment, and hydrodynamic retention, with adsorption being the primary mechanism. This widespread adsorption retention has both advantages and disadvantages. The beneficial aspect is that polymer adsorption and retention can plug large pore throats, thereby enlarging the sweep volume to some extent and consequently improving oil recovery [2]. Simultaneously, a higher retention amount leads to greater resistance factor and residual resistance factor. Studies show that a higher resistance factor enhances the polymer's mobility control capability, favoring the enlargement of the swept volume in the oil layer. A higher residual resistance factor is more conducive to increasing the displacement pressure and sweep volume during subsequent water flooding, thus achieving higher displacement efficiency. Additionally, adsorption retention increases the flow resistance of water molecules and polymer molecules in the aqueous phase, selectively reducing water phase permeability to achieve mobility control objectives, thereby expanding the sweep volume to a certain extent. Furthermore, the adsorbed film formed on the rock surface can prevent pore throat blockage caused by the migration of dispersed, mobile particles [3]. The detrimental aspects include the reduction of the effective polymer solution concentration due to adsorption retention, leading to decreased viscosity and weakened viscoelastic effects, which ultimately reduces the microscopic displacement efficiency and increases operational costs [4, 5]. Moreover, polymer adsorption and retention during flow in porous media can reduce permeability, and in severe cases, cause formation damage. Blockage by polymers and cross-linked polymers, in particular, can make it difficult for subsequent polymer solutions to enter formation pores, potentially leading to a significant decrease or even cessation of well production [6]. Furthermore, after years of development, various polymer-based profile control technologies have emerged, such as high-concentration polymer flooding, weak gel flooding, and pre-formed particle gel (PPG) technologies. These technologies aim to increase recoverable resources by plugging major flow channels. However, adsorbents and polymers may be adsorbed or retained during migration through porous media, affecting their gelation performance in areas far from the well-bore [7].

Although research on polymer adsorption is abundant, a systematic integration focusing on measurement methods, influencing factors, and reservoir anti-adsorption measures is currently lacking. Therefore, this work first systematically introduces the measurement methods for polymer adsorption, providing a comparative analysis of their advantages and disadvantages. Subsequently, it summarizes the factors influencing the adsorption and retention of polyacrylamide. Finally, it reviews several types of chemical anti-adsorption methods. The research outcomes can provide guidance for field applications concerning polymer concentration measurement, polymer molecular structure design, and concentration selection.

2. Measurement Methods for Polymer Adsorption Quantity

The adsorption quantity of a polymer is calculated based on the difference between the initial polymer concentration at the injection point and the concentration at the production end. The core of this calculation lies in measuring the polymer concentration in the produced fluid. Therefore, determining the polymer solution concentration is fundamental for analyzing adsorption issues in polymer flooding technology. The primary methods for quantifying HPAM concentration include turbidimetry, ultraviolet (UV) spectrophotometry, starch-cadmium iodide method, kinematic viscosity method, ultrafiltration membrane drying method, molecular size exclusion chromatography (SEC), and fluorescence spectrophotometry [5].

2.1 Turbidimetry

Turbidimetry is an analytical method that determines sample concentration or particle count by measuring the degree of light scattering by suspended particles in a liquid. Its principle involves the reaction of glacial acetic acid with sodium hypochlorite to generate Cl_2 , which then reacts with HPAM to form insoluble chloramide, causing turbidity in the solution. Under certain conditions, the turbidity value exhibits a linear relationship with the mass concentration of HPAM. The turbidity of the test solution is measured using a turbidimeter or spectrophotometer, and the HPAM concentration is determined from a pre-established standard calibration curve [8].

The main advantages of turbidimetry are its speed, simplicity, and low cost. However, this method requires a relatively large sample volume and is susceptible to interference from heavy metal ions and anionic surfactants, limiting its application. Additionally, factors such as the molecular weight, degree of hydrolysis, and molecular aggregation state of HPAM can influence the results. Given these characteristics, this method is suitable for preliminary field screening, process monitoring, and rough estimation of polymer concentration loss [9].

2.2 UV Spectrophotometry

UV spectrophotometry is based on the ultraviolet absorption of polymers, whose main chains consist of acrylamide and acrylic acid units, at a wavelength of 202.2 nm. Within a concentration range of 10–250 mg/L, the absorbance (E) shows a good linear relationship with concentration. The absorbance of the test solution is measured using a spectrophotometer, and the polymer concentration is calculated by comparison with a standard curve of known concentrations.

This method is simple, rapid, non-toxic, low-cost, and allows for sample recovery, making it an ideal laboratory technique. However, UV spectrophotometry has low sensitivity and is not suitable for low-concentration polymer solutions. Furthermore, it performs poorly in terms of linearity and stability for samples prepared with produced water from oil wells, primarily because organic impurities in the wastewater also absorb UV light. Therefore, when using UV spectrophotometry to determine HPAM concentration, strict requirements must be placed on water quality. If the water contains substances that absorb light around 200 nm, this method should not be used. Consequently, it is only applicable in laboratory settings and not suitable for direct detection of polymer concentration in oilfield produced water [10].

2.3 Starch-Cadmium Iodide Method

The reaction mechanism of this method involves bromine reacting with amide groups to generate N-bromoamide, which hydrolyzes to form hypobromous acid. Excess bromine is reduced and removed using sodium formate. Upon adding the starch-cadmium iodide reagent, I^- is oxidized by hypobromous acid to I_2 , which then forms I_3^- with I^- . I_3^- reacts with starch to form a blue starch-triiodide complex. The absorbance is subsequently measured using a UV spectrophotometer. A standard curve of concentration versus absorbance is plotted using a series of standard solutions with known concentrations. The accurate HPAM content in the test sample is then calculated based on its absorbance. The detection wavelength for the starch-cadmium iodide method ranges from 574 to 590 nm, with a buffer solution pH of 5.0. This colorimetric method offers high detection sensitivity and can accurately measure HPAM concentration in colored samples containing surfactants and dissolved crude oil. It is suitable for analyzing core effluent and large batches of routine samples [9].

2.4 Kinematic Viscosity Method

This method involves conducting simultaneous test and reference group experiments during static adsorption processes. A standard curve of reduced specific viscosity versus mass concentration is established using the reference group. The polymer concentration in the test group is then determined by measuring the solution viscosity and referring to the standard curve. This method controls the specific viscosity increment between 0.2 and 1.5, where the relationship between reduced specific viscosity and mass concentration is linear. The relative standard deviation (RSD) for precision is 0.61%, and for repeatability, it is 0.043%. The kinematic viscosity method features simple procedures, low cost, no need for chemical reagents, and applicability to complex systems containing polymer solutions for static adsorption concentration loss measurement. Its drawbacks include significant influences from molecular weight, degree of hydrolysis, pH, temperature, salinity, and shear history, leading to low measurement accuracy. This method is often used as a reference for polymer concentration detection, primarily for the preliminary determination of concentration loss in static adsorption [11].

2.5 Ultrafiltration Membrane Drying Method

This method primarily utilizes a Millipore tangential flow ultrafiltration system to ultrafilter and concentrate the test water sample. The sample is first filtered through a membrane package

with pores small enough to allow only small molecules to pass, thereby retaining HPAM molecules with larger hydrodynamic radii. By repeatedly diluting and washing the concentrated test solution to purify the HPAM, followed by evaporation, drying, and weighing, the HPAM concentration can be accurately determined. This method is unaffected by variations in polymer molecular weight and degree of hydrolysis, allowing for accurate determination of polymer mass concentration in produced fluids with high precision. However, the procedure is relatively complex and requires ultrafiltration concentration materials, making it costly for many operators [9, 12].

2.6 Molecular Size Exclusion Chromatography

The Molecular Size Exclusion Chromatography (SEC) method is an integrated separation and detection analytical technique. Polymers are separated from small molecule impurities on a chromatographic column with specific pore sizes, and the absorption peak is measured using an ultraviolet (UV) detector. The HPAM concentration is determined based on the proportional relationship between concentration and peak height. Commonly used columns are surface-modified Gly-CPG with controlled pore sizes, on which HPAM adsorption is minimal. Using special eluent systems in standard gel chromatography columns can also eliminate HPAM adsorption. Small molecules in the test sample, such as petroleum sulfonates, crude oil, isopropanol, sodium chloride, and sodium hexametaphosphate, can be effectively separated by SEC without interfering with HPAM concentration measurement. Therefore, severely contaminated and complex samples, such as field-produced emulsions containing surfactants and colored samples containing dissolved crude oil, can be accurately measured using SEC. The analysis speed of the SEC method is moderate, typically processing 3–5 samples per hour. Factors affecting HPAM hydrodynamic volume, such as water salinity, divalent cations, and HPAM hydrolysis under high-temperature alkaline conditions, can influence this method, but they can be mitigated by selecting chromatographic columns with appropriate pore sizes. Additionally, when the HPAM molecular weight distribution is very broad, some molecules might enter the pore structure of the packing material, leading to peak tailing [13].

2.7 Fluorescence Spectrophotometry

This method utilizes the Hofmann rearrangement reaction to convert HPAM into amine derivatives. In the presence of phthalaldehyde and mercaptoethanol, amide derivatives are converted into compounds that emit strong light. Measurements using fluorescence spectrophotometry show a linear relationship within the concentration range of 0–6 mg/L, with an RSD of 5%. HPAM concentrations as low as 20 µg/L can be effectively detected. Interference occurs when concentrations of NaCl, CaCl₂, and AlCl₃ exceed 1×10^{-3} mol/L. The prominent feature of this method is its ability to detect trace amounts of HPAM, and it can be used for measuring anionic, cationic, and non-ionic PAMs. Since the rearrangement reaction proceeds relatively slowly, the measurement process is time-consuming. Thus, it is commonly used for concentration detection when sample quantities are limited [13].

Among the seven measurement methods listed above, the SEC method and the starch-cadmium iodide method exhibit fewer interfering factors and high accuracy, making them widely used in both laboratory research and field sample analysis in oilfields. The kinematic viscosity method is often employed as a reference method for polymer concentration detection, valued for its speed and simplicity in the preliminary determination of static adsorption concentration loss. Turbidimetry offers relatively high accuracy and fewer interfering factors, leading to its relatively widespread application in oilfields. UV spectrophotometry is only suitable for measuring HPAM concentration in pure systems. Produced fluids from field polymer flooding and compound flooding are complex in composition, generally have high salinity, and are often colored due to dissolved crude oil or emulsification. The SEC method, with its function of separating small molecules, is particularly suitable for analyzing such samples. The advantages and disadvantages of each method are summarized in Table 1 below.

Table 1 – Methods for Determining Polymer Adsorption Quantity [8–13]

Method Name	Advantages	Disadvantages	Applicability
Turbidimetry	Rapid, simple, low cost.	Susceptible to interference from heavy metal ions, anionic surfactants, molecular weight, and degree of hydrolysis.	Suitable for preliminary field screening, process monitoring, and rough estimation of polymer concentration loss.
UV Spectrophotometry	Simple, rapid, non-toxic, low cost, sample recovery possible.	Low sensitivity, not suitable for low-concentration solutions; requires high water quality, organics in produced water cause severe interference.	Ideal laboratory method, but only applicable to clean systems with known makeup water composition.
Starch-Cadmium Iodide Method	High sensitivity, strong resistance to interference.	Complex and time-consuming procedure.	Suitable for precise analysis of core effluent and large batches of routine samples.
Kinematic Viscosity Method	Simple operation, low cost, no chemical reagents needed, applicable to complex systems.	Highly influenced by MW, HD, pH, temperature, salinity, shear history; low precision.	Often used as a reference method for preliminary determination of static adsorption concentration loss.
Ultrafiltration Drying Method	High precision, unaffected by polymer MW and HD.	Complex operation, time-consuming, requires specialized ultrafiltration equipment, high cost.	Accurate determination of polymer mass concentration in produced fluids; suitable for high-precision requirements.
Size Exclusion Chromatography (SEC)	Integrated separation and detection, excellent anti-interference ability, can handle complex/s contaminated samples.	Moderate analysis speed; salinity, divalent cations can affect separation, requires column optimization.	Suitable for complex field samples.
Fluorescence Spectrophotometry	Extremely high sensitivity, very low detection limit (down to $\mu\text{g/L}$ level), applicable to various ionic PAMs.	Rearrangement reaction is slow, measurement is time-consuming; high concentrations of inorganic salts ($>1 \times 10^{-3} \text{ mol/L}$) cause interference.	Mainly used for trace polymer detection, suitable for scenarios with limited sample volume and low concentration.

3. Factors Influencing the Adsorption Retention Quantity of HPAM

The adsorption behavior of Partially Hydrolyzed Polyacrylamide (HPAM) is primarily controlled by three categories of factors: the polymer itself, rock mineral characteristics, and reservoir conditions. The intrinsic properties of HPAM include its molecular structure, molecular weight, degree of hydrolysis, and concentration. Rock mineral characteristics encompass mineral type, surface properties, surface charge, structure, and wettability. Reservoir conditions involve temperature, pressure, and the types and ionic strength of electrolytes in the liquid phase. These factors collectively influence the adsorption and retention of HPAM during its flow through the reservoir formation [4].

In recent years, significant attention has been paid to the factors influencing HPAM adsorption in formations, and extensive research has been conducted. However, a holistic understanding of

its adsorption and retention behavior remains somewhat fragmented. Therefore, this review systematically collates and summarizes existing findings in this field.

3.1 Polymer Properties

During polymer flooding, factors such as polymer concentration, particle size, and dissolution rate can lead to polymer adsorption, loss, and retention within the formation [14].

Zhang et al. investigated the influence of HPAM concentration on its own adsorption retention in formations [15]. Their study indicated that within a certain concentration range, the adsorption quantity of HPAM increases with increasing solution concentration. Following this trend, once the concentration exceeds a critical value, the adsorption quantity increases abruptly; below this critical value, the relationship between adsorption quantity and solution concentration is less pronounced. If high molecular weight polymers adsorb onto the rock surface, numerous polymer chains come into contact with the rock. Although the adsorption amount per individual chain might be low, the adsorption of molecular aggregates is typically higher. Consequently, the adsorption isotherm of HPAM exhibits characteristics of strong adsorption: the adsorption density increases rapidly with concentration in the low-concentration region and can reach a plateau in the high-concentration region. The study also found that if the porous medium is first saturated by exposure to a dilute polymer solution, subsequent exposure to a higher HPAM concentration does not result in significant additional adsorption.

Yang Jiping et al. explored the relationship between the degree of hydrolysis and the adsorption quantity of HPAM on kaolinite and montmorillonite [1]. They discovered that the adsorption quantity of HPAM on both kaolinite and montmorillonite decreases with an increasing degree of hydrolysis, reaching a minimum at around 25–30% hydrolysis. Beyond this point, the adsorption quantity increases again with further increases in hydrolysis. Additionally, the amount of polymer adsorbed on montmorillonite was greater than on kaolinite. This is attributed to the significantly larger specific surface area of montmorillonite compared to kaolinite, leading to higher adsorption under the same conditions.

Hu Jingbang et al. studied the relationship between polymer adsorption retention and HPAM concentration within the range of 0–1500 mg/L [16]. They found that the dynamic adsorption retention amount increases with increasing HPAM concentration. The relationship between HPAM concentration and dynamic adsorption quantity can be approximately described by a straight line, with an average relative error of 5.8%, although the rate of increase is not substantial. This is because, at higher polymer concentrations, mechanical entrapment occurs more readily, increasing the number of molecules retained at trapping sites, thereby increasing the retention amount. In this scenario, the increase in adsorption amount itself is modest, so the overall increase in retention is determined by the entrapment amount.

3.2 Rock Mineral Characteristics

Rock mineral characteristics, including mineral type, surface properties, surface charge, structure, and wettability, are crucial factors influencing HPAM adsorption and retention behavior. The interfacial interaction mechanisms between different minerals and HPAM molecules vary significantly, directly determining the dynamic retention of the polymer in the formation and the resulting displacement efficiency. Therefore, studying the influence of mineral characteristics on adsorption behavior is of great importance.

Liu Xiao et al. investigated the adsorption of HPAM on different minerals [4]. Their results showed that the adsorption quantity of HPAM on minerals reaches saturation after a certain time, with no significant increase observed after 10 hours. The time required to reach equilibrium adsorption ranged from 20 to 70 hours, with clay minerals requiring longer times. The order of adsorption quantity on various solid materials was: gypsum > montmorillonite > kaolinite > chlorite > biotite > muscovite > dolomite > calcite > plagioclase > microcline > quartz. The adsorption quantity on gypsum could reach 0.36 mg/g.

Li Jun et al. studied the effect of initial concentration on HPAM adsorption on shale and sandstone [17]. They found that the adsorption quantity on rock surfaces increases with increasing

HPAM solution concentration. This is because a higher polymer concentration means more macromolecular chains are present in the solution, increasing the probability of contact with the rock surface, thus leading to more adsorption. Simultaneously, the adsorption quantity of HPAM on shale was significantly greater than on sandstone across different concentrations, with adsorption quantities ranging from 0.12 to 1.05 mg/g on shale and 0.06 to 0.51 mg/g on sandstone.

Li Yiqiang et al. observed that the presence of varying clay content in clean sand increases the adsorption quantity [3]. Under reservoir conditions at 45°C, the adsorption of HPAM in oil-free quartz sand increased with concentration, exhibiting an “S”-shaped adsorption isotherm. The maximum adsorption occurred around a solution concentration of 1200 mg/L. The adsorption quantity was only 0.953 mg/g in quartz sand without kaolinite but increased to 1.725 mg/g when the kaolinite content reached 5%. Furthermore, the presence of kaolinite in the porous medium significantly increased HPAM retention due to its specific lattice structure. When the kaolinite content increased from 2% to 5%, the retention amount increased from 0.584 mg/g to 0.794 mg/g. Additionally, under identical temperature and kaolinite content conditions, the adsorption amount of HPAM in quartz sand was far greater than the retention amount in the porous medium.

3.3 Reservoir Conditions

Reservoir conditions include temperature, pressure, wettability, and the types and ionic strength of electrolytes in the liquid phase. A deep understanding of the multi-scale regulatory role of reservoir conditions on adsorption behavior is essential for studying polymer adsorption and retention in oil reservoirs.

YANG L, et al. indicated that the adsorption of HPAM on reservoir rocks largely depends on the wettability of the rock surface [15]. The adsorption amount on water-wet models was much higher than on oil-wet models.

Hu Jingbang et al. investigated the effect of salinity (2000–10000 µg/mL) on the dynamic adsorption retention of HPAM in porous media [16]. They found that within this salinity range, the dynamic adsorption quantity initially decreased rapidly with increasing salinity and then tended to stabilize. The study also noted that the saturated dynamic retention quantity is the sum of the dynamic adsorption quantity and the amount of polymer retained in the accessible pore volume of the core. While increased salinity reduces dynamic adsorption, the concomitant shrinkage of HPAM molecular size increases the accessible pore volume of the core, leading to an increase in the amount of polymer retained therein. Consequently, the saturated retention quantity changes more gradually with salinity. Furthermore, permeability also significantly affects HPAM adsorption retention: lower permeability leads to greater dynamic adsorption quantity. Under constant injection rate, the retention mechanisms are primarily adsorption and mechanical entrapment. Low-permeability cores have smaller pore throat radii, resulting in stronger mechanical entrapment. Although their smaller accessible pore volume and adsorption surface area lead to lower pure adsorption, the total dynamic adsorption quantity remains high. As permeability increases, the inaccessible pore volume decreases and the adsorption surface area increases, potentially increasing the adsorption component. However, due to the enlargement of pore throat radii, mechanical entrapment decreases significantly, causing the total dynamic adsorption quantity to drop rapidly.

3.4 Chemical Additives

Besides polymer properties, rock mineral characteristics, and reservoir conditions, chemical additives represent another important factor influencing the adsorption and retention behavior of HPAM.

Li Yiqiang et al. pointed out that the presence of surfactants can reduce the retention of HPAM in porous media. As surfactant concentration increases, HPAM retention decreases [5]. This is attributed to competitive adsorption between the surfactant and polymer on the rock surface.

Zhou Yangfan confirmed that the introduction of surfactants suppresses the adsorption of HPAM on rock surfaces [21]. Their experiments, conducted at 45°C using distilled water to prepare HPAM-surfactant mixed solutions, showed that the adsorption isotherm of HPAM on oil-free formation sand was “L”-shaped, and the equilibrium adsorption quantity in the composite system was significantly

lower than in the polymer-only system. This phenomenon stems from the competitive adsorption between the polymer and surfactant, thereby reducing the adsorption amount of HPAM.

4. Anti-adsorption Methods During Polymer Flooding

To efficiently utilize the oil displacement effect of polymers and reduce adsorption losses, the addition of chemical additives into the formation for anti-adsorption purposes is a common field practice. Based on literature review, three primary types of chemical anti-adsorption methods are currently employed. The first method involves using silane coupling agents as anti-adsorbents to modify the sandstone surface into a hydrophobic one, thereby reducing polymer adsorption. The second method utilizes chemicals like surfactants, which readily adsorb onto reservoir rock surfaces, to compete with the polymer for adsorption sites, thus achieving a reduction in polymer adsorption. The final method employs nano-film agents, which utilize electrostatic interactions between ions to deposit and reach saturation on the reservoir surface, forming a monomolecular film that alters the adsorption properties of the reservoir rock and reduces polymer adsorption.

The mechanism by which silane coupling agents reduce adsorption is as follows: After injection, the coupling agent first hydrolyzes to generate hydroxyl compounds. These hydroxyl compounds undergo etherification reactions with the silanol groups on the sandstone surface, forming Si-O-S groups. This reaction exposes the hydrophobic chains of the anti-adsorbent, thereby modifying the sandstone surface into a hydrophobic one, which in turn reduces the adsorption amount of the polymer.

Fu Meilong et al. addressed the issue of severe production decline in Henan Oilfield caused by polymer blockage in injection wells during polymer flooding [18]. They developed a silane coupling agent anti-adsorbent, with n-octyltriethoxysilane as the main agent, and acidic catalysts and alcoholic co-solvents as additives. The study primarily analyzed the effects of anti-adsorbent mass fraction, salinity, injected volume of anti-adsorbent, treatment duration, and injection mode on the polymer anti-adsorption rate. The results showed that the optimal injection mass fraction for the developed silane coupling agent anti-adsorbent was 0.5%. Within the salinity range of 1500–10000 mg/L, salinity had little effect on the anti-adsorption rate of this system. Under conditions of pH=3, temperature 80°C, core permeability 408.6 mD, salinity 2500 mg/L, anti-adsorbent mass fraction 0.5%, treatment time 12 hours, and polymer injection rate 0.5 mL/min, a larger injected slug volume resulted in lower dynamic adsorption of polymer on the rock surface and a higher polymer anti-adsorption rate. Considering both treatment effectiveness and economic cost, the optimal injected slug volume for the anti-adsorbent was determined to be 0.6 PV. When the treatment time exceeded 12 hours, the anti-adsorption rate was greater than 70%. Under conditions of pH=3, temperature 80°C, core permeability 286.12 mD (Core 5), salinity 2500 mg/L, anti-adsorbent slug volume 0.6 PV, anti-adsorbent mass fraction 0.5%, polymer injection rate 0.5 mL/min, and treatment time 12 hours, the optimal injection sequence was: displacing fluid + 0.6 PV anti-adsorbent slug + 1 PV cleaning agent.

Gao Shenling et al. studied the effects of permeability, temperature, pH, and salinity on the anti-adsorption rate of the silane coupling agent [19]. They found that when permeability was not less than 196 mD, the silane coupling agent maintained a high anti-adsorption rate for the polymer, staying above 70%. Higher temperatures led to lower dynamic adsorption of polymer on the rock surface. The increase in anti-adsorption rate diminished after 80°C, but the anti-adsorption rate remained above 70%. Higher fluid pH values resulted in lower anti-adsorption rates for the treatment agent. The optimal pH range for this anti-adsorbent system to reduce polymer adsorption was between 3 and 5, with the lowest anti-adsorption rate observed at pH=7. Due to the good salt tolerance of this anti-adsorbent system, different salinities had little effect on its anti-adsorption rate, which remained above 70%. A higher mass fraction of the anti-adsorbent system led to lower dynamic adsorption of polymer on the rock surface and a higher polymer anti-adsorption rate. When the mass fraction was not less than 0.5%, the anti-adsorption rate gradually stabilized above 70%.

Surfactants can compete with polymers for adsorption sites. Based on this principle, they can be used to suppress and reduce the adsorption retention of polymers in formation pores.

Guo Yongjun et al. examined the adsorption behavior of a water-soluble hydrophobically associating polymer (acrylamide/sodium acrylate/octadecyldimethyl allyl ammonium chloride copolymer) at the kaolinite/water interface [20]. They reported that the adsorption amount of the copolymer on kaolinite increased with equilibrium concentration, exhibiting characteristics of multilayer adsorption. Increasing the concentration of SDBS (sodium dodecyl benzene sulfonate) in the solution led to a monotonic decrease in the copolymer adsorption amount on kaolinite. Zhou Yangfan conducted experiments on the effect of surfactants on HPAM adsorption [21]. The results showed that under experimental conditions of 45°C, using distilled water to prepare polymer-surfactant mixed solutions, the adsorption isotherm on oil-free formation sand after reaching equilibrium was “L”-shaped. Due to competitive adsorption between the polymer and surfactant components, the adsorption amount of HPAM under the binary composite condition was significantly lower than in the single-component condition without surfactant. This aligns with the general rule that in multi-solute mixed solutions, the adsorption amount of each solute is less than that in its single-component solution. Therefore, the polymer-surfactant binary composite flooding method can be employed to reduce polymer adsorption issues in the formation.

Nano-film agents, after injection into the formation, utilize electrostatic interactions between ions to deposit and reach saturation on the reservoir surface, forming a monomolecular film that alters the adsorption properties of the reservoir rock. Because the film is extremely thin, it does not affect the permeability of the reservoir. When the polymer solution flows, the rock surface covered with the molecular deposition film makes polymer adsorption less likely, reducing flow resistance and thus preventing formation blockage caused by polymer adsorption and retention.

Xu Guomin et al. proposed a nano-film anti-adsorption technology [22]. Their research results indicated that after injection into the formation, the nano-film firmly adsorbs onto the rock surface through electrostatic interactions, forming a nano-scale ultra-thin film. This film combines strongly with the substrate, is resistant to flushing, and plays a significant role in reservoir protection. When polymer solutions flow over rock surfaces with this molecular deposition film, adsorption is less likely to occur, flow resistance is reduced, and thus formation blockage due to polymer adsorption and retention can be prevented.

5. Conclusions and Perspective

This work has systematically reviewed the research methods, influencing factors, and anti-adsorption techniques related to the adsorption behavior of Partially Hydrolyzed Polyacrylamide (HPAM) in reservoir formations. Although existing research has yielded substantial results, further efforts are required to achieve more efficient and economical polymer flooding technology. Based on the content of this review and current development trends, the following future research directions for polymer adsorption studies are proposed.

(1) Digitization of Adsorption Retention and Construction of Predictive Models. With the rapid development of artificial intelligence and big data, it has become feasible to construct predictive models for polymer loss based on factors such as polymer properties, rock mineral characteristics, and reservoir conditions. Utilizing machine learning algorithms to train on vast amounts of experimental and field data can create intelligent tools capable of quickly and accurately predicting polymer adsorption quantity under specific reservoir conditions. This will significantly optimize the design of flooding schemes and promote the intelligent development of polymer flooding technology.

(2) Research on Adsorption Mechanisms under Multi-Physics-Chemical Field Coupling. Current adsorption studies are often conducted under single-variable or idealized conditions. However, actual reservoirs are complex systems characterized by high temperature, high salinity, strong heterogeneity, and the presence of residual oil. Therefore, future research needs to strengthen the investigation of the dynamic adsorption and retention mechanisms of polymer molecules under multi-field coupling (e.g., temperature field, chemical field, stress field) and within micro-nano scale pores. Utilizing technologies such as micro-visualization models and molecular dynamics simulations to reveal the adsorption retention process at the pore scale will provide a more solid theoretical foundation for studying polymer adsorption behavior.

(3) Precision Design of Novel Functional Polymers with Anti-adsorption Properties. Future efforts should focus on precision design at the molecular level, such as developing self-responsive polymers that can adapt their configuration based on the formation environment. Concurrently, bio-based polymers or nano-composite polymers, owing to their unique environmental friendliness and high stability, also represent promising directions for improving oil displacement efficiency.

REFERENCES

- 1 Yang, J., Li, H., Huang, P. The Static adsorption of HPAM on porous media –influence of hydrolysis degree on the quantity of adsorbed HPAM. *Journal of Polymer Science*, 05, 90–94 (1997).
- 2 Lai, N., Zhang, Y., Zhao, X., et al. Adsorption and retention characteristic of hyperbranched polymers in porous media. *Chemical Research and Application*, 028(003), 360–365 (2016).
- 3 Li, Y., Qu, Y. Influence of kaoline to static adsorption and dynamical retention of polyacrylamide [J]. *Offshore oil*, 30(02), 72–76 (2010).
- 4 Liu, X., Chen, Z., Zeng, F. Effect of mineral property on the adsorption of polyacrylamide [J]. *Mineralogy and Petrology*, 01, 97–100 (2004).
- 5 Li, Y., Qu, C. Research on dynamic hold-up of water-soluble polymer in porous medium [J]. *Oil Drilling & Production Technology*, 33(01), 76–79 (2011).
- 6 Sun, X., Zhang, W., Li, J., et al. Mechanism and Performance Analysis of Nanoparticle-Polymer Fluid for Enhanced Oil Recovery: A Review. [J]. *Molecules* (Basel, Switzerland), 2023, 28(11).
- 7 Pi, Y., Liu, J., Cao, R., et al. Visualized Study on a New Preformed Particle Gels (PPG) + Polymer System to Enhance Oil Recovery by Oil Saturation Monitoring Online Flooding Experiment [J]. *Gels*, 9(2), 81 (2023).
- 8 Guan, S., Fan, H., Wu, S., et al. Determination of Polyacrylamide Mass Concentration – Turbidity Method [J]. *Journal of Northeast Petroleum University*, 31(2), 106–109 (2007).
- 9 Zhou, X., Tang, S., Sheng, H., et al. Overview of Polyacrylamide Concentration Measurement Methods [J]. *Guangzhou Chemical Industry*, 47(16), 4 (2019).
- 10 Zhu, H., Cai, Y., Wei, Y., et al. Concentration measurement of novel water-soluble polymers used in EOR with ultraviolet spectroscopy [J]. *Fine and Specialty Chemicals*, 12(6), 4 (2004).
- 11 Li, J., Yan, X. A new method for determining concentration loss of oil displacement polymer by static adsorption and its application [J]. *Petrochemical Technology*, 53(6), 884–889 (2024).
- 12 Li, Q. Research progress on polyacrylamide concentration detection methods [J]. *Chemical Management*, 17, 1 (2016).
- 13 Kong, B. Overview of Polyacrylamide Concentration Measurement Methods [J]. *Oilfield Chemistry*, 13(3), 6 (1996).
- 14 Xu, T., Cao, T., Huang, J., et al. Blockage of Polymers in High Permeability Cores and Mechanism of Reactive Anti-adsorption System [J]. *Contemporary Chemical Industry*, 2021.
- 15 Yang, L., Liu, Y., Wu, X., et al. The research status and summary of adsorption and retention mechanism of polymer [J]. *IOP Conference Series: Materials Science and Engineering*, 892(1), 012023 (2020).
- 16 Hu, J., Li, X., Han, C. An Experimental Study on The Dynamic Adsorption of Partially Hydrolyzed Polyacrylamides in Porous Media [J]. *Oilfield Chemistry*, (4), 5 (1991).
- 17 Li, J., Lan, L., Xiao, M., et al. Research on the Adsorption Behavior of Polyacrylamide on the Surfaces of Shale and Sandstone [J]. *Standard Science*, (S1)(2025).
- 18 Fu, M., Zhang, Z., Hou, B., et al. Polymer plugging mechanism and prevent technology of polymer injection wells in Henan Oilfield [J]. *Journal of Yangtze University (Natural Science Edition)*, 18(03), 37–48 (2021).
- 19 Gao, S., Chen, Y., Wang, X., et al. Research on Dynamic Performance of Anti-polymer Adsorbent in Henan Oilfield [J]. *Contemporary Chemical Industry*, 50(09), 2170–2175 (2021).
- 20 Guo, Y., Li, F., Li, L., et al. Adsorption Behavior of Water-Soluble Hydrophobic Associative Polymer at Water/Kaolin Interface [J]. *Chinese Journal of Applied Chemistry*, (01), 26–29 (2002).
- 21 Zhou, Y. The Interaction between Polymer and Surfactant Binary System [D]. *Southwest Petroleum Institute*, 2005.
- 22 Xue, G., Zhu, H. Nanofilm anti adsorption technology [J]. *Chemical Engineering & Equipment*, (02), 84–85 (2021).

¹Хаобин Ши,
магистрант, e-mail: 2842117353@qq.com
^{1,2*}Хунбин Ян,
доцент, *e-mail: hongbinyang@upc.edu.cn
¹Син Чэнь,
магистрант, e-mail: xin_baobei77@163.com
¹Лян Пэн,
магистрант, e-mail: 1539552876@qq.com
¹Шухэ Чжан,
магистрант, e-mail: 1009383847@qq.com
³Ванли Кан,
профессор, e-mail: kangwanli@126.com
⁴Абдусамиг Қадыров,
профессор, e-mail: abdusamig@rambler.ru

¹Мұнай инженериясы мектебі, Қытай мұнай университеті (Шығыс Қытай),
Циндао қ., ҚХР

²Шаньдун провинциясының мұнай және газ кен орындарының химиясы жөніндегі негізгі
зертханасы, Қытай мұнай университеті (Шығыс Қытай), Циндао қ., ҚХР

³XinJiang Keli New Technology Development Co. Ltd, Қарамай қ., ҚХР

⁴Физикалық химия кафедрасы, Өзбекстан ұлттық университеті,
Ташкент қ., Өзбекстан

ҚАБАТ ЖАҒДАЙЛАРЫНДАҒЫ ЖАРТЫЛАЙ ГИДРОЛИЗДЕНГЕН ПОЛИАКРИЛАМИДТІҢ АНТИАДСОРБЦИЯЛЫҚ ҚАСИЕТТЕРІНЕ ШОЛУ

Аңдатпа

Полимерлі айдау – мұнай өндіруді арттырудың негізгі технологияларының бірі. Жартылай гидролизденген полиакриламид (НРАМ) жоғары тұтқырлық түзуші қасиеттеріне байланысты кеңінен қолданылады. Алайда НРАМ-ның қабаттың кеуекті ортасында адсорбцияланып, ұсталып қалуы екі түрлі әсер туындатады: бір жағынан, фильтрацияға кедергіні арттырып, қамту аймағын кеңейтеді; екінші жағынан, ерітіндідегі тиімді полимер концентрациясы мен тұтқырлығын төмендетіп, мұнай ығыстыру тиімділігін азайтады және технологиялық шығындарды көбейтеді. Осыған байланысты НРАМ адсорбциясын терең түсіну және оны бақылау полимерлі айдаудың нәтижелілігін арттырудың маңызды шарты саналады. Бұл шолуда НРАМ адсорбция мөлшерін анықтаудың жеті негізгі әдісі жүйеленіп, олардың қолдану ауқымы мен шектеулері салыстырмалы талданады. Сонымен қатар адсорбцияға әсер ететін басты факторлар – полимердің молекулалық сипаттамалары, жыныстың минералдық құрамы және қабат жағдайлары бойынша ғылыми қорытындылар сараланады. Химиялық антиадсорбциялық тәсілдер – бәсекелі адсорбция және нанопленкалық қорғаныс – олардың әрекет ету механизмдерімен бірге сипатталады. Сонымен қатар болашақ зерттеу бағыттары ұсынылады: адсорбцияны болжауға арналған модельдерді жетілдіру, көп өрісті әсерлер жағдайындағы адсорбция механизмдерін терең зерделеу және антиадсорбциялық қасиеттері жақсартылған жаңа функционалды полимерлерді әзірлеу.

Тірек сөздер: жартылай гидролизденген полиакриламид, адсорбция және ұсталып қалу, адсорбцияға әсер ететін факторлар, антиадсорбция, мұнай өндіруді арттыру.

¹Хаобин Ши,
магистрант, e-mail: 2842117353@qq.com
^{1,2*}Хунбин Ян,
доцент, *e-mail: hongbinyang@upc.edu.cn
¹Син Чэнь,
магистрант, e-mail: xin_baobei77@163.com
¹Лян Пэн,
магистрант, e-mail: 1539552876@qq.com
¹Шухэ Чжан,
магистрант, e-mail: 1009383847@qq.com
³Ванли Кан,
профессор, e-mail: kangwanli@126.com
⁴Кадыров,
профессор, e-mail: abduamig@rambler.ru

¹Школа нефтяной инженерии, Китайский университет нефти (Восточный Китай),
г. Циндао, КНР

²Шаньдунская ключевая лаборатория химии нефтегазовых месторождений,
Китайский университет нефти (Восточный Китай), г. Циндао, КНР

³XinJiang Keli New Technology Development Co. Ltd, г. Карамай, КНР

⁴Кафедра физической химии, Национальный университет Узбекистана,
г. Ташкент, Узбекистан

ОБЗОР АНТИАДСОРБЦИОННЫХ СВОЙСТВ ЧАСТИЧНО ГИДРОЛИЗОВАННОГО ПОЛИАКРИЛАМИДА В ПЛАСТОВЫХ УСЛОВИЯХ

Аннотация

Полимерное заводнение является одной из ключевых технологий повышения нефтеотдачи. Частично гидролизированный полиакриламид (НРАМ) широко используется благодаря своим превосходным вязкоупругим свойствам. Однако процессы адсорбции и удерживания НРАМ в поровом пространстве коллектора имеют двойственный эффект: с одной стороны, они повышают охват вытеснением за счет увеличения сопротивления фильтрации, с другой – приводят к снижению эффективной концентрации и вязкости полимера, уменьшают эффективность вытеснения и повышают затраты. Поэтому систематическое понимание и контроль адсорбционного поведения НРАМ имеют решающее значение для повышения эффективности полимерного заводнения. В данной работе систематически рассмотрены семь основных методов измерения количества адсорбции НРАМ, проведено сравнение их применимости и ограничений. Обобщены ключевые факторы, влияющие на адсорбцию и удерживание НРАМ, с точки зрения свойств полимера, минералогических характеристик породы и условий пластовой среды. Кроме того, рассмотрены химические методы предотвращения адсорбции, такие как конкурентная адсорбция и нанопленочная защита, и раскрыты их механизмы. Наконец, предложены будущие направления исследований, включая создание моделей предсказания адсорбции, углубление понимания механизмов адсорбции при многополевой сопряженности и разработку новых функциональных полимеров с антиадсорбционными свойствами.

Ключевые слова: частично гидролизированный полиакриламид, адсорбция и удерживание, факторы, влияющие на адсорбцию, антиадсорбция, повышение нефтеотдачи.

Article submission date: 24.11.2025

ЭКОНОМИКА ЖӘНЕ БИЗНЕС
ECONOMY AND BUSINESS
ЭКОНОМИКА И БИЗНЕС

UDC 65.011.4
IRSTI 06.81.19

<https://doi.org/10.55452/1998-6688-2025-22-4-444-455>

¹Sadykov E.,

MBA, ORCID ID: 0000-0004-8035-628X,

e-mail: e.sadykov@kazatomprom.kz

*¹Yesmurzaeva A.,

PhD student, ORCID ID: 0000-0003-0114-5925,

*e-mail: aknuryesmurzayeva@gmail.com

¹International Information Technology University, Almaty, Kazakhstan

MODERNIZATION OF PRODUCTION CAPACITIES
OF JSC “NAC “KAZATOMPROM”: STRATEGIC ANALYSIS

Abstract

The study examines the modernization processes of production capacities at JSC “NAC Kazatomprom” through a comprehensive strategic analysis based on SWOT, PESTLE, and BCG-matrix assessment methods. As the world’s largest uranium producer, the company faces the dual challenge of maintaining its cost leadership and adapting to global market volatility, technological change, and growing ESG requirements. The research identifies key drivers of modernization including increasing extraction efficiency, capital investment growth, digitalization, and technological innovation — as well as constraints related to rising production costs and external geopolitical risks. The analysis demonstrates that modernization is essential for sustaining Kazakhstan’s competitive position in the global nuclear fuel market and ensuring long-term operational stability. Practical implications include recommendations for optimizing investment programs, enhancing technological upgrading, and strengthening strategic resilience in the context of the global energy transition.

Key words: modernization of production capacities, uranium mining industry, strategic analysis, production digitalization, enterprise efficiency.

Introduction

Kazakhstan, the world’s largest uranium producer, accounts for up to 40% of global production, and the sustainability of its position largely depends on the technical condition of its production facilities and their adaptation to changing global market conditions [1].

The majority of JSC NAC Kazatomprom’s assets were established in the 1990s and 2000s, and despite their unique resource base and competitive costs, a significant portion of their equipment and technological solutions require updating. The dominance of in-situ leaching (ISL), which ensures environmentally friendly and relatively cost-effective production, is currently associated with

challenges related to the need to increase uranium recovery rates, optimize chemical reagent costs, and implement digital monitoring systems.

An additional aspect of the problem is the need for modernization in the face of volatile global uranium prices and increased competition from Canada, Australia, and new players in the African market. Long-term contracts with international partners and growing interest in nuclear energy as a low-carbon energy source are creating the preconditions for increased demand for Kazatomprom's products. However, this growth is only possible with timely modernization of capacities, which will improve the efficiency and stability of production and minimize environmental and social risks.

The current challenge lies in the need to transition from exploiting existing production potential to its technological upgrade, which will allow the company to maintain its leadership position in the global uranium market and ensure strategic sustainability in the face of the energy transition and stricter environmental requirements.

The relevance of studying the modernization of Kazatomprom's production capacities is determined by the company's key role in the global uranium industry and its importance to the economy of Kazakhstan. Amid the global energy transition and growing interest in nuclear energy, competition among leading raw material suppliers is intensifying, requiring the company to upgrade its technology, improve the efficiency of its production processes, and implement innovative solutions. Maintaining a leading position in the global uranium market is impossible without modernizing capacity, which ensures cost reduction, sustainable development, and compliance with modern environmental and industrial standards.

The purpose of this study is to conduct a strategic analysis of modernization processes at the company's enterprises, identify the factors determining its effectiveness, and substantiate development directions that can strengthen JSC NAC Kazatomprom's position as a global leader in the uranium mining industry.

The practical significance of this study lies in the potential application of the obtained results to improving Kazatomprom's corporate strategy, developing investment programs, and increasing the efficiency of resource utilization. This analysis can serve as a tool for government agencies in formulating nuclear energy policy and Kazakhstan's export potential, as well as for investors and industry experts assessing the development prospects of the national nuclear industry in the context of the global energy transformation.

Materials and methods

One of the key conditions for ensuring high competitiveness is the modernization of production processes. In the classical sense, modernization (from the English "modern" — updated, contemporary, rapid growth of scientific knowledge), as O. Bobkov points out, is interpreted as "updating an object, bringing it into compliance with new requirements and standards, technical specifications, and quality indicators" [2]. Through modernization, enterprises gain the ability to maintain a modern level of technical and technological potential, ensure high product quality, increase labor productivity, and successfully integrate into foreign markets.

M.E. Nadezhdina points out that modernization presupposes the development of the ability to generate advanced scientific and technological ideas, export the results of research and development, and quickly transition to the use of new information technologies and the production of competitive products [3]. In this regard, enterprise modernization can be implemented as a structural process encompassing a system of interconnected areas. It includes the creation of new types of products or the improvement of existing ones, increased efficiency in the use of technological equipment, a reduction in the labor intensity of production processes, and the optimization of personnel. An important consequence of these changes is the reduction of production cycle times, the minimization of production and non-production losses, and a reduction in production costs, achieved through the use of advanced technologies, new materials, and the rational use of energy and labor resources.

According to A.I. Shinkevich, “modernization of an enterprise’s production capacity is a set of measures aimed at improving and enhancing the enterprise’s technical base to increase production efficiency, expand production capacity, and enhance the company’s competitiveness” [4]. Table 1 shows the types of modernization of enterprise production capacity.

Table 1 – Types of Modernization of Enterprise Production Capacity

Type of modernization	Characteristics	Economic effects	Cost scale
Partial	Replacement of individual equipment elements or subsystems while maintaining the production structure and product range. Disruption to the production cycle is minimal.	Cost reduction, increased reliability of individual processes, and maintenance of current capacity.	Low
Comprehensive	Upgrading significant parts of the production system, including retooling workshops or sections, introducing new technologies and workflow methods.	Temporary reduction in production output, accelerated capacity growth after implementation, and increased production efficiency.	Medium
Full	Complete reconstruction of production using modern equipment and the latest technology and restructuring the organizational structure.	Significant increase in productivity (up to 20–30%), improved product quality, and enhanced competitiveness.	High
Note: compiled by the author based on data from the source [5].			

An analysis of the presented data allows us to conclude that modernizing an enterprise’s production facilities is an essential condition for improving the efficiency of production processes and strengthening its competitive position in the market. The nature and depth of the transformation are determined by the chosen type of modernization, which can range from localized upgrades of individual equipment elements to a comprehensive restructuring of the entire production system.

A detailed comparative analysis of these groups is presented in Figure 1. Technological modernization involves the implementation of the latest technologies, the re-equipping of industrial enterprises with modern equipment, and the upgrading of employee skills. These processes ensure increased labor productivity, improved product quality, and enhanced competitiveness in both domestic and global markets, but require significant investment.

Infrastructure modernization encompasses the banking, financial, social, and medical sectors, as well as the transport and logistics system, which unites industries, enterprises, and organizations that ensure the organizational and economic interaction of economic entities based on the principles of economic efficiency. It involves creating conditions for the sustainable functioning of the national economy and its entities in the face of market fluctuations. Public administration plays a key role in this process, guiding development and supporting the stability of infrastructure [7].

Institutional modernization is a complex of transformations that depends on the degree of involvement of various entities and the nature of their interaction. In some cases, it is implemented primarily through the efforts of the private sector, with the state acting as a coordinator; in others, it is based on the strict dominance of state institutions, which determine the priority areas of transformation. There are also models that assume a balance of interests between the state and the private sector, based on their interaction in the formation and implementation of modernization strategies. Thus, modernization of production facilities is a key factor in improving the efficiency and competitiveness of an enterprise. It ensures technological innovation, increased productivity, and the rational use of resources.

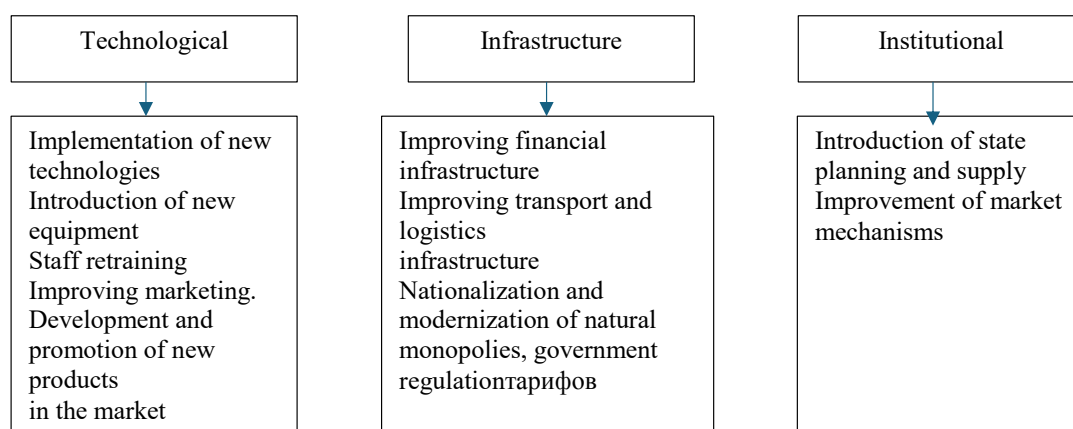


Figure 1 – Types, Methods, and Measures of Modernization

Note: Compiled by the author based on [6].

Comprehensive modernization contributes to strengthening economic resilience, developing competitive advantages, and adapting to environmental changes. Its implementation lays the foundation for the transition to innovative development and long-term growth of the enterprise.

To ensure the scientific validity of the analysis of production facility modernization processes, it is important to define strategic analysis methods that allow us to identify the internal and external factors influencing the effectiveness of an enterprise's development. The use of these methods provides a comprehensive understanding of the company's current state, its competitive advantages, and areas for strategic growth. A comparative analysis of the most common strategic analysis methods is presented in Table 2.

Table 2 – Comparative Analysis of Strategic Analysis Methods

Strategic Analysis Method	Method Contents	Advantages	Limitations	Appropriateness of Use
SWOT-analysis	Assessment of the company's strengths and weaknesses, as well as external opportunities and threats	Ease of use, clarity, integration of qualitative and quantitative factors	Subjectivity of assessments, limited depth of analysis	Effective for a comprehensive assessment of a company's current position and the formation of strategic development directions
PESTLE-analysis	Analysis of political, economic, social, technological, legal, and environmental factors in the external environment	Allows consideration of a wide range of macrofactors, including legal and environmental aspects; provides a more comprehensive understanding of the external environment	Does not reflect the internal characteristics of the enterprise; can be labor-intensive given the large number of factors	Optimal for an in-depth analysis of the external environment's impact on modernization processes and long-term strategic planning

Continuation of table 2

BCG-matrix	Classification of business areas by market growth rate and company share	Forms a strategy for managing a portfolio of business areas	Requires accurate market data; applicable to diversified structures	Suitable for assessing the effectiveness of individual production areas
SPACE-analysis	Determination of the company's strategic position based on an assessment of financial, competitive, industry, and stability factors	Provides a holistic view of strategic sustainability	Difficulty in quantitatively assessing factors	Appropriate for choosing a strategic development model for an enterprise
SNW-analysis	Modification of the SWOT analysis by highlighting neutral characteristics that do not directly impact strategy	Increases the accuracy and objectivity of analysis	Requires high analytical skills	Useful for refining strategic modernization priorities
Note: compiled by the author based on data from the source [8].				

A comparison of strategic analysis methods reveals that SWOT analysis is most applicable to studying production capacity modernization, as it provides a comprehensive assessment of the internal and external factors that determine an enterprise's strategic position. Its use in Chapter 2 will help identify the strengths and weaknesses of JSC NAC Kazatomprom's production system, identify opportunities and threats in the external environment, and substantiate modernization strategies that will improve the company's efficiency and sustainability.

Thus, modern approaches to strategic modernization management are aimed at improving the efficiency, sustainability, and competitiveness of enterprises. Their results depend on a combination of innovation, resource optimization, and adaptation to environmental changes.

The use of strategic analysis methods allows for the informed determination of development directions. The most effective tool is SWOT analysis, which, in Chapter 2, will help identify key modernization factors and define the strategic priorities of JSC NAC Kazatomprom.

The research methodology relies on a systems approach, allowing the company's production complex to be considered as a holistic socio-economic and technological system operating within the global uranium market. Concepts of strategic management, institutional economics, and sustainable development serve as a theoretical framework, enabling analysis of the impact of internal and external factors on modernization processes.

The research methodology combines quantitative and qualitative analytical tools. Corporate reporting data, statistics from the World Nuclear Association and the International Atomic Energy Agency, and analytical materials from industry research centers are used. Comparative analysis and economic-statistical modeling are used to assess production capacity dynamics, enabling the identification of key trends and growth factors. SWOT and PEST analysis elements, as well as the BCG matrix, are used to identify the company's strengths and weaknesses, as well as the opportunities and threats associated with the transformation of production processes. The use of a case study method provides a deeper understanding of the practical aspects of modernization, including technological innovation, digitalization of production chains, and improvements to the corporate governance system.

Results and discussion

Joint Stock Company National Atomic Company Kazatomprom is the world's largest producer of natural uranium, with priority access to one of the most significant resource bases. Since 2018,

the company's shares have been traded on the Astana International Financial Centre (AIFC) and the London Stock Exchange, demonstrating its high level of transparency and investment appeal.

Kazatomprom is the national operator of the Republic of Kazakhstan for the export and import of uranium, its compounds, nuclear fuel for nuclear power plants, and specialized equipment and technologies. This status grants the company unique rights, including the ability to obtain subsoil use licenses through direct negotiations with the Government of Kazakhstan. This system of interaction guarantees priority access to deposits with high-quality natural uranium reserves suitable for cost-effective in-situ recovery (ISR) mining.

The company's operations are focused exclusively on Kazakhstan, home to the industry's largest uranium reserve base, exceeding 300,000 tonnes. The in-situ leaching (ISL) method used ensures high environmental safety and low production costs, allowing Kazatomprom to maintain its leading position in the global market in terms of production costs. This technology eliminates direct human involvement in the production process, reducing production risks and environmental impact [9]. The industrial safety indicators of JSC NAC Kazatomprom are presented in Table 3.

Table 3 – Industrial Safety Indicators of JSC NAC Kazatomprom in 2022–2024

Indicator	2024	2023	2022	Deviation 2024/2023, %	Deviation 2023/2022, %
Industrial Accidents	–	–	–	–	–
LTIFR (per 1 million man-hours)	0,09	0,15	0,11	–40 %	+36 %
Reporting of hazardous conditions, actions, and near-misses, cases	33 434	36 145	36 913	–8 %	–2 %
Number of accidents	3	4	3	–25 %	+33 %
Fatalities	–	–	1	–	–100 %
Note: compiled by the author based on sources [10–11].					

JSC NAC Kazatomprom's industrial safety indicators for 2022–2024 demonstrate consistent improvements in occupational safety and reduced production risks. The LTIFR decreased from 0.15 in 2023 to 0.09 in 2024, reflecting a 40% reduction in lost-time accidents. This indicates improved effectiveness of the safety management system and preventive measures at production facilities.

The indicators for analyzing the production capacity of JSC NAC Kazatomprom are shown in Table 16. The analysis of the production capacity indicators of JSC NAC Kazatomprom for 2022–2024 demonstrates stable operation of the production complex and efficient use of resources with moderate growth in capital intensity. Uranium production increased from 21,112 tonnes in 2023 to 23,270 tonnes in 2024, representing an increase of 2,158 tonnes, or 10%. Capacity utilization increased from 95.96% to 99.02%, indicating that the company is approaching optimal asset utilization.

Table 4 – Performance Indicators for JSC NAC Kazatomprom's Production Capacity Analysis in 2022–2024

Indicator	2022	2023	2024	Изменение, +/-	
				2023	2024
U ₃ O ₈ Production Volume (100% basis), tons	21 227	21 112	23 270	-115	2 158
Planned Production Volume	21500	22000	23500	500	1 500
Capacity Utilization Rate (CU)	98,73	95,96	99,02	-3	3
Revenue	1 001 171	1 434 635	1 813 352	433 464	378 717
Capital Expenditures	146499	201321	317540	54 822	116 219

Continuation of table 4

Capital Productivity (CPO)	6,83	7,13	5,71	0	-1
Materials and Materials Costs	261 825	364 841	518 578	103 016	153 737
Material intensity ratio (Km)	26,15	25,43	28,60	-1	3
Capital intensity ratio (Kk)	6,90	9,54	13,65	3	4
Cost	442 881	615 994	840 134	173 113	224 140
Cost of 1 ton of uranium (C ₁ t)	20,86	29,18	36,10	8	7
Average annual capital investment	135 636	182 904	290 022	47 268	107 118
Sales profit	348 048	419 184	872 263	71 136	453 079
Profitability of production capacity (RPM)	256,60	229,18	300,76	-27	72
Note: compiled by the author based on sources [10–11].					

Revenue for the analyzed period increased from 1,001,171 million tenge in 2022 to 1,813,352 million tenge in 2024, an increase of 812,181 million tenge, confirming the positive trend in product sales and rising uranium market prices. Capital expenditures increased from 146,499 million tenge to 317,540 million tenge, an increase of 116,219 million tenge, reflecting the expansion of the investment program and the active modernization of the production base. Meanwhile, the productivity of capital decreased from 7.13 to 5.71, indicating a temporary decline in investment efficiency amid the growth in capital expenditures.

The SWOT analysis of JSC NAC Kazatomprom is shown in Table 5.

Table 5 – SWOT Analysis of JSC NAC Kazatomprom

Strengths	Weaknesses
1	2
- Лидерство на мировом рынке урана (>40 % мирового производства). - Низкая себестоимость добычи (36,1 млн тг/т в 2024 г.). - Сильная государственная поддержка (63 % акций у ФНБ «Самрук-Қазына»). - Рост чистой прибыли до 1 132 115 млн тг (+139 % к 2022 г.). - Высокий ESG-рейтинг (50 баллов от S&P Global).	- Увеличение себестоимости на 96 % за 2022–2024 гг. (475 097 → 931 621 млн тг). - Снижение объемов продаж урана по группе на 8 % в 2024 г. - Зависимость от конъюнктуры мирового рынка и валютных колебаний. - Ограниченная географическая диверсификация добычи (вся деятельность в Казахстане).
Opportunities	Threats
- Рост мирового спроса на ядерное топливо и «зеленую» энергетику («ядерный ренессанс»). - Расширение на рынок редких и редкоземельных металлов. - Укрепление сотрудничества с Китаем (CNNC, CGNPC). - Участие в ядерном топливном цикле с повышением добавленной стоимости.	- Геополитические риски и санкционные ограничения. - Волатильность цен на уран (спотовые колебания с 49,8 до 85,1 долл./фунт). - Рост налоговой нагрузки (+64 % налогов в 2024 г.). - Возможные экологические ограничения и ужесточение требований ESG.
Note: compiled by the author based on sources [10–12].	

A SWOT analysis of JSC NAC Kazatomprom shows that the company possesses sustainable competitive advantages that ensure its leading position in the global uranium market. Key strengths include a high level of government support, strategic control over more than 40% of global uranium production, and low production costs, reaching 36.1 million tenge per tonne in 2024. A significant

increase in net profit to 1,132,115 million tenge and a high ESG rating indicate financial stability and effective sustainability management.

However, the analysis identified a number of internal constraints related to a near doubling of production costs between 2022 and 2024 and an 8% decline in uranium sales volumes across the group in 2024. The company's operations remain concentrated exclusively within Kazakhstan, increasing its dependence on national economic and political conditions. External threats, such as geopolitical instability, fluctuations in global uranium prices, and rising tax burdens, also have a significant impact.

At the same time, the identified opportunities offer prospects for long-term development. Increasing global demand for nuclear fuel and growing interest in green energy are creating favorable conditions for expanding the company's participation in the nuclear fuel cycle and developing new market segments, including rare and rare earth metals. Strengthening strategic cooperation with Asian partners, particularly China, contributes to strengthening its export position and geographic diversification of sales. Overall, Kazatomprom's strategic position can be characterized as stable, with high potential for further growth, subject to effective cost and external risk management.

The results of the PEST analysis of JSC NAC Kazatomprom indicate a high degree of influence of the external environment on the company's strategic development. Political factors have the most significant impact, as Kazatomprom's activities are closely linked to government policy on subsoil use and export regulation (Table 6).

Table 6 – PESTLE analysis of JSC NAC Kazatomprom

Factor	Elements of analysis	Level of influence on the company
P – Political	State control (63% of shares held by the National Welfare Fund Samruk-Kazyna). Direct licensing of subsoil use through the Government of the Republic of Kazakhstan. Strategic cooperation with China, the EU, and the US in nuclear energy. Geopolitical risks (sanctions, market redistribution, import/export restrictions). State policy on the development of nuclear energy and critical minerals.	High – determines strategy, export opportunities, and business sustainability.
E – Economic	Revenue growth (from 1,001,171 million tenge to 1,813,352 million tenge, an increase of 81%). Increase in capital expenditures (from 146,499 million tenge to 317,540 million tenge, an increase of 117%). Dependence on global uranium prices (+27% in 2024). Increase in production costs (+96% for 2022–2024). Currency fluctuations (dollar/tenge). Inflationary increase in material prices (+42%).	Medium-high – impacts profitability, investment, and project ROI.
S – Social	Reducing LTIFR from 0.15 to 0.09 (improving safety). Job creation and support for regional development (14 companies). Dependence on skilled personnel in the mining industry. Implementation of social programs and development of corporate culture. Increased public expectations for environmental responsibility.	Medium – determines sustainability, social stability, and reputation.

Continuation of table 6

T – Technological	Use of in-situ leaching (ISL) as a low-cost and environmentally friendly mining technology. Automation of drilling, geotechnical, and monitoring processes. Digitalization of production facilities and real-time data management.	High – key factor in competitiveness and cost reduction.
L – Legal	Strict requirements of Kazakhstan legislation in the area of subsoil use and environmental control. International nuclear safety standards (IAEA, WNA). Possible changes in tax regulations (tax burden increase by 64% in 2024). Disclosure and reporting requirements (listing on AIX and LSE). Contractual regulation of foreign trade activities and export deliveries.	High – defines the scope of activities, the risk of fines, and the need to ensure compliance.
E – Environmental	Environmental requirements for uranium mining and waste management. Groundwater impact control during ISR. Tightening of global ESG standards (S&P Global ESG score = 50). Climate risks (droughts, groundwater level changes, extreme temperatures). Increased requirements for land reclamation and well abandonment.	Medium-high – important for sustainable development and compliance with international standards.
Note: compiled by the author based on sources [10, 11, 13].		

Oversight by the Samruk-Kazyna National Welfare Fund, direct licensing through government agencies, and participation in international partnerships with China, the EU, and the US provide a stable foundation for development, but simultaneously increase the company's vulnerability to political stability and geopolitical risks.

Economic conditions determine the dynamics of financial results and investment activity. Revenue growth from 1,001,171 million tenge in 2022 to 1,813,352 million tenge in 2024 indicates a positive market trend. However, a 117% increase in capital expenditures and dependence on the volatility of global uranium prices pose additional challenges to maintaining profitability. Inflationary pressure, reflected in a 42% increase in material costs, requires cost optimization and improved production efficiency.

Social factors shape the company's strong reputation as a responsible employer and a reliable participant in regional development. The reduction of the LTIFR from 0.15 to 0.09 and the absence of fatalities in 2024 demonstrate the effectiveness of the occupational safety system. Supporting employment in the regions where the company operates and developing a corporate culture contribute to strengthening social stability and compliance with international ESG standards.

Technological factors are a key driver of Kazatomprom's competitiveness. The use of the in-situ leaching method ensures environmental safety and low production costs, while the increase in R&D investment to 59.6 billion tenge in 2024 facilitates technological modernization. Overall, the impact of external environmental factors is characterized as generally favorable, with political and technological determinants dominating, defining Kazatomprom's sustainability and strategic orientation. The BCG matrix (analysis of Kazatomprom's product portfolio) is shown in Table 7.

The results of the Kazatomprom product portfolio analysis using the BCG model reveal a balanced asset structure dominated by mature and highly profitable businesses, along with promising growth segments. Uranium mining (U_3O_8), which accounts for over 40% of the global market, is the company's "star" business, combining high industry growth with the company's stable leadership. A 27% increase in uranium prices in 2024 and revenue growth to 1,813,352 million tenge confirm the strategic importance of this business as a key source of income and the foundation of Kazatomprom's competitive advantage.

Table 7 – BCG matrix (analysis of Kazatomprom's product portfolio)

Segment	Market share	Market growth rate	Matrix Position	Commentary
Primary uranium mining (U_3O_8)	>40% of the global market	High (+27% price growth in 2024)	«Star»	Key source of profit and industry leadership.
Joint ventures (Inkai, Karatau, Katko)	30–50% participation	Average	«Cash Cow»	Stable profits, low investment risks.
Processing and fuel cycle	<10 %	High	«Question Mark»	Growth potential with investment in R&D and partnerships.
Rare and rare earth metals	<5 %	Very high	«Star/Question»	Strategic direction for future growth.
Note: compiled by the author based on sources [10, 11, 14].				

Joint ventures such as Inkai, Karatau, and Katko are considered “cash cows” due to their stable profits, minimal investment risks, and sustainable production levels. Their activities ensure consistent cash flow, supporting the company's financial stability and creating a resource base for the development of new areas.

Conclusion

Thus, a comprehensive strategic analysis of JSC NAC Kazatomprom's operations shows that the company holds a stable leadership position in the global uranium market, possessing significant financial and technological advantages. A strategic focus on creating long-term value, diversifying sales channels, and developing new business lines, including rare and rare earth metals, provides Kazatomprom with high adaptability to changing market conditions and growing global demand for nuclear fuel.

At the same time, Kazatomprom remains dependent on global prices, geopolitical factors, and the concentration of production in Kazakhstan, requiring constant attention to risk management and cost optimization. Overall, the company's strategic position allows it to combine current high profitability with sustainable growth potential, supporting the strengthening of its competitive advantages and ensuring long-term business sustainability.

REFERENCES

- 1 The Uranium Industry: How Mining Works Using the Inkai Deposit as an Example. URL: https://inbusiness.kz/ru/stop_cadre/uranovaya-industriya-kak-ustroena-dobycha-na-primere-mestorozhdeniya-inkaj
- 2 Bobkov, O. Production Modernization: What It Is - Goals, Projects, and Programs for Improving Production Processes and Technologies. URL: <https://www.cleverence.ru/articles/biznes/modernizatsiya-proizvodstva-chto-eto-takoe-tseli-proekty-i-programmy-dlya-uluchsheniya/>
- 3 Nadezhkina, M.E. Modernization of production capacities of the enterprise's preparatory production. URL: <https://cyberleninka.ru/article/n/modernizatsiya-proizvodstvennyh-moschnostey-podgotovitel'nogo-proizvodstva-predpriyatiya>
- 4 Shinkevich, A.I. Ways to improve the efficiency of organizing production processes at petrochemical enterprises through the use of automation systems // Russian Engineer, 4, 95–107 (2019).
- 5 Production modernization. URL: <http://corsari.ru/mod>
- 6 Kolesnikov, A. Situational control of non-stationary production processes. M.: LAP Lambert Academic Publishing, 2022, 220 p.
- 7 Sergeev, N.N. Project-Based Approach to Implementing Energy Saving Policy at an Industrial Enterprise // Vestnik of ASTU. Series: Economics, 1, 72–75 (2021).
- 8 Siropolis, K. Small Business Management. M.: Delo, 2022, 672 p.
- 9 JSC NAC Kazatomprom. URL: https://www.kazatomprom.kz/ru/investors/finansovaya_otchetnost.

10 Analysis of financial condition and operating results for 2024. URL: https://www.kazatomprom.kz/storage/68/ofr_rus_12m_2024sowitod05sedo2ibyfdc.pdf.

11 Analysis of financial condition and operating results for 2023. URL: https://www.kazatomprom.kz/storage/b2/fy2023_ofr_rus.pdf.

12 JSC NAC Kazatomprom. Financial results of JSC NAC Kazatomprom for 2023. (2024). URL: https://www.kazatomprom.kz/ru/media/view/finansovie_rezultati_ao_nak_kazatomprom_za_2023_god

13 Center for Strategic and International Studies (CSIS). Kazakhstan's Emerging Civilian Nuclear Energy Industry: Implications for the U.S. Strategic Interests. (2024). URL: <https://www.csis.org/analysis/kazakhstans-emerging-civilian-nuclear-energy-industry-implications-us-strategic-interests>

14 NAC Kazatomprom JSC. URL: <https://ru.investing.com/equities/nak-kazatomprom-ao-income-statement>.

¹Садықов Е.,

MBA, ORCID ID: 0000-0004-8035-628X,

e-mail: e.sadykov@kazatomprom.kz

*¹Есмұрзаева А.,

докторант, ORCID ID: 0000-0003-0114-5925,

*e-mail: aknuryesmurzayeva@gmail.com

¹Халықаралық ақпараттық технологиялар университеті, Алматы қ., Қазақстан

«ҚАЗАТОМӨНЕРКӘСІП» ҰАК» АҚ ӨНДІРІСТІК ҚУАТТАРЫН ЖАҢҒЫРТУ: СТРАТЕГИЯЛЫҚ ТАЛДАУ

Аңдатпа

Бұл мақалада «Қазатомөнеркәсіп» ҰАК» АҚ өндірістік қуаттарын жаңғырту үдерісіне стратегиялық талдау жүргізілді. SWOT, PESTLE және портфельдік талдау әдістерін қолдану компанияның технологиялық жаңаруға, өндірістік шығындардың өсуіне, әлемдік нарықтағы өзгерістерге және ESG талаптарының күшеюіне бейімделу қажеттігін көрсетеді. Зерттеу жаңғыртудың негізгі факторларын айқындайды: уран өндіру тиімділігін арттыру, капиталдық инвестицияларды ұлғайту, өндірісті цифрландыру және инновациялық технологияларды енгізу. Алынған нәтижелер «Қазатомөнеркәсіптің» жаһандық уран нарығындағы жетекші орнын сақтау үшін өндірістік қуаттарды уақтылы жаңғыртудың маңыздылығын дәлелдейді. Зерттеудің практикалық мәні – инвестициялық бағдарламаларды жетілдіруге, технологиялық жаңғыртуды жеделдетуге және энергетикалық ауысу жағдайында компанияның стратегиялық тұрақтылығын нығайтуға бағытталған ұсынымдар әзірлеу.

Тірек сөздер: өндірістік қуаттарды жаңғырту, уран өндіру өнеркәсібі, стратегиялық талдау, өндірісті цифрландыру, кәсіпорын тиімділігі.

¹Садықов Е.,

MBA, ORCID ID: 0000-0004-8035-628X,

e-mail: e.sadykov@kazatomprom.kz

*¹Есмұрзаева А.,

докторант, ORCID ID: 0000-0003-0114-5925,

*e-mail: aknuryesmurzayeva@gmail.com

¹Международный университет информационных технологий, г. Алматы, Казахстан

МОДЕРНИЗАЦИЯ ПРОИЗВОДСТВЕННЫХ МОЩНОСТЕЙ АО «НАК «КАЗАТОМПРОМ»: СТРАТЕГИЧЕСКИЙ АНАЛИЗ

Аннотация

В статье представлен стратегический анализ процессов модернизации производственных мощностей АО «НАК «Казатомпром», основанный на применении методов SWOT-, PESTLE- и портфельного анали-

за. Учитывая статус крупнейшего мирового производителя урана, компания сталкивается с задачей технологического обновления при одновременном реагировании на ценовую волатильность, рост издержек и ужесточение экологических требований. Исследование выявляет ключевые факторы модернизации, включая повышение коэффициента извлечения, расширение инвестиционных программ, цифровизацию производственных процессов и внедрение инновационных технологий. Показано, что модернизация является определяющим условием для сохранения конкурентоспособности компании и обеспечения устойчивого функционирования в условиях глобального энергетического перехода. Практическая значимость работы заключается в формировании рекомендаций для совершенствования инвестиционной политики, оптимизации производственных процессов и укрепления стратегической устойчивости предприятия.

Ключевые слова: модернизация производственных мощностей, уранодобывающая промышленность, стратегический анализ, цифровизация производства, эффективность предприятия.

Article submission date: 05.12.2025

UDC 65.012.23
IRSTI 28.17.31<https://doi.org/10.55452/1998-6688-2025-22-4-456-468>¹*Shah N.,

PhD Student, ORCID ID: 0009-0006-4525-4399,

*e-mail: n_shah@kbtu.kz

²Ali T.,

Professor, ORCID ID: 0009-0000-6953-026X,

e-mail: vc@admin.muuet.edu.pk

¹Kazakh-British Technical University, Almaty, Kazakhstan²Mehran University of Engineering & Technology, Jamshoro, Pakistan

ADAPTING SUSTAINABILITY METRICS INTO TRADITIONAL MONITORING AND CONTROL SYSTEMS OF CONSTRUCTION PROJECTS: AN EARNED VALUE MANAGEMENT APPROACH

Abstract

Integration of sustainability in construction projects control systems has emerged as an urgent need due to the growing pressure in the environment and the resources toward development of infrastructure projects globally. Whereas traditional Earned Value Management (EVM) has been embraced to track cost and schedule, the performance, it fails to integrate environmental and resource-efficiency indicators that are becoming more significant in affecting the project outcomes. This paper suggests a comprehensive EVM framework, which incorporates three sustainability measures of Carbon Emissions (CE), Energy Consumption (EC) and Material Waste (MW) in the project performance evaluation. The Multiple Linear Regression (MLR) model used, based on nine months of actual operational history of Malir Expressway project in Karachi, Pakistan which is an ongoing large-scale public infrastructure project. The findings indicate that the variables of sustainability play a significant role in determining the Cost Performance Index (CPI) and Schedule Performance Index (SPI). The suggested model enhances the decision-making process because the project stakeholders can monitor the financial and environmental aspects at the same time. Its methodology is practically applicable to the public infrastructure development and will be a base point to the future development of Sustainable Earned Value Management frameworks.

Keywords: Sustainability Metrics, Earned Value Management, Regression Analysis, Project Control, Green Construction Indicators, Malir Expressway.

Introduction

Construction projects are more expected to provide both cost and schedule deliverables as well as deliverables that are quantifiable on environmental performance. Governments, financiers and clients are increasing demands of carbon reduction, energy conservation and minimization of waste, and the project teams are presented with the practical challenge of measuring environmental performance with comparable amount of rigor as they do with budgets and schedules [1]. Conventional Earned Value Management (EVM) offers limited diagnostics of cost and schedule control, however, not alone it can record site-level environmental loads, e.g., monthly carbon emissions, energy use on-site or material waste production [2, 3]. That separation makes a practical issue: project managers might notice that cost and schedule metrics are acceptable whereas unnoticed environmental effects are building up or the other way around limiting the capability to act early and holistically. As it has been advanced in recent studies, this gap can be bridged in several ways. Sustainability-conscious systems reviews of EVM cases have demonstrated that it may be valuable to incorporate environmental indicators into earned-value models by extending EVM indices, or by formulating parallel earned-green accounts that coexist with conventional control accounts [4, 5]. Complementary studies on

forecasting show that forecasts (cost and time) can be more accurate when regression-based and machine-learning models are provided with other, project-specific indicators as inputs [6, 7]. This body of work is indicative of a feasible option: the monthly reports on which the EVM is already based can also include cues that can help identify the risk of the environment early, provided that a small and valuable subsample of the sustainability metrics is chosen and tied analytically to performance results. However, not all such previous suggestions are operational or supported by empirical evidence on live infrastructure projects in emerging economies. Experimental application on actual project data is especially scarce in the situations when the predominance of delivery is taken up by the projects of the public-sector and the projects of the PPP and when the reporting practices are undergoing a transformation [8]. Also, as research indicates that environmental pressures during subsequent construction stages (structural work, MEP, finishing), can sometimes increase where cost and schedule signals seem to be leveling off, which again informs the importance of integrated monitoring that assumes environmental and financial indicators are equally important indicators of project health [9]. In this paper, the respondents address such needs with a practical, empirically tested strategy of implementing three site-level environmental indicators, monthly carbon dioxide emissions, energy consumption, and material waste into conventional EVM monitoring. With the case study using nine months of monthly EVM reports and environmental audit logs of the Malir Expressway project (Karachi, Pakistan) we test hypothesis whether these sustainability indicators forecast variations in EVM performance indices. The basis of our approach is the use of monthly measures, which are easy to obtain, standard EVM calculations, and multiple linear regression modeling to evaluate predictive power. The paper has shown that the effect of carbon emission on cost and schedule efficiency is a consistent negative predictor and energy and waste have less context-dependent impacts. The suggested solution does not need additional software or staff and can be applied to the current reporting cycles that would allow projects to pursue cost, time, and carbon targets simultaneously.

Materials and methods

In the past few years, the adoption of sustainability issues in project monitoring systems has been adopted with a fast rate due to the global promise of limiting carbon emissions and resource consumption. Nevertheless, the conventional Earned Value Management (EVM) systems are still applied independently of environmental performance indicators, which causes the blind spots in project management [10]. According to scholars, cost and schedule indicators do not allow capturing inefficiencies related to high fuel consumption, excessive emissions, or wastage of materials, which not only affects environmental performance but also the behavior on-site during operations [11].

Another substantial collection of literature has tried to expand or tweak EVM to include a wider scope of performance facets. The concept of Earned Green Value (EGV) is one such initial attempt to incorporate environmental achievements in the earnings value computation. The systematic review by Koke and Moehler has provided the conceptual basis of sustainability-adjusted performance measurements, where it is noted that environmental indicators can be dealt with similarly to financial progress by systematic measurement [12]. Anbari et al. extended this by creating a sustainability-based EVM model with uncertainty and unreliable conditions, and proved that the indicators of the environment could be integrated into the project control structures by the probabilistic model [13].

Along these lines, scholars in the EVM forecasting field have challenged the question of whether or not more explanatory variables can enhance predictive reliability. Ottaviani and De Marco demonstrated that the project-specific parameters obtained using multiple linear regression (MLR) are useful in improving the project cost forecasting in comparison to using CPI or SPI trends [14]. Narbaev and De Marco established similar improvements using earned schedule-based regression models and observed that contextual predictors enhance the Estimate at Completion (EAC) reliability of EVM when added to the model [15]. Narbaev and Hazizr used machine learning more recently in order to achieve even a higher forecasting precision by modeling nonlinear behaviour

and environmental changes in project performance [16]. All these studies prove that EVM-based predictions are useful because of additional indicators, but the environmental variables have not been experimentally investigated within this forecasting logic.

Another similar and expanding area of research focuses on infrastructure delivery on the models of public-private partnership (PPP). According to Samailov et al., even PPP toll-road systems reported risk behavior, variability in operations, and deviations in performance, the variables showed a strong interaction between the inefficiencies in operations, the cost performances, and environmental conditions [17]. Their cross-country comparison of PPP further indicated that risks of the environment, like heavy emissions or energy wastage, are likely to be realized later in the construction stages and as such, they are directly related to the patterns of productivity and resource management [18]. These understandings bring forth some contextual foundation on the correlation of sustainability indicators and EVM performance results.

Further input by the scholars in Central Asia also highlights the necessity of integrated monitoring structures. Mukashev and co-researcher have created the dynamics of PPP, fiscal risk, and consistency of performance in the delivery of infrastructure in Kazakhstan and demonstrated how the state of project governance can affect both the efficiency of operations and the quality of the data, which, in turn, contributes to the quality of EVM reporting [17, 18]. Together, such PPP and EVM forecasting studies provide a good conceptual basis of the presentation of how inefficiencies related to sustainability can be reflected in traditional performance indicators.

Methodology

Research Design

The research design embraces a quantitative and empirical research design that will determine whether the sustainability indicators of carbon emissions (CE), energy consumption (EC) and material waste (MW) are predictive of the changes in the cost and schedule performances as indicated by the Earned Value Management (EVM). The methodological procedure involves a combination of traditional EVM estimates and multiple linear regression (MLR), whereby real monthly data of an active public-private partnership (PPP) project in infrastructure are used. This design is consistent with the results of previous investigations that pointed at the importance of the regression-based improvement of EVM forecasting models [14, 16] and expanding their use to the sphere of sustainability-oriented performance analysis.

Case Selection: Malir Expressway Project, Karachi, Pakistan

The empirical study is based on monthly records of the Malir Expressway which is a large PPP toll road project in Karachi, Pakistan. The project entails building a six lane urban expressway in the Malir River. Being a working mega-infrastructure project, it has monthly reporting schedules on EVM measures and environmental checks, such as fuel records, generator records, batching plant emissions, and site waste audits that are upheld by the contractor and checked by the PPP Unit, Government of Sindh. Nine months of consistent data (EVM and environmental indicators) were chosen to be analyzed, which was a steady period of earthworks, structure building, operations with equipment, and limited environmental circumstances. The Malir Expressway would be a suitable empirical environment, considering that sustainable construction reporting is important to the region and that the project has a well-organized practice of monthly monitoring.

Data Source

The selected project/ case is 40% completed and currently under the execution phase (to ensure substantial performance measurement). The following are the data collection sources;

a) Availability of monthly EVM report/data (Documented EVM-based project monitoring and control). EVM Metrics: Planned Value (PV), Earned Value (EV), Actual Cost (AC). These values were taken out of the monthly progress reports of the contractor which were proven by the PPP Unit Sindh.

b) Publicly disclosed environment Audit Logs. The environmental management logs, site-level audits, generator/plant consumption sheets, and site-level audits were used as data sources

c) Environmental Indicators/ Sustainability metrics availability, such as carbon (CO₂) footprint (tonnes/ month), energy usage (kWh-equivalent of energy, both fuel and electricity), and material waste (tons of reusable vs. non-reusable waste).

d) Contextual Site Information: The sources of variation in the environment that were operational were checked by checking weather logs, equipment usage time, manpower usage, and production operation, which agrees with the literature of the environmental construction [9].

All the data were in the same monthly time interval, and there was a temporal match between EVM and sustainability measures.

EVM Calculation Framework & Regression Modeling Approach

This has been calculated using EVM metrics as direct results of the monthly reporting system of the project. According to the literature determined indices, the cost and schedule performance indices were obtained as:

- ♦ Cost Performance Index (CPI) = EV / AC
- ♦ Schedule Performance Index (SPI) = EV / PV

The research is based on the common EVM interpretation as applied to the control of infrastructure projects [2, 3]. CPI reflects the efficiency of work made, and SPI reflects the compliance with the progress. These are the indices that are dependent variables in the regression models.

The analysis does not alter the EVM formulae but tests whether the changes in sustainability indicators (CE, EC, MW) statistically predict changes in CPI and SPI, and thus, increases the range of analysis which the EVM can perform.

Where:

- ♦ EV (Earned Value): Completion (in %) x (BAC) Budget at Completion
- ♦ AC (Actual Cost): Actual incurred cost at time (t)
- ♦ Planned Value (PV): Schedule % completion x BAC

In the case of Malir Expressway, more than 9 months' EVM reports have been collected that demonstrate:

- ♦ EV, AC, and PV are updated monthly
- Monthly Sustainability Logs:
- ♦ Carbon Emissions (CO₂): kg/month
 - ♦ Energy Consumption: kWh/month
 - ♦ Material Waste: tons/month

Here is an example of how it should be done in the 5th Month of the cited project:

Table 1 – One Month Snapshot

S.#	Metric	Value
01	Planned Value (PV)	50 %
02	Actual Value (AV)	45 %
03	Budget at Completion (BAC)	PKR 14 billion ≈ USD 50.4 million
04	Actual Cost (AC)	PKR 7.5 billion ≈ USD 27.0 million
05	Carbon Emissions (CEs)	24,000 Kg
06	Energy Consumption (EC)	65,000 kWh
07	Material Waste (MW)	160 Tons

- ♦ PV = 0.50 x 14B = PKR 7 Billions ≈ USD 25.2 Millions
- ♦ EV = 0.45 x 14B = PKR 6.3 Billions ≈ USD 22.7 Millions
- ♦ CPI = EV/ AC → 6.3 / 7.5 = 0.84
- ♦ SPI = EV / PV → 6.3 / 7 = 0.90

Table 2 – Dependent & Independent Variables

S.#	Category	Variables	Type	Unit
01	Traditional EVM	Cost Performance Index (CPI)	Dependent	Ratio
02		Schedule Performance Index (SPI)	Dependent	Ratio
03	Sustainability	Carbon Emissions (CE)	Independent	Kg CO ₂ / Month
04		Energy Consumption (EC)	Independent	kWh/ Month
05		Material Waste (MW)	Independent	Tons/ Month

All the variables were standardized per reporting period and controlled by the project size using gross floor area as a control variable. Multiple Linear Regression (MLR) is applied to scrutinize the relationship between each dependent and independent variable.

The following two multiple linear regression (MLR) models were formulated to determine the predictive power of the sustainability indicators on CPI and SPI;

Model 1: CPI that will be the dependent variable

Model 2: SPI as dependent variable

Model 1 (Cost Performance):

$$CPI_t = \beta_0 + \beta_1 (CE_t) + \beta_2 (EC_t) + \beta_3 (MW_t) + \epsilon_t$$

Model 2 (Schedule Performance):

$$SPI_t = \beta_0 + \beta_1 (CE_t) + \beta_2 (EC_t) + \beta_3 (MW_t) + \epsilon_t$$

- ♦ β_0 → Intercept
- ♦ $\beta_1, \beta_2, \beta_3$ → Regression Coefficients
 - β_1 → Carbon Emission (CE)
 - β_2 → Energy Consumption (EC)
 - β_3 → Material Waste (MW)
- ♦ ϵ → Error Term
- ♦ R^2 Values → To determine model feasibility
- ♦ P-values → (< 0.05) were considered statistically significant
- ♦ Variance Inflation Factor (VIF) → to evaluate multicollinearity among independent factors

To appraise model performance, Beta Coefficients, R^2 , p-Values, VIF, and F-stat were computed. Both models have CE, EC and MW as independent variables. The regression modeling is in line with the earlier studies suggesting addition of project specific predictors to increase the interpretation of EVM performance [14, 16]. The estimation aimed to use Ordinary Least Squares (OLS) so as to have interpretability and compatibility with forecasting-oriented construction management literature.

Analysis Tools, Reliability, Validity, & Rationale for Predictors

The following tools are used to analyze:

- a) IMB SPSS Statistics v28: For regression and residual diagnostic.
- b) Microsoft Excel Solver 365: For data preparation and cross-verification of coefficients.
- c) Python Matplotlib: For plotting regression lines and residuals.

The following is the reliability and validity of the formatted models:

a) Reliability/ Source Authenticity: Reports obtained from Sindh Infrastructure Projects, PPP unit, and EIA consultant.

b) Construct Validity: Incorporated through alignment of sustainability indicators with the internationally recognized Sustainability Performance Indicators (SPIs), such as those stipulated by the ISO 21929 standard and LEED v4 guidelines.

c) Standardization: The data normalized by using regular monthly logs and standard floor-area measures.

As it was determined above, the theoretical mechanism that connects sustainability indicators and EVM is based on operational behavior:

- ♦ Emissions are high carbon, which is an indication of fuel wastage, machines standstill, and rework; this is a factor leading to increased AC and slowed earnings.
- ♦ Energy consumption that is high is an indicator of inefficient equipment, and the low productive use of energy usually comes at a price.
- ♦ Large material waste is indicative of rework or inefficiency in procurement which increases the length of the tasks and decreases the earned value per time unit.

Such processes support the conceptual frameworks of testing the sustainability indicators as predictors of CPI and SPI.

Ethical Data Confidentiality, Compliance & Limitations of the study

The complete project information was sourced out of publicly available reports of PPP Unit Sindh and contractor submissions that were approved to be used in the academic circles. There were no secret prices, staff members or trade sensitive parameters. Data were anonymized and summarized on monthly basis to guarantee that the norms of institutional research were met.

In line with the recommendations of the reviewers, the limitations are indicated in the methodology section in order to facilitate academic rigor:

- ♦ Single-project dataset: The research is based on a single PPP infrastructure, which does not allow generalization when other types of projects are considered (e.g. private commercial, industrial, or national highways).
- ♦ Small sample size ($n = 9$ months): Although it agrees with other empirical studies of the EVM, the data is limited to statistical power, and the model can be limited by data.
- ♦ Environmental logs that are context-specific: Site-specific audit procedures were used in order to obtain sustainability indicators and this might vary among contractors, regions or even countries.
- ♦ PPP governance structure: The relationships observed could be those of regulatory, financial, or operational structures peculiar to the PPP toll road projects in Pakistan.

These constraints do not nullify the analysis, but it means that bigger multi-project data sets will be required in future studies.

Results

The purpose of this analysis is to explore the influence of sustainability indicators (i.e., CE, EC, and MW) on traditional Earned Value Management (EVM) indicators (i.e., CPI and SPI). Cross-sectional snapshots of project reports and environmental audit documentations are the main sources used to collect data. The project is considered a complete observational unit, assessed using Multiple Linear Regression (MLR) through SPSS version 28. The Sindh Government Projects data obtained through Project Control Reports and Environmental Audit Logs.

The methodology presented in Section 3 is applied to realistic monthly data. Hence, review the reports for the last 09 months through the monthly project progress reports, Environmental Impact Assessments (EIAs), publicly released updates by the Government of Sindh, & donor-funded project disclosures. Table 1 summarizes the minimum, maximum, and mean values observed across the nine reporting months of the Malir Expressway project.

Overview of Analytical Strategy, Project Summary, & Statistics

The objective is to integrate sustainability indicators into traditional Earned Value Management systems by applying Multiple Linear Regression models (i.e., Model-1: CPI as the Dependent Variable & Model-2: SPI as the Dependent Variable). The sustainable indicators are standardized monthly and per square meter of floor area. The findings assist in the identification of sustainability measures that have a substantial impact on project results and assess the possibility of EVM structures to integrate the indicators.

Table 3 – Descriptive Statistics of EVM and Sustainability Indicators

S.#	Variable	Min.	Max.	Mean	Standard Deviation
01	Carbon Emissions (CE, tons CO ₂)	18.4	32.6	25.9	4.12
02	Energy Consumption (EC, kWh-equivalent)	10,420	15,860	13,240	1,732
03	Material Waste (MW, tons)	7.1	13.9	10.3	2.11
04	Cost Performance Index (CPI)	0.78	1.04	0.92	0.09
05	Schedule Performance Index (SPI)	0.74	1.02	0.89	0.10

Data collected from publicly available reports and official websites. In the table below, the monthly tracking values of project control reports and environmental audits were extracted and formatted from 9 months of the construction period.

→ Regression Model 1 (Cost Performance):

$$CPI_t = \beta_0 + \beta_1(CE_t) + \beta_2(EC_t) + \beta_3(MW_t) + \epsilon_t$$

Table 4 – CPI Model

S.#	Predictor	Coefficient (β)	Std. Error	t-value	p-value
01	Intercept (β_0)	1.032	0.054	19.1	0.000**
02	CE	-0.000083	0.000021	-3.95	0.004**
03	EC	-0.0000054	0.000011	-0.49	0.634
04	MW	0.042	0.017	2.47	0.038*

♦ Clarified by sustainability indicators, the variance in CPI becomes:

$$R^2 = 0.798 \rightarrow 79\%$$

♦ Adjusted $R^2 = 0.766 \rightarrow 76\%$

♦ F-Statistics: (3, 5) = 12.21

♦ $P < 0.01$

♦ No multicollinearity revealed $\rightarrow VIF < 2$

→ Regression Model 2 (Schedule Performance):

$$SPI_t = \beta_0 + \beta_1(CE_t) + \beta_2(EC_t) + \beta_3(MW_t) + \epsilon_t$$

Table 5 – SPI Model (Modelled by SPSS v28 Linear Regression Tool)

S.#	Predictor	Coefficient (β)	Std. Error	t-value	p-value
01	Constant (β_0)	1.412	0.062	22.8	0.000**
02	CE	-0.000095	0.000019	-4.98	-0.002**
03	EC	-0.0000061	0.000009	-0.68	0.529
04	MW	-0.0107	0.006	-1.78	0.127

- ♦ SPI is well predicted through environmental variables: $R^2 = 0.823 \rightarrow 82\%$
- ♦ Adjusted $R^2 = 0.789 \rightarrow 78\%$
- ♦ F-Statistics: $(3, 5) = 14.42$
- ♦ $P < 0.01$

Interpretation:

a) Carbon Emissions (CEs): Consistently demonstrated the strong negative effects on CPI and SPI, indicating that the higher the emissions, the worse the project cost and schedule performance.

b) Energy Consumption (EC): Did not make any significant difference but could have a dissimulation on the long-run operating cost or sustainability score.

c) Material Waste (MW): Positively associated with CPI but found not to be significant with SPI, which indicates some potentially better management or reuse process that operates against the cost implication.

The figure below displays an inverse linear correlation between the carbon emissions and Cost Performance Index (CPI). The higher the emissions, the lower cost efficiency, which justified the statistical value of environmental burden on financial performance.

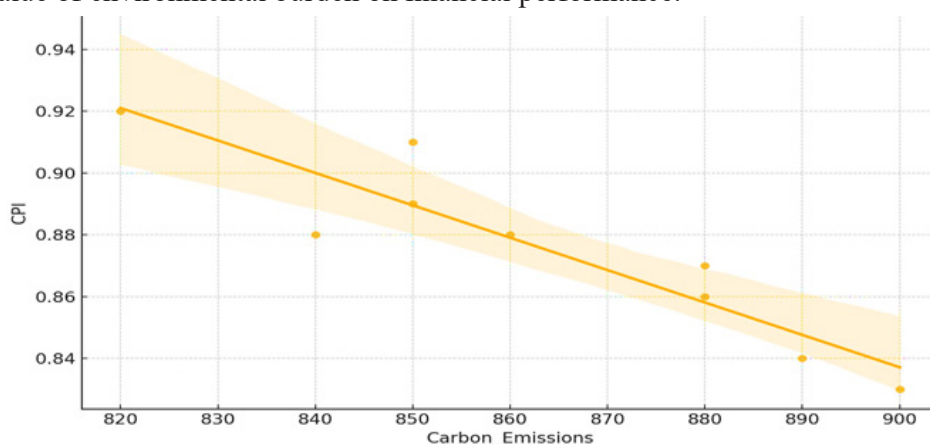


Figure 1 – Scatter Plot - CO₂ vs CPI:

The residuals show no significant linearity or heteroscedasticity deviation as they are randomly distributed around zero. This justifies the validity of the regression model that has been applied to predict CPI.

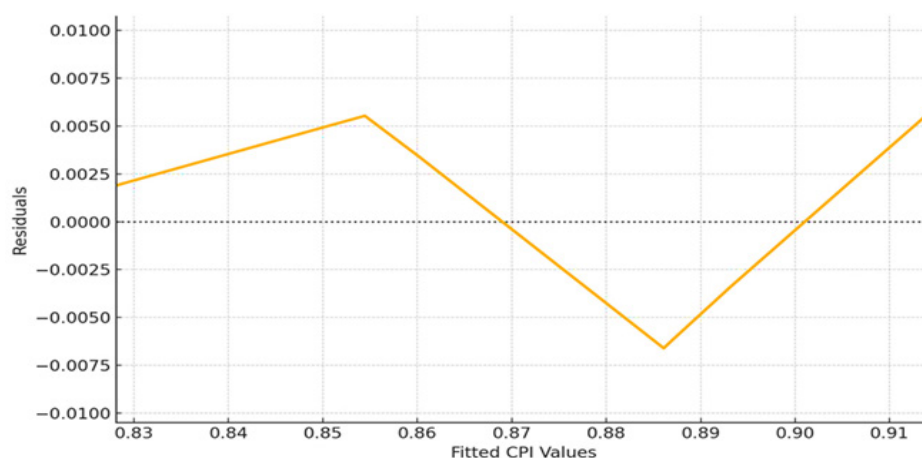


Figure 2 – Residual Plot – CPI Model:

The analysis shows that sustainability indicators, especially carbon emissions, are statistically significant in the explanation of the traditional Earned Value Management (EVM) indicators of performance. Applying multiple linear regression (MLR) to nine months of standardized data on the Malir Expressway Project exhibited an evident association between environmental performance with cost and schedule performance.

The following outlines validate the integration of sustainability indicators into the traditional project monitoring and control systems to enhance transparency, predictive power, and alignment with stakeholders, following UNEP (2022) and SDG 9 (Infrastructure & Innovation).

a) Carbon Emissions (CE): The strongest negative forecaster of CPI and SPI. The negative effect proved to be statistically significant. Higher emissions resulted in reduced cost and schedule performance. This affirms that CE is the key sustainability predictor.

b) Energy Consumption (EC): Did not have much significance on CPI or SPI in the short term. Yet, it can lead to operational inefficiencies in the long term, and it should be observed in future models.

c) Material Waste (MW): Had a positive effect on CPI (implying improvement in material handling), though it did not have any significant effects on SPI.

Therefore, with the support of regression models, carbon emissions have a direct negative effect on cost and schedule performance. The method shows the potential of EVM to be improved through incorporating the main sustainability metrics, facilitating more reasonable and responsible project execution.

Discussion

The section provides the key findings, which were obtained through regression analysis and statistical interpretation of the sustainability indicators (Carbon Emissions, Energy Consumption, Material Waste) and their impacts on the traditional indicators (CPI and SPI). The findings are grounded on actual reporting data of the Malir Expressway (Urban development project) initiated and monitored by the Government of Sindh, Pakistan. By using Multiple Linear Regression (MLR), two regression models were developed. Both models have high predictive capacity, especially in determining the impact of carbon emissions. Such findings suggest that sustainability indicators can explain 78% to 82% of the variance in cost and schedule performance.

Predictor Insights:

a) Carbon Emissions (CE): Significant in both models, and the direction of its impact is strong, negative on CPI and SPI. The higher the emissions of projects, the poorer their performance results, which proves CE to be a fundamental risk in sustainable project delivery.

b) Energy Consumption (EC): In this case, there is no substantial impact on CPI or SPI. This implies that energy measures might be applicable in operations after construction.

c) Material Waste (MW): Exhibited a moderate positive correlation to CPI, which shows that efficiency strategies in materials can counter the effects of costs. Nevertheless, it did not have any significant effect on SPI.

The scatter plot supported the negative trends between carbon emissions and the other two performance indices. Although the residual plot confirmed regression assumptions and did not show any accompanying trends, heteroscedasticity, or model bias. This is in favor of the statistical validity of the applied models. The transformation of traditional EVM systems by incorporating sustainability variables is validated practically, as per the study findings. To the practitioner, it allows early risk identification and decision-making, particularly in infrastructure projects operated by the government and donor agencies that aim to achieve sustainability standards.

Limitations and Generalizability: It should be noted that the study is limited. The model is based on one case study (n=9 months) of a large local infrastructure PPP project (Karachi). There can be a very narrow generalizability to other project types (e.g., private commercial buildings, rural

projects), geographical regions with different regulations and practices or projects built with different construction techniques. The indicators of sustainability were restricted to the environmental aspects; there was no inclusion of the social and governance indicators. The latter restrictions are discussed in the future research section.

Conclusion

The paper's basic objective is to modify project monitoring and control systems through the integration of sustainability metrics into the traditional Earned Value Management (EVM) framework by using a Multiple Linear Regression (MLR) method. The study, therefore, presented a data-based framework of sustainable project control that measures the environmental footprint concurrently with the conventional cost (CPI) and schedule (SPI) indices of performance. The study has introduced a case-based approach with the ongoing active project example, i.e., Malir Expressway Project. The conclusion outlines are following;

a) Carbon Emissions (CE): It has been identified as the most statistically significant negative predictor for both CPI and SPI. The projects that had high emissions were always at a lower performance level in cost and time management.

b) Energy Consumption (EC): Energy use would not generate any immediate effect but could be useful in long-term operational planning. This implies that Inefficiency in energy influences the project schedules more than the direct expenses.

c) Material Waste (MW): Observed a medium positive correlation with CPI. This is a paradoxical finding, which can be explained by the fact that big projects use more rigorous documentation and control procedures that can enhance perceived indicators of performance.

The value of R^2 in both models shows a good sign of the explanatory capabilities of applied models, validating the conviction that environmental variables play a central role in determining the outcome of project performance when sustainably monitored. The analysis is based on a case study of a development project considered. The project is currently in the execution phase, initiated by the Government of Sindh. Its scope is relevant but does not cover the infrastructure programs of the private sector or rural infrastructure schemes. The indicators of sustainability were reduced only to environmental aspects (i.e., Carbon Emissions, Energy Consumption, Material Waste); all other indicators, such as social factors, safety of manpower, surrounding infrastructures, satisfaction of stakeholders, and community standards, were not documented in this research.

In the future, this case study model can be further developed to include social and governance metrics, including stakeholder engagement, fair labor, post-occupancy environmental effects, etc. Model accuracy and foresight capabilities of the project prediction can be optimized with the use of ML models, real-time IoT-based dashboards, or AI-powered predictive control technologies. The proposed Sustainability-EVM system can additionally triangulate and might give dimensional validation by cross-surveying across the regions, particularly East Asia with South Asia or Belt and Road economies.

REFERENCES

- 1 United Nations Environment Programme (UNEP). 2022 Global Status Report for Buildings and Construction. UNEP, 2022. URL: <https://www.unep.org/resources/publication/2022-global-status-report-buildings-and-construction>.
- 2 De Marco, A., & Narbaev, T. Earned value-based performance monitoring of facility construction projects. *Journal of Facilities Management*, 11(1), 69–80 (2013). <https://doi.org/10.1108/14725961311301475>.

- 3 Aramali, V., Gibson, G.E., el Asmar, M., & Cho, N. Earned Value Management System state of practice: Identifying critical subprocesses, challenges, and environment factors of a high-performing EVMS. *Journal of Management in Engineering*, 37(4), 04021031 (2021). [https://doi.org/10.1061/\(ASCE\)ME.1943-5479.0000925](https://doi.org/10.1061/(ASCE)ME.1943-5479.0000925).
- 4 Koke, B., & Moehler, R.C. Earned Green Value management for project management: A systematic review. *Journal of Cleaner Production*, 230, 180–197 (2019). <https://doi.org/10.1016/j.jclepro.2019.05.079>.
- 5 Anbari, M.B., Moehler, R.C., & Gholamreza, S.F. Sustainability assessment in construction projects: A sustainable earned value management model under uncertain and unreliable conditions. *Environment Systems and Decisions*, 44(1), 45–68 (2024). <https://doi.org/10.1007/s10669-023-09913-2>.
- 6 Ottaviani, F.M., & De Marco, A. Multiple linear regression model for improved project cost forecasting. *Procedia Computer Science*, 196, 808–815 (2022). <https://doi.org/10.1016/j.procs.2021.12.079>.
- 7 İnan, T., Narbaev, T., & Hazır, Ö. A machine learning study to enhance project cost forecasting. *IFAC-PapersOnLine*, 55(10), 3286–3291 (2022). <https://doi.org/10.1016/j.ifacol.2022.10.127>.
- 8 Moshood, T.D., Rotimi, J.O.B., & Shahzad, W. Enhancing sustainability considerations in construction industry projects. *Environment, Development and Sustainability* (2024). <https://doi.org/10.1007/s10668-024-04946-2>.
- 9 Chen, L., Huang, L., Hua, J., Chen, Z., Wei, L., Osman, A. I., Fawzy, S., Rooney, D.W., Dong, L., & Yap, P.S. Green construction for low-carbon cities: a review. *Environmental Chemistry Letters*, 21(3), 1627–1657 (2023). <https://doi.org/10.1007/s10311-022-01544-4>.
- 10 Rajabi, S., El-Sayegh, S., & Romdhane, L. Identification and assessment of sustainability performance indicators for construction projects. *Environmental and Sustainability Indicators*, 15, 100193 (2022). <https://doi.org/10.1016/j.indic.2022.100193>.
- 11 Rajabi, S., El-Sayegh, S., & Romdhane, L. Project controls model for sustainable construction projects. *Sustainability*, 16(20), 9042 (2024). <https://doi.org/10.3390/su16209042>.
- 12 Koke, B., & Moehler, R. C. Earned Green Value management for project management: A systematic review. *Journal of Cleaner Production*, 230, 180–197 (2019). <https://doi.org/10.1016/j.jclepro.2019.05.079>.
- 13 Anbari, M.B., Moehler, R.C., & Gholamreza, S.F. Sustainability assessment in construction projects: A sustainable earned value management model under uncertain and unreliable conditions. *Environment Systems and Decisions*, 44, 45–68 (2024). <https://doi.org/10.1007/s10669-023-09913-2>.
- 14 Ottaviani, F.M., & De Marco, A. Multiple linear regression model for improved project cost forecasting. *Procedia Computer Science*, 196, 808–815 (2022). <https://doi.org/10.1016/j.procs.2021.12.079>.
- 15 Narbaev, T., & De Marco, A. An Earned Schedule-based regression model to improve cost estimate at completion. *International Journal of Project Management*, 32(6), 1007–1018 (2014). <https://doi.org/10.1016/j.ijproman.2013.12.005>.
- 16 Narbaev, T., Hazır, Ö., Khamitova, B., & Talgat, S. A machine learning study to improve the reliability of project cost estimates. *International Journal of Production Research*, 62(12), 4372–4388 (2024). <https://doi.org/10.1080/00207543.2023.2262051>.
- 17 Samoilov, A., Narbaev, T., Castelblanco, G., & Mukashev, Y. Evaluating public–private partnership dynamics: The Kazakhstan toll road case. *Eurasian Journal of Economic and Business Studies*, 68(1), 131–142 (2024). <https://doi.org/10.47703/ejeb.v68i1.2024.131>.
- 18 Samoilov, A., Osei-Kyei, R., Kussaiyn, M., Mamyrbayev, A., & Mukashev, Y. Cross-country comparison of risk factors in public–private partnerships in infrastructure development. *Sustainability*, 16(13), 5712 (2024). <https://doi.org/10.3390/su16135712>.
- 19 Yin, R. *Case Study Research: Design and Methods*. Sage Publications (2018).
- 20 Project Management Institute (PMI). *Practice Standard for Earned Value Management* (2nd ed.). PMI (2011).
- 21 Durdyev, S., Zavadskas, E. K., Thurnell, D., Banaitis, A., & Ihtiyar, A. Sustainable construction across developing countries: Status and future directions. *Journal of Cleaner Production*, 264, 121–134 (2020). <https://doi.org/10.1016/j.jclepro.2020.121-134>

¹*Шах Н.,

докторант, ORCID ID: 0009-0006-4525-4399,

*e-mail: n_shah@kbtu.kz

²Али Т.,

профессор, ORCID ID: 0009-0000-6953-026X,

e-mail: vc@admin.muet.edu.pk

¹Қазақстан-Британ техникалық университеті, Алматы қ., Қазақстан

²Мехран инженерия және технология университеті, Джамшоро қ., Пәкістан

ҚҰРЫЛЫС ЖОБАЛАРЫНЫҢ ДӘСТҮРЛІ МОНИТОРИНГ ЖӘНЕ БАҚЫЛАУ ЖҮЙЕЛЕРІНЕ ТҰРАҚТЫЛЫҚ МЕТРИКАЛАРЫН ЕНГІЗУ: ҚОЗҒАЛТЫЛҒАН ҚҰН ӘДІСІНЕ НЕГІЗДЕЛГЕН ТӘСІЛ

Аңдатпа

Құрылыс жобаларын бақылау жүйелерінде тұрақтылықты интеграциялау инфрақұрылымдық жобаларды әлемдік деңгейде дамыту барысында қоршаған орта мен ресурстарға өсіп жатқан қысым салдарынан шұғыл қажеттілікке айналды. Дәстүрлі Осынған Баға Басқару (Earned Value Management, EVM) шығындар мен кестені қадағалау үшін қолданылғанымен, ол жобаның нәтижелеріне әсер ететін экологиялық және ресурстық тиімділік көрсеткіштерін біріктірмейді. Бұл мақала жобаның тиімділігін бағалауда тұрақтылықтың үш көрсеткішін: көмірқышқыл газын шығару (Carbon Emissions, CE), энергия тұтыну (Energy Consumption, EC) және материалдық қалдықтар (Material Waste, MW) қамтитын кешенді EVM құрылымын ұсынады. Пайдаланылған Көптік Сызықтық Регрессия (Multiple Linear Regression, MLR) моделі Пәкістанның Карачидегі Малир Экспрессвей жобасының тоғыз айлық нақты эксплуатациялық тарихына негізделген, бұл ірі ауқымды қоғамдық инфрақұрылымдық жоба болып табылады. Нәтижелер тұрақтылық көрсеткіштері шығындар тиімділік индексі (Cost Performance Index, CPI) және кесте тиімділік индексі (Schedule Performance Index, SPI) анықтауда маңызды рөл атқаратынын көрсетеді. Ұсынылған модель шешім қабылдау процесін жетілдіреді, себебі жобаның мүдделі тараптары қаржылық және экологиялық аспектілерді бір уақытта бақылай алады. Оның әдістемесі қоғамдық инфрақұрылымдық дамуға практикалық түрде қолдануға болады және Тұрақты Осынған Баға Басқару (Sustainable Earned Value Management) жүйелерін болашақта дамыту үшін негізгі база болады.

Тірек сөздер: тұрақтылық метрикалары, қозғалтылған құн әдісі, регрессиялық талдау, жоба бақылауы, жасыл құрылыс индикаторлары, Малир Экспрессвей.

¹*Шах Н.,

докторант, ORCID ID: 0009-0006-4525-4399,

*e-mail: n_shah@kbtu.kz

²Али Т.,

профессор, ORCID ID: 0009-0000-6953-026X,

e-mail: vc@admin.muet.edu.pk

¹Казахстанско-Британский технический университет, г. Алматы, Казахстан

²Мехранский университет инженерии и технологии, г. Джамшоро, Пакистан

ВНЕДРЕНИЕ ПОКАЗАТЕЛЕЙ УСТОЙЧИВОСТИ В ТРАДИЦИОННУЮ СИСТЕМУ МОНИТОРИНГА И КОНТРОЛЯ СТРОИТЕЛЬНЫХ ПРОЕКТОВ: ПОДХОД НА ОСНОВЕ МЕТОДА ОСВОЕННОГО ОБЪЕМА

Аннотация

Интеграция устойчивого развития в системы контроля строительных проектов стала насущной необходимостью из-за растущего давления на окружающую среду и ресурсы в процессе реализации инфраструк-

турных проектов по всему миру. В то время как традиционное управление освоенной стоимостью (Earned Value Management, EVM) используется для отслеживания затрат и графика выполнения, оно не учитывает показатели экологической и ресурсной эффективности, которые становятся все более значимыми для результатов проекта. В данной статье предлагается комплексная структура EVM, включающая три показателя устойчивого развития: выбросы углекислого газа (Carbon Emissions, CE), потребление энергии (Energy Consumption, EC) и потери материалов (Material Waste, MW) при оценке эффективности проекта. Используемая модель множественной линейной регрессии (Multiple Linear Regression, MLR) основана на девяти месяцах фактической эксплуатационной истории проекта Малир Экспрессвей в Карачи, Пакистан, который является крупномасштабным общественным инфраструктурным проектом. Результаты показывают, что переменные устойчивого развития играют значительную роль в определении индекса эффективности затрат (Cost Performance Index, CPI) и индекса эффективности графика (Schedule Performance Index, SPI). Предложенная модель улучшает процесс принятия решений, поскольку заинтересованные стороны проекта могут одновременно контролировать финансовые и экологические аспекты. Ее методология практически применима к развитию общественной инфраструктуры и станет отправной точкой для будущей разработки рамок устойчивого управления освоенной стоимостью (Sustainable Earned Value Management).

Ключевые слова: показатели устойчивости, метод освоенного объема, регрессионный анализ, контроль проекта, индикаторы зеленого строительства, Малир Экспрессвей.

Article submission date: 10.12.2025

REQUIREMENTS TO THE DESIGN OF ARTICLES PUBLISHED IN THE JOURNAL KBTU BULLETIN

Previously **unpublished material by the author (s)** corresponding to the profile of the journal in Russian, Kazakh or English is accepted for publication. The article is provided in electronic format (in formats .doc, .docx,) only by downloading it through the functionality of the journal's website (<https://vestnik.kbtu.edu.kz/jour/about>), which has detailed instructions for authors on the process of submitting articles, as well as on payment for publication.

Mandatory requirements for registration:

The MRNTI index is placed in the upper-left corner. (Times New Roman 14 size)

One single interval below the authors' surnames and initials in bold. Detailed names of organizations, cities, countries where the work was performed, e-mail of one of the authors. Specify footnotes to the place of work of the authors.

Below, in one interval, the full title of the article is in bold.

Below, a brief (150–250 words) abstract (abstract) at one interval. Abstract should outline the general orientation of the article and formulate its main results and be self-sufficient, it should not contain references to literary sources or to any sections and elements of the article to which it relates.

Key words related to the submitted article.

The text of the article is typed in a straight font Times New Roman, size 14. The line spacing is single. All fields are 2 cm long.

The article should be divided into sections: Introduction, Materials and Methods, Results and Discussion, Conclusion, Information on funding (if available).

At the end, a "List of References" is given. Literary references in the text are in square brackets [1]. The list of references is drawn up twice: according to the above GOST. The list should fully correspond to the main text of the article, contain all the works mentioned (and only mentioned) in it. Next, the same list of references should be presented in the generally accepted English transliteration (see the sample of the literature design below).

The article undergoes "blind" review, reviewers are appointed by the editorial board of the journal.

At the end of the article please indicate full name, return address, phone numbers, fax, e-mail address.

The author(s) are responsible for the content of the article. The opinion of the Scientific and Editorial Board does not always coincide with the opinion of the author (authors). The Editorial Board reserves the right to publish or reject articles.

The reference to our edition in reprinting is obligatory.

Our address: **59, Tole Bi St., Almaty, 050000, Kazakhstan-British Technical University.**

Website: <https://vestnik.kbtu.edu.kz/jour/pages/view/EditorialC> Editorial Board.

**ЖУРНАЛЫНДА ЖАРИЯЛАНАТЫН МАҚАЛАЛАРДЫ
РӘСІМДЕУГЕ ҚОЙЫЛАТЫН ТАЛАПТАР**

Жариялау үшін **автормен (авторлармен) бұрын жарияланбаған**, журнал бейініне сәйкес келетін орыс, қазақ немесе ағылшын тілдеріндегі материалдар қабылданады. Мақала журнал сайтының (<https://vestnik.kbtu.edu.kz/jour/about>) функционалдығы арқылы жүктеу жолымен тек электронды форматта (.doc, .docx форматтарында) қабылданады, сайтта авторларға мақаланы жүктеу, ұсыну үрдісі туралы, сондай-ақ жарияланымға ақы төлеу туралы толық нұсқаулық бар.

Ресімдеу бойынша міндетті талаптар:

Жоғарғы сол жақ бұрышта FTAMI индексі жазылады. (Times New Roman 14 өлшем)

Бір бірлік интервалдан кейін авторлардың тегі мен аты-жөні жартылай қалың қаріппен жазылады. Жұмыс орындалған ұйымдардың, қалалардың, елдердің толық атаулары, авторлардың бірінің e-mail поштасы көрсетіледі. Авторлардың жұмыс орнына сілтеме беріледі.

Бір интервалдан кейін мақаланың толық атауы жартылай қалың қаріппен жазылады.

Бір интервалдан кейін қысқаша (150–250 сөз) аңдатпа (abstract) ұсынылады. Аңдатпа мақаланың жалпы бағытын көрсетіп, оның негізгі нәтижелерін тұжырымдауы тиіс, аңдатпада әдеби дереккөздерге де, оған қатысты мақаланың кез-келген бөлімдері мен элементтеріне сілтеме жасауға болмайды.

Ұсынылған мақалаға қатысты кілт сөздер (тірек сөздер) жазылады.

Мақала мәтіні 14 өлшемді Times New Roman шрифтімен теріледі. Жолдар арасында бірлік интервал сақталады. Шеттер – барлық жағынан 2 см.

Мақала келесі бөлімдерді қамтуы тиіс: Кіріспе, Материалдар және әдістер, Нәтижелер мен талқылау, Қорытынды, Қаржыландыру туралы ақпарат (болған жағдайда көрсетіледі).

Соңында «Әдебиеттер тізімі (References)» жазылады. Мәтіндегі әдеби сілтемелер төртбұрышты жақшада ұсынылады [1]. Әдебиеттер тізімі MEMCT-ке сәйкес екі рет рәсімделеді. Тізім мақаланың негізгі мәтініне толық сәйкес келіп, мақалада аталған (тек аталған) барлық жұмыстарды қамтуы керек. Әрі қарай, дәл осы сілтемелер жалпы қабылданған ағылшын транслитерациясында жазылады.

Мақала «жасырын» рецензиялаудан өтеді, рецензенттерді журналдың редакциялық алқасы тағайындайды.

Мақаланың соңында авторлардың толық аты-жөнін, кері байланыс мекенжайын, телефондар, факс және электрондық пошта мекенжайын (e-mail) жазуыңызды сұраймыз.

Мақаланың мазмұнына автор (авторлар) жауапты. Ғылыми-редакция алқасының пікірі әрдайым автордың (авторлардың) пікірімен сәйкес келе бермейді. Редакция алқасы мақалаларды жариялау немесе қабылдамау құқығын өзіне қалдырады.

Қайта басылым кезінде біздің жарияланымға сілтеме жасау керек.

Мекенжайымыз: **050000, Алматы қ., Төле би көш., 59, Қазақстан-Британ техникалық университеті.**

Сайт: <https://vestnik.kbtu.edu.kz/jour/pages/view/EditorialC> Редакция алқасы.

ТРЕБОВАНИЯ К ОФОРМЛЕНИЮ СТАТЕЙ, ПУБЛИКУЕМЫХ В ЖУРНАЛЕ ВЕСТНИК КБТУ

Для публикации принимается **ранее неопубликованный автором (авторами) материал**, соответствующий профилю журнала, на русском, казахском или английском языках. Статья предоставляется в электронном формате (в форматах .doc, .docx,) только путем ее загрузки через функционал сайта журнала (<https://vestnik.kbtu.edu.kz/jour/about>), в котором имеется подробная инструкция для авторов о процессе подачи статей, а также об оплате публикации.

Обязательные требования по оформлению:

В левом верхнем углу ставится индекс МРНТИ (Times New Roman 14 размер).

Через один одинарный интервал ниже фамилии и инициалы авторов полужирным шрифтом. Развернутые названия организаций, городов, стран, где выполнена работа, e-mail одного из авторов. Указать сноски на место работы авторов.

Ниже через один интервал полное название статьи полужирным шрифтом.

Ниже через один интервал краткая (150–250 слов) аннотация (abstract). Аннотация должна обрисовывать общую направленность статьи, формулировать ее основные результаты и быть самодостаточной, в ней недопустимы ссылки ни на литературные источники, ни на какие-либо разделы и элементы статьи, к которой она относится.

Ключевые слова (Key words), относящиеся к представляемой статье.

Текст статьи набирается прямым шрифтом Times New Roman, размер 14. Интервал между строками одинарный. Поля все по 2 см.

Статья должна быть разделена на разделы: Введение, Материалы и методы, Результаты и обсуждение, Заключение, Информация о финансировании (при наличии).

В конце дается «Список литературы (References)». Литературные ссылки в тексте в квадратных скобках [1]. Список литературы оформляется дважды: согласно ГОСТу. Список должен полностью соответствовать основному тексту статьи, содержать все упомянутые (и только упомянутые) в нем работы. Далее тот же список литературы представить в общепринятой английской транслитерации.

Статья проходит «слепое» рецензирование, рецензенты назначаются редколлегией журнала.

В конце статьи просим указать Ф.И.О. полностью, обратный адрес, телефоны, факс, адрес электронной почты (e-mail).

Ответственность за содержание статьи несет автор (авторы). Мнение Научно-редакционного совета не всегда совпадает с мнением автора (авторов). Редакционный совет оставляет за собой право публикации или отклонения статей.

Ссылка на наше издание при перепечатке обязательна.

Наш адрес: **050000, г. Алматы, ул. Толе Би, 59, Казахстанско-Британский технический университет.**

Сайт: <https://vestnik.kbtu.edu.kz/jour/pages/view/EditorialC> Редакционная коллегия.

ҚАЗАҚСТАН-БРИТАН ТЕХНИКАЛЫҚ УНИВЕРСИТЕТІНІҢ
ХАБАРШЫСЫ
HERALD
OF THE KAZAKH-BRITISH TECHNICAL UNIVERSITY
ВЕСТНИК
КАЗАХСТАНСКО-БРИТАНСКОГО ТЕХНИЧЕСКОГО УНИВЕРСИТЕТА

Ответственный за выпуск
Алагузов Рашид Жуманович

Редакторы
Кабылдаева Толганай Кайсаровна,
Скуратова Ирина Михайловна

Компьютерная верстка
Жуйкова Майя Антоновна

Подписано в печать 12.12.2025 г.
Тираж 300 экз. Формат 60х84 1/16.
Бумага тип. Усл.печ.л. 51,0 Уч.-изд.л. 53,0
Заказ № 512

Редакция журнала «Вестник КБТУ» не несет ответственность за содержание публикуемых статей. Содержания статей целиком принадлежат авторам, и размещаются в журнале исключительно под их ответственность.

Издание Казахстанско-Британского технического университета
Издательский центр КБТУ, Алматы, ул. Толе би, 59