

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

ҚАЗАҚСТАН РЕСПУБЛИКАСЫ
ҰЛТТЫҚ ҒЫЛЫМ АКАДЕМИЯСЫ

әл-Фараби атындағы Қазақ ұлттық университетінің

Х А Б А Р Л А Р Ы

ИЗВЕСТИЯ

НАЦИОНАЛЬНОЙ АКАДЕМИИ
НАУК РЕСПУБЛИКИ КАЗАХСТАН
Казахский национальный
университет имени аль-Фараби

N E W S

OF THE ACADEMY OF SCIENCES
OF THE REPUBLIC OF
KAZAKHSTAN
al-Farabi Kazakh National University

SERIES
PHYSICS AND INFORMATION TECHNOLOGY

2 (346)

APRIL – JUNE 2023

PUBLISHED SINCE JANUARY 1963

PUBLISHED 4 TIMES A YEAR

ALMATY, NAS RK

БАС РЕДАКТОР:

МУТАНОВ Ғалымқайыр Мұтанұлы, техника ғылымдарының докторы, профессор, ҚР ҰҒА академигі, ҚР БҒМ ҒК «Ақпараттық және есептеу технологиялары институты» бас директорының м.а. (Алматы, Қазақстан), **Н=5**

БАС РЕДАКТОРДЫҢ ОРЫНБАСАРЫ:

МАМЫРБАЕВ Өркен Жұмажанұлы, ақпараттық жүйелер мамандығы бойынша философия докторы (Ph.D), ҚР БҒМ Ғылым комитеті «Ақпараттық және есептеуші технологиялар институты» РМК жауапты хатшысы (Алматы, Қазақстан), **Н=5**

РЕДАКЦИЯ АЛҚАСЫ:

ҚАЛИМОЛДАЕВ Мақсат Нұрәділұлы, физика-математика ғылымдарының докторы, профессор, ҚР ҰҒА академигі (Алматы, Қазақстан), **Н=7**

БАЙГУНЧЕКОВ Жұмаділ Жанабайұлы, техника ғылымдарының докторы, профессор, ҚР ҰҒА академигі, Кибернетика және ақпараттық технологиялар институты, Сатпаев университетінің Қолданбалы механика және инженерлік графика кафедрасы, (Алматы, Қазақстан), **Н=3**

ВОЙЧИК Вальдемар, техника ғылымдарының докторы (физика), Люблин технологиялық университетінің профессоры (Люблин, Польша), **Н=23**

БОШКАЕВ Қуантай Авгазыұлы, Ph.D. Теориялық және ядролық физика кафедрасының доценті, әл-Фараби атындағы Қазақ ұлттық университеті (Алматы, Қазақстан), **Н=10**

QUEVEDO Nemandó, профессор, Ядролық ғылымдар институты (Мехико, Мексика), **Н=28**

ЖҮСПОВ Марат Абжанұлы, физика-математика ғылымдарының докторы, теориялық және ядролық физика кафедрасының профессоры, әл-Фараби атындағы Қазақ ұлттық университеті (Алматы, Қазақстан), **Н=7**

КОВАЛЕВ Александр Михайлович, физика-математика ғылымдарының докторы, Украина ҰҒА академигі, Қолданбалы математика және механика институты (Донецк, Украина), **Н=5**

РАМАЗАНОВ Тілекқабұл Сәбитұлы, физика-математика ғылымдарының докторы, профессор, ҚР ҰҒА академигі, әл-Фараби атындағы Қазақ ұлттық университетінің ғылыми-инновациялық қызмет жөніндегі проректоры, (Алматы, Қазақстан), **Н=26**

ТАКИБАЕВ Нұрғали Жәбағаұлы, физика-математика ғылымдарының докторы, профессор, ҚР ҰҒА академигі, әл-Фараби атындағы Қазақ ұлттық университеті (Алматы, Қазақстан), **Н=5**

ТИГИНЯНУ Ион Михайлович, физика-математика ғылымдарының докторы, академик, Молдова Ғылым Академиясының президенті, Молдова техникалық университеті (Кишинев, Молдова), **Н=42**

ХАРИН Станислав Николаевич, физика-математика ғылымдарының докторы, профессор, ҚР ҰҒА академигі, Қазақстан-Британ техникалық университеті (Алматы, Қазақстан), **Н=10**

ДАВЛЕТОВ Асқар Ербуланович, физика-математика ғылымдарының докторы, профессор, әл-Фараби атындағы Қазақ ұлттық университеті (Алматы, Қазақстан), **Н=12**

КАЛАНДРА Пьетро, Ph.D (физика), Наноқұрылымды материалдарды зерттеу институтының профессоры (Рим, Италия), **Н=26**

«ҚР ҰҒА Хабарлары. Физика және информатика сериясы».

ISSN 2518-1726 (Online),

ISSN 1991-346X (Print)

Меншіктеуші: «Қазақстан Республикасының Ұлттық ғылым академиясы» РҚБ (Алматы қ.). Қазақстан Республикасының Ақпарат және қоғамдық даму министрлігінің Ақпарат комитетінде 14.02.2018 ж. берілген **№ 16906-Ж** мерзімдік басылым тіркеуіне қойылу туралы куәлік.

Тақырыптық бағыты: *физика және ақпараттық коммуникациялық технологиялар сериясы.*

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

Мерзімділігі: *жылына 4 рет.*

Тиражы: *300 дана.*

Редакцияның мекен-жайы: *050010, Алматы қ., Шевченко көш., 28, 219 бөл., тел.: 272-13-19*
<http://www.physico-mathematical.kz/index.php/en/>

© Қазақстан Республикасының Ұлттық ғылым академиясы, 2023

Типографияның мекен-жайы: «Аруна» ЖК, Алматы қ., Мұратбаев көш., 75.

ГЛАВНЫЙ РЕДАКТОР:

МУТАНОВ Галимжаир Мутанович, доктор технических наук, профессор, академик НАН РК, и.о. генерального директора «Института информационных и вычислительных технологий» КН МОН РК (Алматы, Казахстан), **Н=5**

ЗАМЕСТИТЕЛЬ ГЛАВНОГО РЕДАКТОРА:

МАМЫРБАЕВ Оркен Жумажанович, доктор философии (PhD) по специальности Информационные системы, ответственный секретарь РГП «Института информационных и вычислительных технологий» Комитета науки МОН РК (Алматы, Казахстан), **Н=5**

РЕДАКЦИОННАЯ КОЛЛЕГИЯ:

КАЛИМОЛДАЕВ Максат Нурадилович, доктор физико-математических наук, профессор, академик НАН РК (Алматы, Казахстан), **Н=7**

БАЙГУНЧЕКОВ Жумадил Жанабаевич, доктор технических наук, профессор, академик НАН РК, Институт кибернетики и информационных технологий, кафедра прикладной механики и инженерной графики, Университет Сатпаева (Алматы, Казахстан), **Н=3**

ВОЙЧИК Вальдемар, доктор технических наук (физ.-мат.), профессор Люблинского технологического университета (Люблин, Польша), **Н=23**

БОШКАЕВ Куантай Авгазыевич, доктор Ph.D, преподаватель, доцент кафедры теоретической и ядерной физики, Казахский национальный университет им. аль-Фараби (Алматы, Казахстан), **Н=10**

QUEVEDO Hemando, профессор, Национальный автономный университет Мексики (UNAM), Институт ядерных наук (Мехико, Мексика), **Н=28**

ЖУСУПОВ Марат Абжанович, доктор физико-математических наук, профессор кафедры теоретической и ядерной физики, Казахский национальный университет им. аль-Фараби (Алматы, Казахстан), **Н=7**

КОВАЛЕВ Александр Михайлович, доктор физико-математических наук, академик НАН Украины, Институт прикладной математики и механики (Донецк, Украина), **Н=5**

РАМАЗАНОВ Тлеккабул Сабитович, доктор физико-математических наук, профессор, академик НАН РК, проректор по научно-инновационной деятельности, Казахский национальный университет им. аль-Фараби (Алматы, Казахстан), **Н=26**

ТАКИБАЕВ Нурғали Жабагаевич, доктор физико-математических наук, профессор, академик НАН РК, Казахский национальный университет им. аль-Фараби (Алматы, Казахстан), **Н=5**

ТИГИНЯНУ Ион Михайлович, доктор физико-математических наук, академик, президент Академии наук Молдовы, Технический университет Молдовы (Кишинев, Молдова), **Н=42**

ХАРИН Станислав Николаевич, доктор физико-математических наук, профессор, академик НАН РК, Казахстанско-Британский технический университет (Алматы, Казахстан), **Н=10**

ДАВЛЕТОВ Аскар Ербуланович, доктор физико-математических наук, профессор, Казахский национальный университет им. аль-Фараби (Алматы, Казахстан), **Н=12**

КАЛАНДРА Пьетро, доктор философии (Ph.D, физика), профессор Института по изучению наноструктурированных материалов (Рим, Италия), **Н=26**

«Известия НАН РК. Серия физика и информатики».

ISSN 2518-1726 (Online),

ISSN 1991-346X (Print)

Собственник: *Республиканское общественное объединение «Национальная академия наук Республики Казахстан» (г. Алматы).*

Свидетельство о постановке на учет периодического печатного издания в Комитете информации Министерства информации и общественного развития Республики Казахстан **№ 16906-Ж** выданное 14.02.2018 г.

Тематическая направленность: *серия физика и информационные коммуникационные технологии.* В настоящее время: *вошел в список журналов, рекомендованных ККСОН МОН РК по направлению «информационные коммуникационные технологии».*

Периодичность: *4 раз в год.*

Тираж: *300 экземпляров.*

Адрес редакции: *050010, г. Алматы, ул. Шевченко, 28, оф. 219, тел.: 272-13-19*

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

© Национальная академия наук Республики Казахстан, 2023
Адрес типографии: ИП «Аруна», г. Алматы, ул. Муратбаева, 75.

EDITOR IN CHIEF:

MUTANOV Galimkair Mutanovich, doctor of technical Sciences, Professor, Academician of NAS RK, acting director of the Institute of Information and Computing Technologies of SC MES RK (Almaty, Kazakhstan), **H=5**

DEPUTY EDITOR-IN-CHIEF

MAMYRBAYEV Orken Zhumazhanovich, Ph.D. in the specialty information systems, executive secretary of the RSE “Institute of Information and Computational Technologies”, Committee of Science MES RK (Almaty, Kazakhstan) **H=5**

EDITORIAL BOARD:

KALIMOLDAYEV Maksat Nuradilovich, doctor in Physics and Mathematics, Professor, Academician of NAS RK (Almaty, Kazakhstan), **H=7**

BAYGUNCHEKOV Zhumadil Zhanabayevich, doctor of Technical Sciences, Professor, Academician of NAS RK, Institute of Cybernetics and Information Technologies, Department of Applied Mechanics and Engineering Graphics, Satbayev University (Almaty, Kazakhstan), **H=3**

WOICIK Waldemar, Doctor of Phys.-Math. Sciences, Professor, Lublin University of Technology (Lublin, Poland), **H=23**

BOSHKAYEV Kuantai Avgazievich, PhD, Lecturer, Associate Professor of the Department of Theoretical and Nuclear Physics, Al-Farabi Kazakh National University (Almaty, Kazakhstan), **H=10**

QUEVEDO Hemando, Professor, National Autonomous University of Mexico (UNAM), Institute of Nuclear Sciences (Mexico City, Mexico), **H=28**

ZHUSSUPOV Marat Abzhanovich, Doctor in Physics and Mathematics, Professor of the Department of Theoretical and Nuclear Physics, al-Farabi Kazakh National University (Almaty, Kazakhstan), **H=7**

KOVALEV Alexander Mikhailovich, Doctor in Physics and Mathematics, Academician of NAS of Ukraine, Director of the State Institution «Institute of Applied Mathematics and Mechanics» DPR (Donetsk, Ukraine), **H=5**

RAMAZANOV Tlekkabul Sabitovich, Doctor in Physics and Mathematics, Professor, Academician of NAS RK, Vice-Rector for Scientific and Innovative Activity, al-Farabi Kazakh National University (Almaty, Kazakhstan), **H=26**

TAKIBAYEV Nurgali Zhabagaevich, Doctor in Physics and Mathematics, Professor, Academician of NAS RK, al-Farabi Kazakh National University (Almaty, Kazakhstan), **H=5**

TIGHINEANU Ion Mikhailovich, Doctor in Physics and Mathematics, Academician, Full Member of the Academy of Sciences of Moldova, President of the AS of Moldova, Technical University of Moldova (Chisinau, Moldova), **H=42**

KHARIN Stanislav Nikolayevich, Doctor in Physics and Mathematics, Professor, Academician of NAS RK, Kazakh-British Technical University (Almaty, Kazakhstan), **H=10**

DAVLETOV Askar Erbulanovich, Doctor in Physics and Mathematics, Professor, al-Farabi Kazakh National University (Almaty, Kazakhstan), **H=12**

CALANDRA Pietro, PhD in Physics, Professor at the Institute of Nanostructured Materials (Monterotondo Station Rome, Italy), **H=26**

News of the National Academy of Sciences of the Republic of Kazakhstan.

Series of physics and informatics.

ISSN 2518-1726 (Online),

ISSN 1991-346X (Print)

Owner: RPA «National Academy of Sciences of the Republic of Kazakhstan» (Almaty). The certificate of registration of a periodical printed publication in the Committee of information of the Ministry of Information and Social Development of the Republic of Kazakhstan **No. 16906-Ж**, issued 14.02.2018
Thematic scope: *series physics and information technology*.

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

Periodicity: *4 times a year.*

Circulation: *300 copies.*

Editorial address: *28, Shevchenko str., of. 219, Almaty, 050010, tel. 272-13-19*

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

© National Academy of Sciences of the Republic of Kazakhstan, 2023

Address of printing house: ST «Aruna», 75, Muratbayev str, Almaty.

NEWS OF THE NATIONAL ACADEMY OF SCIENCES OF THE REPUBLIC OF
KAZAKHSTAN
PHYSICO-MATHEMATICAL SERIES
ISSN 1991-346X
Volume 2. Number 346 (2023). 5–20
<https://doi.org/10.32014/2023.2518-1726.180>

UDK 004-056-5

© A. Adamova^{1*}, T. Zhukabayeva², Y. Mardenov³, 2023

¹ Astana IT University, Astana, Kazakhstan;

² L. Gumilyov Eurasian National University, Astana IT University,
Astana, Kazakhstan;

³ Astana International University, Astana, Kazakhstan.

E-mail: aigul.adamova@astanait.edu.kz

INTERNET OF THINGS: STATUS AND PROSPECTS FOR THE DEVELOPMENT OF LIGHTWEIGHT ALGORITHMS

Adamova Aigul — Doctor PhD of Computing and Software Engineering, Assistant Professor, Department of Computer Engineering, Astana IT University, Astana, Kazakhstan

E-mail: aigul.adamova@astanait.edu.kz. ORCID ID: 0000-0001-7773-9522;

Zhukabayeva Tamara — PhD (Informatics, Computer Engineering and Control), Associate Professor, Department of Information Systems, L. Gumilyov Eurasian National University, Astana IT University, Astana, Kazakhstan

E-mail: tamara.kokenovna@gmail.com. ORCID ID: 0000-0001-6345-5211;

Mardenov Yerik — PhD. Doctoral student. Astana International University, Astana, Kazakhstan

E-mail: emardenov@gmail.com. ORCID ID: 0000-0001-9284-9797.

Abstract. There is a high probability of various attacks coming from unknown network devices, in this regard, ensuring the security and confidentiality of data is relevant and one of the main problems today. It is important to note that IoT has a number of limitations in terms of power supply, memory and dimensions. Thus, it is necessary to define more resource-optimized and security-related inferences to solve the problems generated in the network. Along with this, device resources are consumed at a higher rate due to complex cryptographic maintenance algorithms, so it is necessary to determine an appropriate encryption procedure for an automated IoT network, taking into account data integrity. This article presents an analysis of lightweight algorithms for hardware and software implementation, the levels of various data flow architectures in the IoT network are considered, an overview of scientific papers is given that notes the relevance of ensuring security in the interaction of IoT devices and the development of lightweight cryptography for 2022–2023. As a result, the results of a comparative analysis of lightweight algorithms in terms of several indicators for hardware implementation are given; when implemented in software, relative to the amount of memory and time delay.

Keywords: Internet of Things, security, privacy, lightweight encryption, lightweight cryptography

Financing: This research has been/was/is funded by the Science Committee of the Ministry of Education and Science of the Republic of Kazakhstan (Grant No. AP14973006).

Conflict of interest: The authors declare that there is no conflict of interest.

© **А. Адамова^{1*}, Т. Жукабаева², Е. Марденов³, 2023**

¹ Astana IT University, Астана, Қазақстан;

² Евразийский национальный университет им. Л. Гумилева,
Астана, Қазақстан;

³ Международный университет Астана, Астана, Қазақстан.

E-mail: aigul.adamova@astanait.edu.kz

ЗАТТАР ИНТЕРНЕТІ: ЖЕҢІЛДІК АЛГОРИТМДЕРДІҢ ДАМУЫ ЖӘНЕ БОЛАШАҒЫ

Адамова Айгуль — Есептеу техникасы және бағдарламалық қамтамасыз ету мамандығы бойынша философия докторы (PhD), Компьютерлік инженерия департаментінің профессор ассистенті. Astana IT University. Астана, Қазақстан

E-mail: aigul.adamova@astanait.edu.kz. ORCID ID: 0000-0001-7773-9522;

Zhukabayeva Tamara — Информатика, есептеу техникасы және бақылау мамандығы бойынша философия докторы (PhD), Ақпараттық жүйелер кафедрасының профессор м.а., қауымдастырылған профессор. Л.Н.Гумилев атындағы Еуразия ұлттық университеті. Астана, Қазақстан

E-mail: tamara.kokenovna@gmail.com. ORCID ID: 0000-0001-6345-5211;

Mardenov Yerik — PhD Докторанты. Астана халықаралық уни верситеті. Астана, Қазақстан

E-mail: emardenov@gmail.com. ORCID ID: 0000-0001-9284-9797.

Аннотация. Белгісіз желілік құрылғылардан келетін әртүрлі шабуылдардың үлкен ықтималдығы бар, осыған байланысты деректердің қауіпсіздігі мен құпиялылығын қамтамасыз ету өзекті және бүгінгі күні негізгі мәселелердің бірі болып табылады. IoT-те қуат, жады және бірқатар өлшемдер бойынша шектеулер бар екенін ескеру маңызды. Осылайша, желіде туындаған мәселелерді шешу үшін ресурстардың онтайландырылған және қауіпсіздікке қатысты қорытындыларды анықтау қажет. Сонымен қатар, күрделі криптографиялық қызмет алгоритмдеріне байланысты құрылғы ресурстары жоғары жылдамдықпен қолданады, сондықтан деректердің тұтастығын ескере отырып, автоматтандырылған IoT желісі үшін қолайлы шифрлау процедурасын анықтау қажет. Бұл мақалада аппараттық және бағдарламалық қамтамасыз етуді іске асырудағы жеңіл салмақ алгоритмдерін талдау, IoT желісіндегі әртүрлі деректер ағынының архитектураларының деңгейлері қарастырылған, IoT құрылғыларының өзара әрекеттесуіндегі қауіпсіздікті қамтамасыз етудің өзектілігін және 2022–2023 жылдардағы жеңіл салмақ криптографиясының дамуын атап өтетін ғылыми жұмыстарға

шолу берілген. Нәтижелер ретінде аппараттық іске асыру кезінде бірнеше көрсеткіштер бойынша жеңіл салмақтағы алгоритмдерді салыстырмалы талдау нәтижелері келтірілген; онымен қоса жады көлеміне және уақыт кідірісіне қатысты бағдарламалық іске асыру қорытындылары көрсетілген.

Түйін сөздер: Интернет заттар, қауіпсіздік, құпиялылық, жеңіл шифрлау, жеңіл криптография

Қаржыландыру: Бұл зерттеу Қазақстан Республикасы Ғылым және жоғары білім министрлігі, Ғылым комитетімен қаржыландырған (Грант № BR10262555).

Мүдделер қақтығысы: Авторлар осы мақалада мүдделер қақтығысы жоқ деп мәлімдемейді.

© А. Адамова^{1*}, Т. Жукабаева², Е. Марденов³, 2023

¹ Astana IT University, Астана, Қазақстан;

² Евразийский национальный университет им. Л. Гумилева, Астана, Қазақстан;

³ Международный университет Астана, Астана, Қазақстан.

E-mail: aigul.adamova@astanait.edu.kz

ИНТЕРНЕТ ВЕЩЕЙ: СОСТОЯНИЕ И ПЕРСПЕКТИВЫ РАЗВИТИЯ ЛЕГКОВЕСНЫХ АЛГОРИТМОВ

Адамова Айгуль — доктор PhD. Ассистент профессора. Департамент компьютерной инженерии. Astana IT University. Астана, Қазақстан

E-mail: aigul.adamova@astanait.edu.kz. ORCID ID: 0000-0001-7773-9522;

Zhukabayeva Tamara — доктор PhD. и.о. профессора, ассоциированный профессор. Кафедра информационных систем. Евразийский национальный университет им.Л.Н.Гумилева. Астана, Қазақстан

E-mail: tamara.kokenovna@gmail.com. ORCID ID: 0000-0001-6345-5211;

Mardenov Yerik — Докторант PhD. Международный университет Астана. Астана, Қазақстан

E-mail: emardenov@gmail.com. ORCID ID: 0000-0001-9284-9797.

Аннотация. Существует большая вероятность различных атак, исходящих от неизвестных устройств сети, в связи с этим обеспечение безопасности и конфиденциальности данных являются актуальным и одним из основных проблем на сегодняшний день. Важно отметить, что IoT имеет ряд ограничений по запасу электропитания, памяти и габаритов. Таким образом, необходимо определить более оптимизированные для ресурсов и связанные с безопасностью выводы для решения проблем, генерируемых в сети. Вместе с этим, ресурсы устройств потребляются с более высокой скоростью из-за сложных криптографических алгоритмов обслуживания, поэтому необходимо определить подходящую процедуру шифрования для автоматизированной сети IoT с учетом целостности данных. В данной статье представлен анализ легковесных алгоритмов при аппаратной и программной

реализации, рассмотрены уровни различных архитектур потока данных в сети IoT, приведены обзор научных работ отмечающих актуальность обеспечения безопасности при взаимодействии IoT устройств и развитие легковесной криптографии за 2022–2023 года. В качестве результатов приведены результаты сравнительного анализа легковесных алгоритмов по нескольким показателям при аппаратной реализации; при программной реализации относительно объема памяти и задержки по времени.

Ключевые слова: Интернет вещей, безопасность, конфиденциальность, легковесное шифрование, легковесная криптография

Финансирование: Данное исследование финансировалось Комитетом науки Министерства науки и высшего образования Республики Казахстан (Грант № BR10262555).

Конфликт интересов: авторы заявляют об отсутствии конфликта интересов.

Introduction

The Internet of Things (IoT) is an intelligent infrastructure formed using several self-organizing devices. Currently, the “IoT” can be defined as an intelligent infrastructure of interacting physical objects in the form of sensors and actuators with the digital world (Vermesan et al., 2009: 52), which transmits information using a network. There are many applications where systems are deployed using IoT. In everyday life, people use many devices, which include sensors to detect problems, transmit information, monitor, control, and so on. The information is provided in real time and can be used to make a decision; therefore, the security and confidentiality of information is a very important aspect. It is also important that the IoT has a number of limitations in terms of power supply, memory and dimensions. Along with this, it is important to determine suitable encryption methods for the IoT network in order to ensure data integrity (Eryk et al., 2022: 18).

According to Statista, the number of attacks on the Internet of Things in 2022 exceeded 10.54 million. Ensuring security in IoT interaction is one of the important tasks of the present time. One of the current methods for ensuring the security and confidentiality of transmitted data in a network of IoT devices is lightweight cryptography algorithms. The National Institute of Standards and Technology NIST conducts research in this area and scientists around the world in search of optimally suitable lightweight algorithms used in IoT devices. So scientists from Graz Technology University (Austria), chip manufacturing company Infineon Technologies (Germany), Lamarr Security Research (Austria) and Radboud University (Netherlands) are working on security and privacy issues,

Encryption is an effective solution to ensure the confidentiality of information and its integrity. Today, IoT applies encryption to touch devices in environments with various restrictions, such as limited memory, low computing power, small physical area, devices with limited power consumption, and at the same time, processes must

take place in real time, that have not previously been encrypted. When designing IoT there are a number of risks directly related to energy consumption and data security. If standard cryptography methods are applied to IoT devices, they may not support the given performance and not only, respectively, for these problems, the solution can be - lightweight cryptography (Nurlan et al., 2021: 19).

This article consists of two sections, which discusses the IoT architecture, development stages, parameters, lightweight cryptography standards, an overview of the work, and also provides an analysis of the numerical characteristics of the hardware and software implementation of lightweight algorithms.

IoT architecture

There are several types of IoT architecture, such as three-, four-, five-layer architectures (Figure 1). Table 1 shows scientists - whose research is carried out on various architectures. To determine the relevance and degree of resolution of the issue under consideration, a search and review of scientific papers in the databases of digital libraries "Web Of Science", "Google Scholar", "IEEE Xplore" was carried out. The review paper (Muhammad et al., 2022: 12) discusses the importance of IoT network security, considers various threats depending on the IoT architectural model. The paper also presents lightweight cryptographic algorithms and protocols for data protection in the IoT environment. S.L. Keoh et al presented an overview of the IEFT (Internet Engineering Task Force) requirements for standardizing security solutions in the IoT network (Keoh, et al., 2014: 10). P. Gaikweid et al presented the IoT architecture associated with the attack model. As mentioned above, IoT is used in various applications, for example, a smart home network can control home devices and appliances, controlling and remotely controlling through various connection methods. The paper also presents some problems in the mode and real time (Gaikwad et al., 2015: 6). It is important to note that depending on the IoT architecture and the level under consideration, various security solutions can be selected for the interaction of IoT devices.

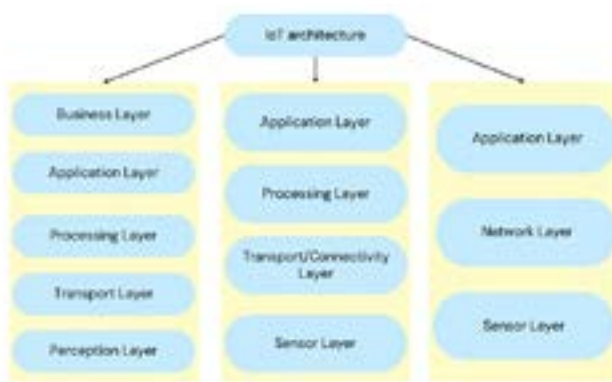


Fig. 1. Layers of different IoT data flow architectures

Consider a three-layer architecture:

- The level of sensors is the physical level, which consists of sensors and sensors responsible for collecting information about the surrounding world. At the sensor level, the IEEE 802.15.4 standard is used as a specification. It's an affordable solution that provides security, but still needs to address existing threat protection loopholes. For example, RFID, various sensors for location, motion, voice, etc. Possible attacks at the sensor level — jamming, tampering, radio interference, unfairness, exhaustion, collisions (Singh et al., 2020: 12);

- Network layer - serves to communicate with other network objects, such as servers, network devices, etc. Here we can note the method of interaction with each other — a wireless sensor network or Internet protocols. The network layer works with the physical layer data. It is also used to disseminate and analyze sensor data. At the network level, the message is divided into packets in order to route packages from source to destination using IPv6. With IoT networks growing rapidly, IPv4 address space has an advantage over IPv6 with more address spaces. Built-in cryptographic conventions such as AES and DES can be updated with IPsec at this level. Attacks possible at the network level — Sinkhole, blackhole, wormhole, misdirection, homing (Singh et al., 2020: 12);

- Application level — the level of applications that provide a special service for working with data and managing them. The application layer is responsible for providing the user with resources regarding the application being used. This layer supports services for client and programmatic functions. As an example, various classes of IoT solutions can be noted: Smart city, smarthome, digital factories, precision farming oil, etc. Attacks possible at the application level — Reprogram, Overwhelm (Singh et al., 2020: 12).

Table - 1. IoT architecture

Type of Architecture	Year	Authors	Name of the Journal
Three-layer architecture	2015	I. Mashal et al.	Ad Hoc Networks Journal
	2022	Aguru, A.D. et al.	Algorithms
	2022	B. Paul	Lecture Notes in Networks and Systems
	2020	M. Parto	Procedia Manufacturing
	2016	F. Bing	2016 2nd International Conference on Cloud Computing and IoT
	2022	B. Paul	ICT Analysis and Applications
Four-layer architecture	2017	C. Zhong et al.	16th International Symposium on Distributed Computing and Applications to Business, Engineering and Science
	2013	J. Gubbi et al.	Future Generation Computer Systems Journal
	2022	Denner Silva et al.	Journal of Network & Systems Management
	2019	J. Li et al.	IEEE Access
	2017	S. Singh et al.	Journal of Ambient Intelligence and Humanized Computing
	2018	C. Kejun	Journal of Hardware and Systems Security
	2015	D. Darwish	International Journal of Computing Academic Research

Five-layer architecture	2022	Raja Gopal, S. et al.	International Journal of System Assurance Engineering and Management
	2017	P. Sethi et al.	Journal of Electrical and Computer Engineering
	2022	Jinyuan Xu et al.	Artificial Intelligence in Agriculture
	2013	S. Omar et al.	International Journal of Computer Networks
	2022	A. Khaled et al.	Journal of Cloud Computing
	2021	M. Yildirim et al.	European Journal of Science and Technology

IoT device security concerns arise in a variety of situations that include technological, ethical, and privacy issues. IoT devices will be considered secure if the following security requirements are met, such as secure authentication, secure download and data transfer, IoT data security, secure access to data by an authorized person. Figure 2 shows the threats and security requirements of a three-layer architecture. The sensor layer is the perception layer and is responsible for identifying devices and collecting information from them. The sensors are selected according to the requirements of the applications. Information that is collected at the level of sensors can be information about location, changes in the air, about the environment, about movement, about vibration, etc. At the same time, this information is the main goal of attackers who want to use them to achieve their own goals. Therefore, most threats are related to the level of sensors and lightweight encryption methods have a special role.

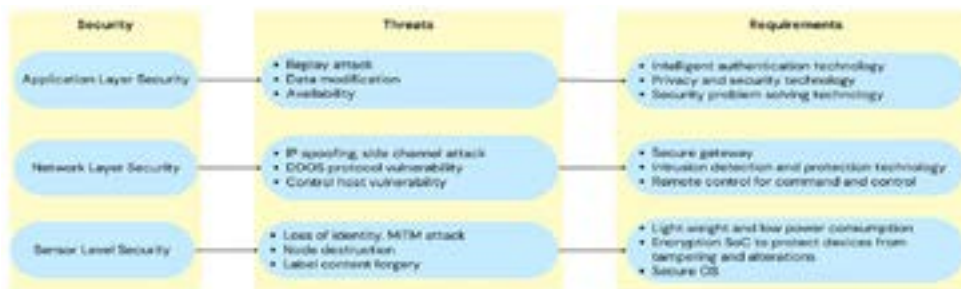


Fig. 2. Threats and security requirements based on a three-layer architecture

Lightweight cryptography in IoT

Lightweight cryptography is divided into two ciphers: a symmetric cipher and an asymmetric cipher. If a symmetric cipher uses the same key for both encryption and decryption, an asymmetric cipher uses the public key for encryption and the private key for decryption. Symmetric encryption provides security and high speed, along with this, an asymmetric cipher is complex and relatively slow, but at the same time, it ensures the confidentiality and integrity of data. The symmetric cipher is divided into three groups: the lightweight block cipher, the lightweight hash function, and the lightweight stream cipher. The classification of lightweight cryptography, the most common algorithms and attacks is shown in Figure 3.



Fig. 3. Classification of algorithms by structure

Research in the direction of lightweight cryptography was started in 2004. In 2007, the PRESENT block cipher was developed and published, which was registered in the ISO / IEC 29192 standard. The US National Security Agency published the SIMON / SPECK lightweight block cipher with a small ROM size and implemented it on a microprocessor (2014). The main stages in the development of lightweight cryptography are shown in Figure 4.



Fig. 4. Stages of development of lightweight cryptography

Lightweight Cryptography Options

One of the important properties of lightweight algorithms is the non-linearity of the coordinate functions of the round transformation (Poschmann, 2009: 516). The nonlinearity of the coordinate function of the output block is determined using the minimum number of rounds.

There are randomly selected options. For these parameters $x, x', p \in S_n, p \neq 0$.

, vectors are encrypted $x, x', x + p, x' + p$. The transformation is non-linear if the condition is satisfied $f(x) + f(x + p) \neq f(x') + f(x' + p)$.

The superposition of non-linear functions produces a system of linear equations:

$$f_1(x, y) = x \oplus y \oplus xy$$

$$f_2(x, y) = x \oplus xy$$

When setting functions instead of f , the following system of equations is obtained:

$$f_1(f_1(x, y), f_2(x, y)) = x \oplus y \oplus xy$$

When learning lightweight cryptography, you need to consider the main criteria for security, cost and performance. In the case of block ciphers, the key length provides a trade-off between security and cost, the number of rounds is a trade-off between security and performance, and the hardware is a trade-off between cost and performance (Figure 5). Considering all three trade-offs at the same time is challenging, and two design goals are mainly considered — safety and low cost, safety and productivity, or low cost and productivity. For example, a secure and high performance hardware implementation can be achieved with a pipelined architecture that also includes many side-channel attack countermeasures. The resulting structure will occupy a large area, which is correlated with high costs. On the other hand, it is possible to develop a secure and inexpensive hardware implementation with the disadvantage of limited performance.

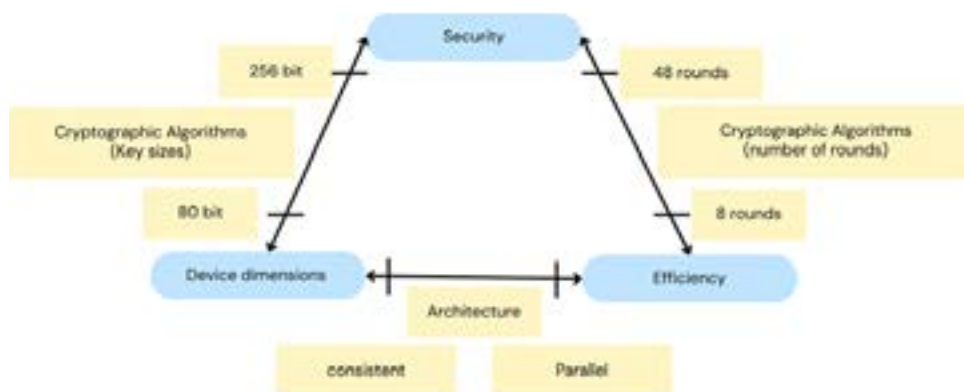


Fig. 5. Lightweight Cryptography Options

Table 2 summarizes the main challenges of implementing cryptographic algorithms in IoT devices and lightweight cryptography solutions.

Table 2. Lightweight Cryptography Options

Implementation challenges in IoT devices	Lightweight Cryptography Solutions
limited memory (registers, RAM, ROM)	small block size (64-bit or less)
low processing power	small key size (80-bit or less)
small physical area	simple logic and simple calculations
low battery (no battery)	simple key planning
work in real time	reliable structure (SPN or FN) [5]

Shannon in his work proposed to use several stages of replacement and permutation, thereby creating a reliable block cipher (Shannon E., et al, 1949:59). Such a scheme is called an SP or FN network. SP- processes the data through a series of substitutions (S-box) and permutations (table), changing the data and finalizing it for the next round. FN is a multi-round cipher that divides the input message into two parts and operates on only one part in each round of encryption or decryption.

ISO/IEC 29192 "Lightweight Cryptography" is an eight-part standard that defines lightweight cryptographic algorithms for privacy, authentication, identification, security, and key exchange (Table-3) (<https://webstore.iec.ch>).

Table - 3. Lightweight Cryptography Standards

Standard	Description	Summary
ISO/IEC 29192-1:2012	General Information	Terms and Definitions; Safety requirements, Classification requirements; Implementation requirements for mechanisms
ISO/IEC 29192-2:2019	Block Encryption	Three block ciphers are described that are suitable for applications requiring lightweight cryptographic implementations: - PRESENT: simplified block cipher / block size 64 bits / key size 80/128 bits; - CLEFIA: lightweight block cipher / block size 128 bits / key size 128/192/256 bits; - LEA: Lightweight block cipher/block size 128 bits/key size 128/192/256 bits
ISO/IEC 29192-3:2012	Stream Ciphers	Defines keystream generators for lightweight stream ciphers: - Enocoro: lightweight keystream generator with 80/128 bit key size; - Trivium: lightweight keystream generator with 80 bit key size
ISO/IEC 29192-4:2013	Mechanisms using asymmetric methods	Defines lightweight mechanisms that use asymmetric methods: - one-way authentication mechanism based on discrete logarithms on elliptic curves; - Authenticated Lightweight Key Exchange (ALIKE) mechanism for one-way authentication and session key establishment; - identification-based signature mechanism
ISO/IEC 29192-4:2013/AMD1:2016	Mechanisms using asymmetric methods	Updated version

ISO/IEC 29192-5:2016	hash function	Defines hash functions suitable for applications requiring lightweight cryptographic implementations. - PHOTON: lightweight hash function/permutation size 100/144/196/256/288 bits/calculated hash codes 80/128/160/224/256 bits. - SPONGENT: lightweight hash function/permutation size 88/136/176/240/272 bits/calculated hash codes 88/128/160/224/256 bits. - Lesamnta-LW: lightweight hash function/permutation size 384 bits/computed hash code 256 bits
ISO/IEC 29192-6:2019	Message Authentication Code (MACs)	MAC algorithms suitable for applications requiring lightweight cryptographic mechanisms have been defined: - a mechanism for ensuring data integrity; - message authentication mechanism
ISO/IEC 29192-7:2019	Broadcast Authentication Protocols	Defines broadcast authentication protocols, which are protocols that provide data integrity and entity authentication in a broadcast setting
ISO/IEC 29192-8:2022	Authenticated Encryption	An authenticated encryption method suitable for applications requiring lightweight cryptographic mechanisms is described. The method processes the data string with the following security objectives: - data confidentiality; - data integrity

In the table-4 presents scientific papers published in 2022–2023 devoted to the problems of cryptography and security in the context of the Internet of things (IoT). The table lists the titles of articles, authors, year of publication, and a brief abstract for each article.

Table - 4. Review of works on lightweight IoT encryption for 2022–2023.

Authors	Year	Annotation
P. Prakasam et al.	2022	The paper proposes and implements a hybrid lightweight cryptographic authentication scheme with low latency, area and optimal power, which uses the 8-bit keying principle.
S. Blank et al.	2022	The paper presents the practical results of research of 12 cryptographic algorithms on a test bench
A. Kumar et al.	2022	The paper presents a literature review of post-quantum cryptography for IoT networks, discusses the problems and directions of research in real-time applications.
K. Tsantikidou et al.	2022	The paper investigates well-known lightweight cryptographic algorithms and their architecture. The analysis of security algorithms, architecture and hardware limitations in healthcare applications is given.
S. Alshehri et al.	2022	The paper proposes an attribute-based access control scheme for IoT using the Hyperledger Fabric blockchain to solve security problems. Performance metrics are measured based on latency, throughput, and storage overhead,
T. Goyal et al.	2022	The paper presents the results of hardware implementation of PRESENT, AES, ECDH, DH and RSA cryptography algorithms.

Sheeja S. et al.	2022	The paper presents approaches to authentication in the IoT-Cloud architecture.
M. Jammula et al.	2022	The paper presents the LWC-ABE method for improving security performance against various attacks in the IoT environment.
A. Ahmed et al.	2023	The paper investigates integration systems between authentication and encryption to preserve confidentiality when passing messages between IoT devices.
J. Chauhan et al.	2023	This paper presents a comparison of their memory performance, latency and throughput, area (GE), key and block size, and other parameters of hardware and software efficient lightweight algorithms.

Lightweight algorithms are designed to protect the transmitted information during the interaction of the Internet of things, at the same time, lightweight algorithms are used in miniature devices where a limited amount of electronic resources is used to ensure security. Today, many researchers from different countries continue to search for weaknesses and identify reliable and efficient algorithms for lightweight cryptography. The choice of the optimal cryptographic algorithm depends on the specific requirements and constraints of the system, including hardware resources, power, power consumption, and performance requirements.

The hardware implementation of lightweight algorithms is determined by the following parameters (Figure 6):

- key size;
- block size;
- logical process;
- Energy consumption;
- throughput;
- hardware performance.

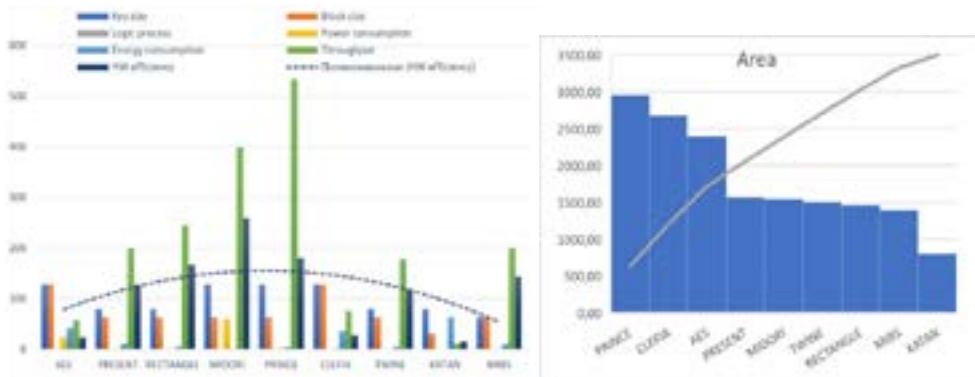


Fig. 6. Numerical characteristics of the hardware implementation of lightweight algorithms

The diagram shows the number of gates in hardware implementation and you can see that the high throughput is achieved by the “PRINCE” algorithm with a

minimum power consumption indicator (Figure 6). Various FPGA families are used as a platform for implementing lightweight algorithms.

The software implementation of lightweight algorithms is described using parameters (Figure 7):

- key size;
- block size;
- RAM size;
- Energy consumption;
- throughput;
- software performance;

Additionally, diagrams are given with the dimensions of the read-only memory and the time delay.

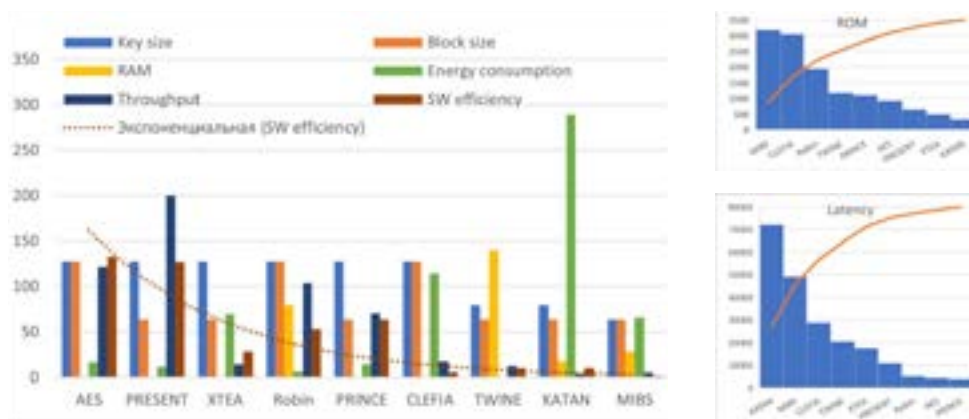


Fig.7. Numerical characteristics of software implementation of lightweight algorithms

The diagram shows several lightweight algorithms such as SPN and FN. According to the diagram, in software implementation, the “PRESENT” algorithm dominates in terms of bandwidth, the “MIBS” algorithm in terms of power consumption and memory space occupied in ROM, and the lowest time delay is demonstrated by the “PRINCE” algorithm (Figure 7).

Conclusion

The Internet of Things (IoT) is rapidly finding its way into our modern lives, seeking to improve the quality of life by connecting various smart devices, technologies and applications. In general, if you choose the right lightweight security algorithm, then IoT will automate everything that surrounds us. The wide range of IoT applications in various fields creates a demand for lightweight cryptographic algorithms with different requirements. Smart home appliances such as smart TV, smart refrigerator, smart kettle, smart light bulbs, etc. require little memory and little processing power. Lightweight algorithms must support simple hardware and software implementation for fast execution, so that it can be

easily embedded in software for encryption, protection transmission and storage data in real applications. When using IoT, an organization faces the challenge of managing, monitoring and securing huge amounts of data and connections from disparate devices.

The paper presents the architecture of IoT devices, an overview of lightweight cryptographic algorithms, threats and security requirements based on a three-layer architecture, lightweight cryptography primitives regarding key size, block length, number of rounds and structure, as well as classification and parameters of lightweight cryptography. Descriptions of sections of ISO/IEC 29192 "Lightweight cryptography" were presented. The works of the authors, who in their studies considered various types of architectures in relation to the tasks to be solved, are summarized. The security problems of IoT devices are considered, as well as a review of works on lightweight IoT encryption for 2022-2023.

This article provides an overview of various studies about layered IoT architectures. Along with the exponential growth in the number of connected devices, every thing in the IoT transmits data packets that require reliable connectivity, storage, and security. The hardware and software implementation of lightweight algorithms is analyzed in terms of such parameters as key size, block size, logical process, power consumption, bandwidth, hardware performance, RAM size, software. Taking into account the growth of attacks and threats on the IoT infrastructure, the discussed issues of lightweight cryptography need to be further studied and developed to ensure sufficient security.

REFERENCES

- Aguru A., Babu E., Nayak S., Sethy A., Verma A., 2022 — *Aguru A., Babu E., Nayak S., Sethy A., Verma A.* Integrated Industrial Reference Architecture for Smart Healthcare in Internet of Things: A Systematic Investigation. *Algorithms*, 15: 309. <https://doi.org/10.3390/a15090309>. (in Eng.).
- Ahmed A.A., Malebary S.J., Ali W., Alzahrani A.A., 2023 — *Ahmed A.A., Malebary S.J., Ali W., Alzahrani A.A.* A Provable Secure Cybersecurity Mechanism Based on Combination of Lightweight Cryptography and Authentication for Internet of Things. *Mathematics*, 11(1):220. (in Eng.).
- Alaghbari K., Md S., Mohamad H., Hussain A., Alam M., 2022 — *Alaghbari K., Md S., Mohamad H., Hussain A., Alam M.* Complex event processing for physical and cyber security in datacentres - recent progress, challenges and recommendations. *Journal of Cloud Computing*. 11:65. (in Eng.).
- Alshehri S., Bamasag O., 2022 — *Alshehri S., Bamasag O.* AAC-IoT: Attribute Access Control Scheme for IoT Using Lightweight Cryptography and Hyperledger Fabric Blockchain. *Applied Sciences*, 12(16):8111. (in Eng.).
- Blanc S., Lahmadi A., Gouguec K., Minier M., Sleem L., 2022 — *Blanc S., Lahmadi A., Gouguec K., Minier M., Sleem L.* Benchmarking of lightweight cryptographic algorithms for wireless IoT networks. *Wireless Networks*. 28(8). Pp. 3453–3476. (in Eng.).
- Bonani P., 2022 — *Bonani P.* Internet of Things (IoT), Three-Layer Architecture, Security Issues and Counter Measures. *ICT Analysis and Applications*. Springer Nature Singapore, 2022. (in Eng.).
- Chauhan J.A., Patel R.A., Parikh S., Modi N., 2022 — *Chauhan J.A., Patel R.A., Parikh S., Modi N.* An Analysis of Lightweight Cryptographic Algorithms for IoT-Applications. *Advancements in Smart Computing and Information Security*. Springer Nature Switzerland, 2023. Pp. 201–216. (in Eng.).
- Chen K., Zhang S., Li Z., Zhang Y., Deng Q., Ray S., Jin Y., 2018 — *Chen K., Zhang S., Li Z., Zhang Y., Deng Q., Ray S., Jin Y.* Internet-of-Things Security and Vulnerabilities: Taxonomy,

Challenges, and Practice. *Journal of Hardware and Systems Security*, 2(2). Pp. 97–110. (in Eng.).

Darwish D., 2015 — *Darwish D.*, "Improved layered architecture for Internet of Things," *Int. J. Comput. Acad. Res. (IJCAR)*, 4. Pp. 214–223. (in Eng.).

Eryk C., Andy A., Jara F., Jonathan S., Michael Z., Burkhard S., 2022 — *Eryk C., Andy A., Jara F., Jonathan S., Michael Z., Burkhard S.* Landscape of IoT security, *Computer Science Review*, 44: 100467. ISSN 1574–0137. <https://doi.org/10.1016/j.cosrev.2022.100467>. (in Eng.).

Fu Bing, 2016 — *Fu Bing*. The research of IOT of agriculture based on three layers architecture. 2016 2nd International Conference on Cloud Computing and Internet of Things (CCIoT), 2016. Pp. 162–165. doi: 10.1109/CCIoT.2016.7868325.

Gaikwad P.P., Gabhane J.P., Golait S.S., 2015 — *Gaikwad P.P., Gabhane J.P., Golait S.S.* 3-level secure Kerberos authentication for Smart Home Systems using IoT. 2015 1st International Conference on Next Generation Computing Technologies (NGCT), 2015. Pp. 262–268, doi: 10.1109/NGCT.2015.7375123. (in Eng.).

Goyal T.K., Sahula V., Kumawat D., 2022 — *Goyal T.K., Sahula V., Kumawat D.* Energy efficient lightweight cryptography algorithms for IoT devices. *IETE Journal of Research*, 68(3). Pp. 1722–1735. (in Eng.).

Gubbi J., Buyya R., Marusic S., Palaniswami M., 2013 — *Gubbi J., Buyya R., Marusic S., Palaniswami M.* Internet of Things (IoT): a vision, architectural elements, and future directions. *Future Generation Computer Systems*, 29(7). Pp.1645–1660. (in Eng.).

IEC, 2023 - *IEC* <https://webstore.iec.ch/> (in Eng.).

Jammula M., Vakamulla V.M., Kondoju S.K., 2022 — *Jammula M., Vakamulla V.M., Kondoju S.K.* Hybrid lightweight cryptography with attribute-based encryption standard for secure and scalable IoT system. *Connection Science*, 34(1). Pp. 2431–2447. (in Eng.).

Jinyuan Xu, Baoxing Gu, Guangzhao Tian, 2022 — *Jinyuan Xu, Baoxing Gu, Guangzhao Tian.* Review of agricultural IoT technology. *Artificial Intelligence in Agriculture*, 6. Pp.10–22. ISSN 2589–7217. <https://doi.org/10.1016/j.aiaa.2022.01.001>. (in Eng.).

Keoh S.L., Kumar S.S., Tschofenig H., 2014 — *Keoh S.L., Kumar S.S., Tschofenig H.* Securing the Internet of Things: A Standardization Perspective. *IEEE Internet of Things Journal*, 1(3). Pp. 265–275. doi: 10.1109/JIOT.2014.2323395. (in Eng.).

Kumar A. et al., 2022 — *Kumar A. et al.* Securing the future internet of things with postquantum cryptography. *Security and Privacy*, 5(2): e200. (in Eng.).

Li J. et al., 2019 — *Li J. et al.* A Remote Monitoring and Diagnosis Method Based on Four-Layer IoT Frame Perception. *IEEE Access*, 7:144324–144338. doi: 10.1109/ACCESS.2019.2945076. (in Eng.).

Mahmoud Parto, Christopher Saldana, Thomas Kurfess, 2020 — *Mahmoud Parto, Christopher Saldana, Thomas Kurfess.* A Novel Three-Layer IoT Architecture for Shared, Private, Scalable, and Real-time Machine Learning from Ubiquitous Cyber-Physical Systems. *Procedia Manufacturing*, 48. Pp. 959–967. ISSN 2351–9789. <https://doi.org/10.1016/j.promfg.2020.05.135> (in Eng.).

Mashal I., Alsaryrah O., Chung T., Yang C., Kuo W., Agrawal D., 2015 — *Mashal I., Alsaryrah O., Chung T., Yang C., Kuo W., Agrawal D.* Choices for interaction with things on Internet and underlying issues. *Ad Hoc Netw*, 28. Pp. 68–90. (in Eng.).

Nurlan Z., Kokenovna T.Z., Othman M., Adamova A., 2021 — *Nurlan Z., Kokenovna T.Z., Othman M., Adamova A.* Resource Allocation Approach for Optimal Routing in IoT Wireless Mesh Networks. *IEEE Access*, 9:153926–153942. doi: 10.1109/ACCESS.2021.3123903. (in Eng.).

Omar S., Masud M., 2013 — *Omar S., Masud M.* Towards Internet of Things: Survey and Future Vision. *International Journal of Computer Networks*, 5(1): 17. (in Eng.).

Pallavi S., Smruti R.S., 2017 — *Pallavi S., Smruti R.S.* Internet of Things: Architectures, Protocols, and Applications. *Journal of Electrical and Computer Engineering*, 2017:9324035. <https://doi.org/10.1155/2017/9324035>. (in Eng.).

Paul B., 2022 — *Paul B.* Internet of Things (IoT), Three-Layer Architecture, Security Issues and Counter Measures. *ICT Analysis and Applications*, 314. Pp. 23–34. https://doi.org/10.1007/978-981-16-5655-2_3. (in Eng.).

Poschmann A., 2009 — *Poschmann A.* Lightweight Cryptography: Cryptographic Engineering for a Pervasive World. Ph.D. Thesis. Ruhr University Bochum, 2009. (in Eng.).

Prakasam P. et al., 2022 — *Prakasam P. et al.* Low latency, area and optimal power hybrid lightweight cryptography authentication scheme for internet of things applications. *Wireless Personal Communications*, 126(1). Pp. 351–365. (in Eng.).

Raja G.S., Prabhakar V.S., 2022 — *Raja G.S., Prabhakar V.S.* Intelligent edge based smart farming with LoRa and IoT. *International Journal of System Assurance Engineering and Management*, 2022. Pp. 1–7. (in Eng.).

Rana Muhammad, Quazi Mamun, Rafiqul Islam, 2022 — *Rana Muhammad, Quazi Mamun, Rafiqul Islam*, Lightweight cryptography in IoT networks: A survey. *Future Generation Computer Systems*, 129. Pp. 77–89. ISSN 0167–739X. <https://doi.org/10.1016/j.future.2021.11.011>. (in Eng.).

Shannon C.E., 1949 — *Shannon C.E.* Communication theory of secrecy systems. *The Bell System Technical Journal*, 28(4). Pp. 656–715. doi: 10.1002/j.1538-7305.1949.tb00928.x. (in Eng.).

Sheeja S. et al., 2022 — *Sheeja S. et al.* Towards an Optimal Security Using Multifactor Scalable Lightweight Cryptography for IoT. 3rd International Conference on Communication, Computing and Industry 4.0 (C2I4). IEEE, 2022. Pp. 1–6. (in Eng.).

Silva D., Heideker A., Zyrianoff I.D. et al., 2022 — *Silva D., Heideker A., Zyrianoff I.D. et al.* A Management Architecture for IoT Smart Solutions: Design and Implementation. *J Netw Syst Manage*, 30(35). (in Eng.).

Singh D., Pati B., Panigrahi C., Swagatika S., 2020 — *Singh D., Pati B., Panigrahi C., Swagatika S.* Security Issues in IoT and their Countermeasures. *Smart City Applications*, 10. (in Eng.).

Singh S., Sharma P.K., Moon S.Y. et al., 2017 — *Singh S., Sharma P.K., Moon S.Y. et al.* Advanced lightweight encryption algorithms for IoT devices: survey, challenges and solutions. *J Ambient Intell Human Comput*. <https://doi.org/10.1007/s12652-017-0494-4>. (in Eng.).

Tsantikidou K., Sklavos N., 2022 — *Tsantikidou K., Sklavos N.* Hardware Limitations of Lightweight Cryptographic Designs for IoT in Healthcare. *Cryptography*, 6(3):45 (in Eng.).

Vermesan O., Friess P., Guillemin P., Gusmeroli S., Sundmaeker H., Bassi A., Jubert I., Mazura M., Harrison M., Eisenhauer M., Doody P., 2009. — *Vermesan O., Friess P., Guillemin P., Gusmeroli S., Sundmaeker H., Bassi A., Jubert I., Mazura M., Harrison M., Eisenhauer M., Doody P.* Internet of Things Strategic Research Roadmap. (in Eng.).

Yildirim Muhammed, Demiroğlu Uğur, Şenol Bilal, 2021 — *Yildirim Muhammed, Demiroğlu Uğur, Şenol Bilal.* An in-depth exam of IoT, IoT Core Components, IoT Layers, and Attack Types. *European Journal of Science and Technology*. 10.31590/ejosat.1010023 (in Eng.).

Zhong C., Zhu Z., Huang R., 2017 — *Zhong C., Zhu Z., Huang R.* Study on the IOT Architecture and Access Technology. 16th International Symposium on Distributed Computing and Applications to Business, Engineering and Science, 2017. Pp. 113–116. doi: 10.1109/DCABES.2017.32. (in Eng.).

NEWS OF THE NATIONAL ACADEMY OF SCIENCES OF THE REPUBLIC OF KAZAKHSTAN
PHYSICO-MATHEMATICAL SERIES

ISSN 1991-346X

Volume 2. Number 346 (2023). 21–31

<https://doi.org/10.32014/2023.2518-1726.181>

UDC: 004.49

IRSTI: 28.23.25

© **G. Alpysbay^{1*}, A. Bedelbayev¹, O. Ussatova^{1,2}, A. Zhumabekova¹,
Edzard Höfig³, 2023**

¹Al-Farabi Kazakh National University, Almaty, Kazakhstan;

²Institute of Information and Computational Technologies, Almaty, Kazakhstan;

³Berlin University of Applied Sciences and Technology, Berlin, Germany.

E-mail: gulbanu.alpysbay@gmail.com

APPLICATION OF MACHINE LEARNING ALGORITHM IN THE ANALYSIS OF MALICIOUS SOFTWARE

Alpysbay G.E. — Lecturer. Al-Farabi Kazakh National University. Department of Information technology. 050040. Almaty, Kazakhstan

E-mail: gulbanu.alpysbay@gmail.com. ORCID: 0000-0002-3766-2396;

Bedelbayev A.A. — Candidate of science (physical and mathematical), professor. Al-Farabi Kazakh National University. Department of Information technology. 050040. Almaty, Kazakhstan

E-mail: agyn08@yandex.ru. ORCID: 0000-0001-9839-4156;

Ussatova O.A. — PhD, chief scientific secretary, senior researcher. Institute of Information and Computational Technologies. 050010. Almaty, Kazakhstan. PhD, acting associate professor. Al-Farabi Kazakh National University. Department of Information technology. 050040. Almaty, Kazakhstan

E-mail: uoa_olga@mail.ru. ORCID: 0000-0002-5276-6118;

Zhumabekova A.T. — Lecturer. Al-Farabi Kazakh National University. Department of Information technology. 050040. Almaty, Kazakhstan

E-mail: zhumabekova2702@gmail.com. ORCID: 0000-0003-4242-7988;

Edzard Höfig — Prof. Dr.-Ing., Berlin University of Applied Sciences and Technology. Department of Media and Computer Science, Berlin, Germany

E-mail: edzard.hoefig@bht-berlin.de.

Abstract. Rapid development and widespread use of information technologies and the Internet have many benefits, it is important not to ignore the negative situations that occur there. One of these negative situations is the proliferation and development of malicious programs, which lead to the failure of many devices and software equipment, and the disclosure of confidential information. Malware is the vehicle for many computer attacks and security breaches. Malware analysis uses techniques from several different fields, such as program analysis and network analysis, to study malicious patterns to gain a deeper understanding of several aspects, including their behavior and how they evolve over time. The day-by-day development of malicious software equipment, the increase in types and complexity of their structure has greatly complicated the work of identifying and classifying

them. For this reason, the importance of using artificial intelligence systems, including machine learning algorithms, is increasing in the field of information protection. Machine learning is a branch of artificial intelligence that involves algorithms and processes that "learn" and can generalize past data and insights to predict future outcomes. Basically, machine learning is a set of mathematical methods implemented in computer systems that provide the process of obtaining information, identifying patterns, and drawing conclusions from data. The article analyzes the type of malicious software in a certain format and classifies it with the help of a machine learning algorithm based on specific selected features. Here, the role of the selection of received features for training is very important. Machine learning works well when the input data is good.

Keywords: information security, malware analysis, machine learning, feature selection, kNN

© Г. Алпысбай^{1*}, А. Бедельбаев¹, О. Усатова^{1,2}, А. Жұмабекова¹,
Эдзард Хофиг³, 2023

¹Әл-Фараби атындағы Қазақ ұлттық университеті, Алматы, Қазақстан;

²Ақпараттық және есептеуіш технологиялар институты, Алматы, Қазақстан;

³Берлин қолданбалы ғылымдар және технологиялар университеті,
Берлин, Германия.

E-mail: gulbanu.alpysbay@gmail.com

ЗИЯНДЫ БАҒДАРЛАМАЛЫҚ ЖАБДЫҚТАРДЫ ТАЛДАУДА МАШИНАЛЫҚ ОҚЫТУ АЛГОРИТМІН ҚОЛДАНУ

Алпысбай Г.Е. — Оқытушы. Әл-Фараби атындағы ҚазҰУ. Ақпараттық технологиялар факультеті. 050040. Алматы, Қазақстан

E-mail: gulbanu.alpysbay@gmail.com. ORCID: 0000-0002-3766-2396;

Бедельбаев А.А. — Физика-математика ғылымдарының кандидаты, профессор. Әл-Фараби атындағы ҚазҰУ. Ақпараттық технологиялар факультеті. 050040. Алматы, Қазақстан

E-mail: agyn08@yandex.ru. ORCID: 0000-0001-9839-4156;

Усатова О.А. — PhD, бас ғылыми хатшы, аға ғылыми қызметкер. Ақпараттық және есептеу технологиялары институты. 050010. Алматы, Қазақстан. PhD, доцент м.а. Әл-Фараби атындағы ҚазҰУ. Ақпараттық технологиялар факультеті. 050040. Алматы, Қазақстан

E-mail: uoa_olga@mail.ru. ORCID: 0000-0002-5276-6118;

Жұмабекова А.Т. — Оқытушы. Әл-Фараби атындағы ҚазҰУ. Ақпараттық технологиялар факультеті. 050040. Алматы, Қазақстан

E-mail: zhumabekova2702@gmail.com. ORCID: 0000-0003-4242-7988;

Эдзард Хофиг — Берлин қолданбалы ғылымдар және технологиялар университеті профессоры. Медиа және информатика кафедрасы, Берлин, Германия

E-mail: edzard.hoefig@bht-berlin.de.

Аннотация. Ақпараттық технологиялар мен интернет желісінің қарқынды дамуы мен кеңінен қолданылуының тиімді тұстары көп болғанымен, ондағы орын алатын келеңсіз жағдайларды назардан тыс қалдырмау керек. Осындай келеңсіз жағдайлардың бірі — зиянды бағдарламалардың көбеюі және

дамуы, бұл көптеген құрылғылар мен бағдарламалық құрал-жабдықтардың істен шығуына, құпия ақпараттың ашылуына әкеліп соғуда. Зиянды бағдарлама көптеген компьютерлік шабуылдар мен қауіпсіздікті бұзудың құралы болып табылады. Зиянды бағдарламалық құралды талдау бірнеше аспектілерді, соның ішінде олардың мінез-құлқы және уақыт өте келе қалай дамып жатқанын тереңірек түсіну үшін зиянды үлгілерді зерттеу үшін бағдарламаларды талдау және желіні талдау сияқты бірнеше түрлі салалардағы әдістерді пайдаланады. Зиянды бағдарламалық жабдықтың күн санап дамуы, олардың түрлерінің ұлғаюы және құрылымының күрделілігі оларды анықтау және жіктеу жұмысын айтарлықтай қиындатады. Осы себепті ақпаратты қорғау саласында жасанды интеллект жүйелерін, соның ішінде машиналық оқыту алгоритмдерін пайдаланудың маңыздылығы артып келеді. Машиналық оқыту — бұл «үйренетін» алгоритмдер мен процестерді қамтитын және болашақ нәтижелерді болжау үшін өткен деректер мен түсініктерді жалпылай алатын жасанды интеллект саласы. Негізінде, машиналық оқыту — бұл ақпарат алу, заңдылықтарды анықтау және деректерден қорытындылар жасау процесін қамтамасыз ететін компьютерлік жүйелерде жүзеге асырылатын математикалық әдістердің жиынтығы. Мақалада нақты бір зиянды бағдарламалық жасақтаманың түріне талдау жасалынды және таңдап алынған белгілерге негізделіп машиналық оқыту алгоритмінің көмегімен жіктеледі. Мұнда оқыту үшін алынған белгілерді таңдаудың рөлі өте маңызды. Кіріс деректері сапалы болған кезде машиналық оқыту жақсы жұмыс істейді.

Түйін сөздер: ақпараттық қауіпсіздік, зиянды бағдарламаларды талдау, машиналық оқыту, белгілерді таңдау, kNN

© Г. Алпысбай^{1*}, А. Бедельбаев¹, О. Усатова^{1,2}, А. Жумабекова¹,
Эдзарт Хофиг³, 2023

¹КазНУ имени аль-Фараби, Алматы, Казахстан;

²Институт информационно-вычислительных технологий,
Алматы, Казахстан;

³Берлинский университет прикладных наук и технологий, Берлин, Германия.
E-mail: gulbanu.alpysbay@gmail.com

ПРИМЕНЕНИЕ АЛГОРИТМА МАШИННОГО ОБУЧЕНИЯ ДЛЯ АНАЛИЗА ВРЕДНОСНОГО ПО

Алпысбай Г.Е. — Преподаватель. КазНУ им. аль-Фараби. Факультет Информационных технологий. 050040. Алматы, Казахстан

E-mail: gulbanu.alpysbay@gmail.com. ORCID: 0000-0002-3766-2396;

Бедельбаев А.А. — Кандидат физико-математических наук, профессор, КазНУ им. аль-Фараби. Факультет Информационных технологий. 050040. Алматы, Казахстан

E-mail: agun08@yandex.ru. ORCID: 0000-0001-9839-4156;

Усатова О.А. — PhD, главный ученый секретарь, старший научный сотрудник. Институт Информационно-вычислительных технологий. 050010. Алматы, Казахстан. И.о. доцента.

КазНУ им. аль-Фараби. Факультет Информационных технологий. 050040. Алматы, Казахстан
E-mail: uoa_olga@mail.ru. ORCID: 0000-0002-5276-6118;

Жумабекова А.Т. — Преподаватель, КазНУ им. аль-Фараби. Факультет Информационных технологий. 050040. Алматы, Казахстан
E-mail: zhumabekova2702@gmail.com. ОРЦИД: 0000-0003-4242-7988;

Эдзард Хофиг — Профессор, доктор технических наук, Берлинский университет прикладных наук и технологий. Департамент медиа и компьютерных наук. Берлин, Германия
E-mail: edzard.hoefig@bht-berlin.de.

Аннотация. Быстрое развитие и широкое использование информационных технологий и Интернета имеют множество преимуществ, важно не игнорировать возникающие там негативные ситуации. Одной из таких негативных ситуаций является распространение и развитие вредоносных программ, приводящих к выходу из строя многих устройств и программного оборудования, разглашению конфиденциальной информации. Вредоносное ПО является средством для многих компьютерных атак и нарушений безопасности. Анализ вредоносных программ использует методы из нескольких различных областей, таких как анализ программ и сетевой анализ для изучения вредоносных шаблонов, чтобы получить более глубокое понимание нескольких аспектов, включая их поведение и то, как они развиваются с течением времени. Ежедневное развитие вредоносных программных средств, увеличение типов и сложности их структуры значительно усложнили работу по их выявлению и классификации. По этой причине возрастает важность использования систем искусственного интеллекта, в том числе алгоритмов машинного обучения, в сфере защиты информации. Машинное обучение — это область искусственного интеллекта, которая включает алгоритмы и процессы, которые «обучаются» и могут обобщать прошлые данные и идеи для прогнозирования будущих результатов. По сути, машинное обучение — это набор математических методов, реализованных в компьютерных системах, которые обеспечивают процесс получения информации, выявления закономерностей и вывода из данных. В статье анализируется тип вредоносного ПО в определенном формате и классифицируется с помощью алгоритма машинного обучения на основе конкретных выбранных признаков. Здесь очень важна роль отбора полученных признаков для обучения. Машинное обучение работает хорошо, когда входные данные хороши.

Ключевые слова: информационная безопасность, анализ вредоносных программ, машинное обучение, выбор признаков, kNN

Introduction

Malware is program that harms a user, computer or network and includes viruses, trojans, worms, rootkits, spyware, adware, etc. Studying the behavior of malicious programs remains an urgent problem in the field of information security. Since the appearance of the first malware, the variety of malware, its complexity, the number of new models and the speed with which they appear, as well as the range

of threats, have increased. The number of malicious programs in existing databases is huge, and new malicious programs appear every day. For example, the German laboratory AV-TEST registers more than 450,000 new malicious programs per day. Traditional anti-virus (reactive protection) based on checking signatures against a database of malicious codes cannot cope with emerging, previously unknown threats supplemented by an expert analyst. Therefore, tools for static analysis of a portable executable (PE) file and dynamic analysis of malware are emerging that can study pattern behavior that is not stored in a signature database. In addition, methods of working with anti-virus scanners by malicious programs are becoming more complicated (AV-TEST, 2022).

Malware analysis is the process of taking programs apart to understand how they work, and how to detect, disable, and remove them. There are millions of malwares on the Internet and this number is growing every day, so analyzing them is very important for everyone responsible for computer security. There are two main approaches to malware analysis: static and dynamic. Static analysis is the study of malware without running it. During dynamic analysis, the malware must be running. Both categories include basic and advanced techniques (Egele et al., 2012). The malware analysis process can often be sped up by making an educated guess about the malware's target and then confirming it. Of course, the accuracy of your predictions will depend on your knowledge of what malware typically does.

Methods and materials

Machine learning is a class of artificial intelligence methods, the characteristic feature of which is not the direct solution of a problem but learning by applying solutions to many similar problems. From the point of view of the application of machine learning in the field of information security, it can be classified into two main categories: pattern recognition and anomaly detection. It is not always possible to distinguish between pattern recognition and anomaly detection, but there is a well-defined task for each specific problem. In pattern recognition, we try to find obvious or undefined characteristics hidden in the data. These characteristics, labeled and combined into a set of features, can be used to train an algorithm that identifies the type of data with the same set of characteristics. Anomaly detection is the exact opposite of the task of obtaining data (Clarence Chio et al., 2018; Michael Sikorski et al., 2018). The main goal here is to determine the concept of normality that describes most of the data in the studied set instead of learning the characteristic patterns in the actual data set. After that, a deviation from the established normality is considered an anomaly.

A common misconception is that anomaly detection is the process of recognizing a set of "normal" patterns and distinguishing them from a set of "abnormal" patterns. The patterns obtained by the pattern recognition method must, of course, be taken from the studied data used to pre-train (train) the algorithm. On the other hand, when using the anomaly detection methodology, there may be an infinite number of anomalous samples with characteristics that match the given description of the

errors (outliers), even from hypothetical data that do not actually exist in the study (Ucci et al., 2017).

Malicious code detection by machine learning consists of several stages (figure 1):

- Obtaining information from the operating system about the files, processes, and software in use. All the data required for this machine learning can be obtained using Process Monitor and Process Hacker utilities for Windows operating system.
- Preparation of data for machine learning.
- Normalization of data received during work with files, registry, deletion of unnecessary settings (feature selection).
- Using machine learning methods.
- Output of results. Issuing the results of determining whether the process is harmful or not harmful.

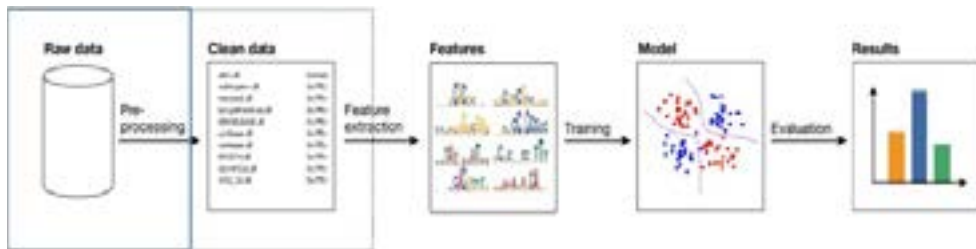


Fig. 1. Stages of detecting malicious code using machine learning

Malware can inject itself into different binary file formats that work in completely different ways. For example, PE files (extensions exe., dll., efi, i.e. portable executable files) in the Windows operating system, which we will discuss in our experiments throughout this article, have completely different internal file structures and require different execution contexts. Of course, the specific additional requirements for parsing each class of executables also differ significantly. Also note that malware can come in forms other than standalone executable binaries. There are common malicious components that infiltrate doc, pdf and rtf document files and use macros and dynamic executables in the document structure to perform malicious actions. Malware can also take the form of extensions and plug-ins for common software platforms, such as web browsers and complex web environments (Dilhara, 2021; Damin Moon et al., 2021).

The reverse engineering method is used to use the data with malicious program codes during machine learning, and since it is a large and extensive work, we preferred to work with a ready-made database. Open access malware databases:

- MalwareTrafficAnalysis.net website contains comprehensive, fully researched 600 samples of malicious code.
- VirusShare.com site 30 million provides an integrated database of malicious code patterns;

- VX Heaven team collected 270,000 malicious code samples for scientific use;
- In 2015, Kaggle and Microsoft managed to collect more than 10,000 malicious code samples into a single database by organizing the Malware Classification Challenge project.

In most cases, mindlessly loading a huge number of features into machine learning algorithms creates unnecessary noise and detrimentally affects the accuracy and efficiency of the model. Therefore, it is important to select only the most important and significant features for use in learning algorithms. This process is commonly known as feature selection. Feature selection can be done manually, based on domain expertise and information gained from the data mining phase. Features can be selected automatically using statistical methods and algorithms. In addition, there are unsupervised machine learning techniques often used for deep learning. (Malware Traffic Analysis, 2022).

One of the widely used methods of feature selection is the use of human practical experience. Human experts can provide a process for guiding machine learning models, which manifests itself primarily in the form of manually mined features, which are considered the most important pieces of information used in the human learning process. (VirusShare, 2022).

Classification of features depending on a specific model (model-specific feature ranking): when machine learning algorithms are applied to a certain set of data, the model of the final assessment can sometimes be expressed in the form of symbolic combinations of input features. (VX Heaven, 2010). For example, for a linear regression model, in which the value of Y is predicted based on a three-dimensional data set (in which the features are designated as x_a , x_b , and x_c), the regression model can be represented by the following formula (without accounting for deviations):

$$Y = W_a x_a + W_b x_b + W_c x_c \tag{1}$$

After the training phase, the coefficients (weights) W_a and W_b will be assigned some values, for example:

$$Y = 3,72x_a + 1,94x_b + 0,138x_c \tag{2}$$

Even in this simplest example, it can be clearly seen that features x_a and x_b have higher weights than feature x_c . Assuming that these features are appropriately normalized (their values are comparable values), you can eliminate the feature x_c , knowing that it will not significantly affect the performance of the regression model.

Table 1. Features

№	Feature name	№	Feature name
1	sha256	26	minor_operating_system_version name
2	appeared	27	major_subsystem_version

3	label	28	minor_subsystem_version
4	file_size	29	sizeof_code
5	vsize	30	sizeof_headers
6	has_debug	31	sizeof_heap_commit
7	exports	32	imports
8	imports	33	exports
9	has_relocations	34	entry
10	has_resources	35	name_of_section
11	has_signature	36	size_of_section
12	has_tls	37	vsize_of_section
13	symbols	38	entropy
14	header	39	props
15	timestamp	40	histogram
16	machine	41	byte_entropy
17	characteristics	42	strings
18	subsystem	43	num_strings
19	dll_characteristics	44	avlength
20	magic	45	printabledist
21	major_image_version	46	printables
22	minor_image_version	47	paths
23	major_linker_version	48	urls
24	minor_linker_version	49	registry
25	major_operating_system_version	50	MZ

For example, let's say we have the following symbols describing a file in the PE format:

- histogram of byte values – a histogram of the distribution of byte values in a binary file;

- byte entropy histogram – a two-dimensional histogram of byte entropy approximating “the combined probability of containing a value in a byte and local entropy”;

- strings: an array of string statistics extracted from the original byte stream, defined by five or more consecutive characters with ASCII values between 0x20 (space) and 0x7f (del), or special strings such as C:\, HKEY_, http:// , and contains characteristics such as number of lines, average line length, number of C:\ paths, URL instances, registry keys, and a character distribution histogram;

- some “syntactically parsed characteristics”, for example:

- general information about the file - high-level details about the PE file, such as whether it was compiled with debug symbols, number of functions exported/imported, etc.;

- file header information - details that can be obtained from the header section of a PE file related to the computer, architecture, OS, linker, etc.;

- information about sections of a binary file - names, sizes, entropy;

- import information – information about imported libraries and functions that can be used in the analyzed PE file;

- export information – information about functions exported from the analyzed PE file.

From among the above-mentioned signs, we further sort out the signs we need and prepare the data necessary for machine learning.

We will use these sets of extracted features (table 1) to classify files into malicious or benign types. To do this, we use the k-nearest neighbor algorithm.

The k-nearest neighbors (kNN) algorithm is the most widely known example of a lazy learning technique. Due to its simplicity, the kNN method is often used as a practical example to introduce machine learning concepts. This type of machine learning technique defers most of the computation to classification time instead of doing the work during training. (Kaggle and Microsoft, 2018). Lazy learning models do not learn data generalizations during the training phase. Instead, they capture (record) all the training data points that are passed to them and use this information to create local generalizations on the test set during classification. (Benjamin Bengfort et al., 2019).

The k nearest neighbors method is one of the simplest machine learning algorithms with the following characteristics: (Hyrum et al., 2018):

- the training stage consists of a simple recording (fixation) of all feature vectors and corresponding label elements in this model;
- the classification prediction is simply the most frequently occurring label among the k nearest neighbors in the test sample (according to the name of the technique).

Distance metrics for determining how “close” points are to each other in an n-dimensional feature space (where n is the size of the feature vectors) are typically the Euclidean metric (distance) for continuous variables and the Hamming distance for discrete variables. (Henrik Brink et al., 2017).

Results

The classification report is presented in Table 2.

Table 2. Classification report

	Precision	Recall	F1-score	Support
0	0.90	0.78	0.83	7458
1	0.84	0.94	0.89	9711
avg/total	0.87	0.87	0.86	17169

Accuracy is our most common evaluation metric and is easy to understand, i.e. the number of samples to be matched divided by the number of all samples. In general, the higher the accuracy, the better the classifier. The degree of accuracy is indeed a very good and intuitive measure of the estimate, but sometimes a high degree of accuracy does not reflect the algorithm. The accuracy of the model was **0.866620071058**.

Precision is for the results of our predictions and shows how many samples whose predictions are positive are correct. Then there are two possibilities to predict the positive class, one is to predict the positive class as class positive (TP) and the other is to predict the negative class as class positive (FP).

Recall is for our original sample and indicates how many positive examples in the sample are predicted correctly. There are also two possibilities, one is to predict the original positive class as class positive (TP) and the other is to predict the original positive class as class negative (FN). The recall rate is a measure of coverage.

F1-score. Indicators Precision and Recall sometimes have contradictions, so they need to be considered comprehensively. The most common method is F-Measure (also known as F-Score). F-Measure is the weighted harmonic mean of precision and recall.

Support is the number of instances of each class.

Conclusion

This article provides information on how high the level of distribution of malicious software equipment is at the present time and the importance of their detection using machine learning algorithms, as well as data obtained from practical work. The classifier model is trained on data normalized by feature importance. The trained model takes only external observation data (on the end user's machine) as input. Based on these data, the model tries to reconstruct the complete pattern of events. Model training and training is an independent study. The machine learning module is constantly trained on the results of running the suspicious software, and the model on end-user machines is updated in a continuous integration mode.

REFERENCES

- AV-TEST: The Independent IT-Security Institute, 2022. Malware. AV-TEST URL: <https://www.av-test.org/en/statistics/malware/>
- Egele M., Scholte Th., Kirda E., Kruegel C., 2012. A survey on automated dynamic malware analysis techniques and tools. *ACM Computing Surveys*, 44(2). Pp. 22–28.
- Wikipedia, 2022. Antivirus software. URL: https://en.wikipedia.org/wiki/Antivirus_software
- Michael Sikorski, Andrew Honig, 2018. *Practical Malware Analysis. The Hands-On Guide to Dissecting Malicious Software*. San Francisco, DC: No Starch Press.
- Clarence Chio, David Freeman, 2018. *Machine Learning and Security: Protecting Systems with Data and Algorithms*. Sebastopol, DC: O'Reilly.
- Ucci D., Aniello L., Baldoni R., 2017. Survey on the Usage of Machine Learning Techniques for Malware Analysis. *ACM Transactions on the Web*, 1. Pp. 3–58.
- B.A.S. Dilhara, 2021. Classification of Malware using Machine Learning and Deep learning Techniques *International Journal of Computer Applications* (0975–8887). 183–32.
- Damin Moon, JaeKoo Lee, MyungKeun Yoon, 2021. Compact feature hashing for machine learning based malware detection *Information & Communications Technology Express*, DOI: 10.1016/j.ict.2021.08.005.
- Malware Traffic Analysis, 2022. A source for packet capture (pcap) files and malware samples. URL: <https://www.malware-traffic-analysis.net/>
- VirusShare, 2022. URL: <https://virusshare.com/>
- VX Heaven, 2010. VX Heaven Virus Collection. URL: <https://web.archive.org/web/20170611163424/http://vxheaven.org/>

Kaggle and Microsoft, 2018. Microsoft Malware Prediction. URL: <https://www.kaggle.com/c/microsoft-malware-prediction>

Benjamin Bengfort, Rebecca Bilbro, Tony Ojeda, 2019. Applied Text Analysis with Python: Enabling Language-Aware Data Products with Machine Learning. Sebastopol, DC: O'Reilly.

Hyrum S. Anderson, Phil Roth, 2018. An Open Dataset for Training Static PE MalwareMachine Learning Models.

Henrik Brink, Joseph W. Richards, Mark Fetherolf, 2017. Real-World Machine Learning. Manning: Shelter, Island.

© **A.U. Altaeva, A.S. Kaipova, A.U. Mukhamejanova*, G.K. Ospanova, 2023**

«Astana Medical university» NJSC, Astana, Kazakhstan.

E-mail: sarsenova_a@mail.ru

PROSPECTS OF USING CHATBOTS IN MEDICINE

A.U. Altaeva — the senior teacher, the master of engineering science. «Astana Medical university» NJSC, Astana, Kazakhstan

E-mail: altaeva.a.u@mail.ru;

A.S. Kaipova — the senior teacher, the master of engineering science. «Astana Medical university» NJSC, Astana, Kazakhstan

E-mail: aru.as@mail.ru;

A.U. Mukhamejanova — the senior teacher, Master of applied mathematics and computer science. «Astana Medical university» NJSC, Astana, Kazakhstan

E-mail: sarsenova_a@mail.ru;

G.K. Ospanova — the senior teacher, the master of engineering science. «Astana Medical university» NJSC, Astana, Kazakhstan

E-mail: galiya_astana2011@mail.ru.

Abstract. Chatbots are applications that can conduct conversations with users in natural language. In the medical field, chatbots have been developed and used for various purposes. They provide patients with timely information that can be crucial in some scenarios. This study was aimed at developing a chatbot in the medical field. The chatbot is configured to convert natural language queries into corresponding SQL queries. SQL queries are executed using the knowledge base, and the results are returned to the user in a natural dialog box. The chatbot was designed to solve common template queries. Methods of knowledge output were used to generate responses to queries. As a result of the development of the chatbot, an approbation work was carried out to test the service.

Keywords: chatbot, medicine, Python, variables, Framework, telegram

© А.У. Алтаева, А.Ш. Каипова, А.У. Мухамеджанова*,
Г.К. Оспанова, 2023

НАО Медицинский университет Астана, Астана, Казахстан.

E-mail: sarsenova_a@mail.ru

МЕДИЦИНАДА ЧАТ-БОТТАРДЫ ҚОЛДАНУ ПЕРСПЕКТИВАЛАРЫ

Аннотация. Чат-боттар-бұл қолданушылармен табиғи тілде сөйлесе алатын қосымшалар. Медициналық салада чат-боттар ауруды алдын-алу, психикалық жағдайды анықтау сияқты әртүрлі мақсаттарда жасалынған. Олар пациенттерге кейбір сценарийлерде маңызды болуы мүмкін болатындай уақтылы ақпарат береді. Бұл мақалада Python бағдарламасының көмегімен медицина саласындағы чат-ботты әзірлеу қарастырылған. Чат-бот табиғи тілдегі сұраныстарды тиісті SQL сұрауларына түрлендіру үшін конфигурацияланған. SQL сұраулары білім базасын қолдана отырып жасалады және нәтижелер табиғи тіл қатысу терезесінде қолданушыға қайтарылады. Чат-бот жалпы шаблондық сұраныстарды шешуге арналған. Сұрауларға жауап беру үшін білімді алу әдістері қолданылды. Чат-ботты әзірлеу нәтижесінде сервисті тестілеу үшін апробациялық жұмыстар тізбегі жүргізілді.

Түйін сөздер: чат-бот, медицина, Python, айнымалы, Фреймворк, телеграмм

© А.У. Алтаева, А.Ш. Каипова, А.У. Мухамеджанова*,
Г.К. Оспанова, 2023

НАО Медицинский университет Астана, Астана, Казахстан.

E-mail: sarsenova_a@mail.ru

ПЕРСПЕКТИВЫ ИСПОЛЬЗОВАНИЯ ЧАТ-БОТОВ В МЕДИЦИНЕ

Аннотация. Чат-боты — это приложения, которые могут вести беседы с пользователями на естественном языке. В медицинской области чат-боты были разработаны и использовались для различных целей. Они предоставляют пациентам своевременную информацию, которая может иметь решающее значение в некоторых сценариях. Данное исследование было направлено на разработку чат-бота в медицинской сфере. Чат-бот настроен для преобразования запросов на естественном языке в соответствующие SQL-запросы. SQL-запросы выполняются с использованием базы знаний, и результаты возвращаются пользователю в естественном диалоговом окне. Чат-бот был разработан для решения распространенных шаблонных запросов. Методы вывода знаний были применены для генерации ответов на запросы. В результате разработки чат-бота была проведена апробационная работа для тестирования сервиса.

Ключевые слова: чат-бот, медицина, Python, переменные, Фреймворк, телеграмм

Введение

Чат-боты — это системы, способные общаться с пользователями на естественном языке таким образом, который имитирует взаимодействие с реальным человеком. Разработка чат-ботов давно привлекала внимание исследователей. Eliza была одной из первых попыток разработки разговорной системы. В 1966 году профессор Массачусетского технологического института Джозеф Вейценбаум разработал компьютерную программу под названием Элиза. Считается, что это первый чат-бот в истории. Элиза была простым чат-ботом на основе ключевых слов, имитирующим человека-психиатра. Программа общалась, сопоставляя вопросы пользователей со сценариями ответов, введенными в ее базу данных. С тех пор разработка чат-ботов стала неотъемлемой частью многих областей применения. Наиболее ярким примером является использование чат-ботов в качестве личных помощников, таких как Apple Siri и Google Assistant. Чат-боты также были разработаны и используются в различных областях применения, таких как маркетинг и предоставление различных видов услуг.

С первых лет своего развития люди применяли различные методы проектирования и разработки чат-ботов. В последние годы, с растущим внедрением методов машинного обучения и искусственного интеллекта в различных областях применения, современные методы создания диалогов и управления диалогами все чаще используют методы машинного обучения и глубокого обучения.

Медицина — это область применения, где чат-боты все чаще используются в качестве инструмента для облегчения доступа к информации со стороны пациента и снижения нагрузки на врача. Многие коммерческие приложения для чат-ботов, доступные в виде веб-приложений или мобильных приложений, были разработаны для взаимодействия с пациентами. Некоторыми примерами чат-ботов для здравоохранения являются OneRemission1, который был разработан, чтобы помочь выжившим после рака; Babylon Health, программа для проверки симптомов; и Wusa, чат-бот для психического здоровья, который взаимодействует с пользователем, чтобы помочь с признаками беспокойства и депрессии. Поэтому целью этой работы является улучшение качества обслуживания клиентов путем разработки чат-бота в сфере медицины.

Материалы и основные методы

По многим причинам важно знать текущее состояние различных методов и техник, используемых при разработке чат-ботов в медицинской сфере. Проведение такого опроса поможет исследователям в будущем определить различные методы, которые использовались, и использовать существующие подходы для разработки более интеллектуальных чат-ботов, которые

обеспечивают более естественный опыт для пользователя. Также важно видеть, каково текущее состояние разработки чат-ботов по отношению к разработке чат-ботов для других приложений. Поэтому в этой работе мы провели предварительный обзор доступной литературы по разработке чат-ботов в области медицины, построили и определили основные компоненты, участвующие в разработке чат-ботов, а также описание методов, используемых при разработке каждого из выявленных компонентов.

Как правило, чат-боты состоят из 7 компонентов и имеют следующую структуру (Рисунок 1):

Обработка естественного языка. Обработка естественного языка (NLP) позволяет чат-ботам преобразовывать текст и речь пользователей в структурированные данные, понятные машине. NLP состоит из следующих шагов:

Токенизация: также называемая лексическим анализом, представляет собой процесс разделения строки слов, образующих предложение, на более мелкие части «токены» на основе ее значения и ее отношения ко всему предложению.

Нормализация: также называемая синтаксическим анализом, представляет собой процесс проверки слов на наличие опечаток и приведения их к стандартной форме. Например, слово «tmrw» будет преобразовано в «завтра».

Распознавание объектов: процесс поиска ключевых слов для определения темы разговора.

Семантический анализ: процесс определения значения предложения путем понимания значения каждого слова и его отношения к общей структуре.



Рисунок 1. Структура чат бота.

Понимание естественного языка. Понимание естественного языка (NLU) – это подраздел НЛП, который фокусируется на понимании смысла человеческой речи путем распознавания закономерностей в неструктурированном речевом вводе. Решения NLU состоят из 3 компонентов:

- Словарь для определения значения слова;
- Парсер, чтобы определить, соответствует ли синтаксис текста правилам языка;

-Грамматические правила для разбивки ввода на основе структуры предложения и пунктуации.

NLU позволяет чат-ботам классифицировать намерения пользователей и генерировать ответ на основе обучающих данных.

Баззнаний.Базазнаний—этобиблиотекаинформации,накоторуюопирается чат-бот для получения данных, используемых для ответа пользователям. Базы знаний различаются в зависимости от потребностей бизнеса. Например, база знаний чат-бота для веб-сайта электронной коммерции будет содержать информацию о продуктах, функциях и ценах, тогда как база знаний чат-бота для здравоохранения будет содержать информацию о календарях врачей, часах работы больниц и обязанностях аптек. Кроме того, некоторые чат-боты интегрированы с веб-скраперами для извлечения данных из онлайн-ресурсов и отображения их пользователям.

Хранилище данных. Разработчики чат-ботов могут сохранять разговоры для обслуживания клиентов, а также для обучения и тестирования ботов. Разговоры чат-бота могут храниться в форме SQL локально или в облаке.

Диспетчер диалогов. Диспетчер диалогов — это компонент, отвечающий за ход разговора между пользователем и чат-ботом. Он ведет учет взаимодействий в рамках одного разговора, чтобы решить, как реагировать. Например, если пользователь говорит «Я хочу заказать клубничное мороженое», а затем в разговоре говорит «Изменить мой заказ на шоколадное мороженое», диалоговый менеджер позволит боту обнаружить изменение «клубничного» на «шоколадное».

Генерация естественного языка. Генерация естественного языка (NLG) — это процесс преобразования машинных структурированных данных в удобочитаемый текст. После понимания намерений пользователей у NLG есть 6 шага для создания ответа:

Определение содержания: фильтрация существующих данных в базе знаний, чтобы выбрать, что включить в ответ.

Интерпретация данных: понимание шаблонов и ответов, доступных в базе знаний.

Планирование документа: структурирование ответа в повествовательной манере.

Агрегация предложений: составление выражений и слов для каждого предложения в ответе.

Грамматизация: применение правил грамматики, таких как пунктуация и проверка орфографии.

Реализации языка: ввод данных в языковые шаблоны для обеспечения естественного представления ответа.

Пользовательские интерфейсы. Диалоговые пользовательские интерфейсы — это внешний интерфейс чат-бота, который обеспечивает физическое представление разговора. Они делятся на текстовые или голосовые

помощники. И их можно интегрировать в разные платформы, такие как Facebook Messenger, WhatsApp, Slack, Google Teams и т. д.

Для создания чат-бота прежде всего устанавливаем фреймворк Telegram Bot. Для его реализации используем командную строку:

```
$ pip install python-telegram-bot
```

Для выявления каких-либо ошибок в процессе работы устанавливаем фреймворк телеграм регистрации Python(Рисунок 2).

```
from telegram.ext import Updater, CommandHandler, MessageHandler, Filters, Dispatcher
import logging
logging.basicConfig(format='%(levelname)s - %(message)s',
                    level=logging.DEBUG)
logger = logging.getLogger(__name__)
updater = None
def start_bot():
    global updater
    updater = Updater(
        '### SAG KUM APTEHINAKAM BOCA TELEGRAM KAVXINTON ZDICH ###', use_context=True)
    updater.start_polling()
    updater.idle()
start_bot()
```

Рисунок 2- Установка фреймворков.

В коде переменная Updater является глобальным переменным. Он обеспечивает легкий передний конец для работы с ботом и запустить наш новый бот с помощью Updater. Переменная UPDATATER.START_POLLING () запускает бот и начинает опрос для любых обновлений чата по телеграмме. Бот начнет опрос в своих отдельных потоках, чтобы не остановить сценарий Python.

Комана UPDATER.IDLE () используется для того, чтобы заблокировать скрипт, пока пользователь не отправит команду, чтобы вырваться из скрипта Python.

При запуске этого сценария вы сможете увидеть сообщения, которые запустили бот и обновления, а также запущено ряд потоков(Рисунок 3).

```
@app.route("/").format(TOKEN), methods=['POST'])
def respond():
    update = telegram.Update.de_json(request.get_json(force=True), bot)
    chat_id = update.message.chat_id
    msg_id = update.message.message_id
    text = update.message.text.encode('utf-8').decode()
    print("got text message:", text)
    if text == "/start":
        bot_welcome = """
        Добро пожаловать в наш новый чат-бот "Саймон"
        """
        bot.sendMessage(chat_id=chat_id, text=bot_welcome, reply_to_message_id=msg_id)
    else:
        url = re.sub("http", "https", text)
        bot.sendPhoto(chat_id=chat_id, photo=url, reply_to_message_id=msg_id)
    except Exception:
        bot.sendMessage(chat_id=chat_id, text="Ошибка xxx", reply_to_message_id=msg_id)
    return 'ok'
```

Рисунок 3 - Запуск программы

Результаты

Интуитивно понятный способ заставить эту функцию работать состоит в том, что мы будем вызывать ее каждую секунду, чтобы она проверяла, пришло ли новое сообщение(Рисунок 4).



Рисунок 4 – Первая страница реализации чат-бота.

Система собирает симптомы, историю болезни и подготавливает данные для врача. Таким образом, во время приема доктор может сосредоточиться на диагностике и лечении. Вся информация о пациенте собрана заранее.

Обсуждение

Для проверки работоспособности чат-бота была проведена тестовая работа. В общей сложности 87 участников приняли участие. Возрастной диапазон 18–55 года, средний возраст 36 лет. По итогам тестовой работы, был проведен опрос(Таблица 1).

Таблица 1

Результат анкетирования

№	Название вопроса	Да(%)	Нет (%)
1	Удобно ли использовать чат бот?	86	14
2	Преимущества: экономия времени	70	30
3	Преимущества: доступность	100	0
4	Преимущества: эффективное вербование информации	100	0
5	Эффективная обработка данных	85	15

Как мы видим в таблице, 86 % респондентам удобно использовать сервис, а также 70 % считают, что можно сократить время на консультацию, так как

сбор информации занимает 60 % времени консультации и длится до 15 минут. Алгоритмы сокращают время сбора в несколько раз и позволяют снизить вероятность ошибки. Доступность сервиса и эффективное вербование информации составляет 100 %. Таким образом, наш чат-бот:

Анализирует жалобы пациента и уточняет детали симптомов;

Спрашивает про наиболее вероятные и важные симптомы и их уточнения;

Структурированно собирает анамнез жизни и другие факторы;

3.1 Структурированная информация по анамнезу предоставляется врачу;

В зависимости от поставленной задачи выдает результат.

Заключение

И так в заключении, цель нашего чат-бота состоит в:

Сокращении времени, которое пациент и доктор тратят на консультацию. Сервис собирает анамнез за 1–3 минуты, что намного быстрее, чем отвечать на аналогичные вопросы врачу.

Повышения качества. Система задает множество вопросов, агрегируя большой объем информации и снижая риск того, что доктор что-то пропустит.

Структурированный сбор данных о жалобах и симптомах. Если раньше они собирались в текстовом виде в свободной форме, то благодаря боту они собираются структурированно и имеют контекст, так как по большинству симптомов пациент отвечает на дополнительные вопросы.

REFERENCES

- Weizenbaum J., 1966. ELIZA - A Computer Program For The Study of Natural Language Communications Between Man and Machine. Communications of the ACM, 1966. 9(1). Pp. 36–45.
- Shawar B.A. and Atwell E., 2007. Chatbots: are they really useful? Journal of Computational Linguistics and Language Technology. Vol. 22. № 1. Pp.29–49.
- Gnewuch U., Morana S., Maedche A., 2017. Towards Designing Cooperative and Social Conversational Agents for Customer Service. 2017 Presented at: International Conference on Information Systems (ICIS) 2017; December 10–13, 2017; Seoul, South Korea.
- Wallace R.S., 2010. Alicebot. Retrieved June 2. 2010 from the World Wide Web: <http://www.alicebot.org/aimlbots.html>
- Guo J., Lu S., Cai H., Zhang W., Yu Y., Wang J., 2018. Long Text Generation via Adversarial Training with Leaked Information. 2018 Presented at: The Thirty-Second AAAI Conference on Artificial Intelligence (AAAI-18); February 2–7, 2018; New Orleans, LA.
- Aivazov A.L., 2019. Ispolzovanie vozmojnostei chatbota kak raznovidnosti iskusstvennogo intellekta v deiatelnosti HR-spezialistov / A.L. Aivazov, T.A. Kornienko, M.S. Gorovaia // Aleia nauki. - 2019. T. 1. - № 2 (28). - Pp. 223–231.
- Dzhoshi P., 2019. Iskusstvennyi intellekt s primerami na Python: sozdanie prilozhenii iskusstvennogo intellekta s pomoshchi Python dlya vzaimodeistvia s okrujaiushim mirom / Pratik Djoshi. - Moskva: Sankt-Peterburg: Dialektika, 2019. - 444 p.
- Kucherbaev P., 2019. Cheloveko-mashinnye chat-boty / P. Kucherbaev, A. Bozon, G.-J. Hyben // Otkrytye sistemy. SUBD. – 2019. - № 1
- Luchano Ramalo Python. K vershinam masterstva. – M.: DMK Pres, 2016. – 768 p.
- Shampandar Aleks Dj., 2007. Iskusstvennyi intellekt v komputernyh igrakh: kak obuchit virtualnye personaji reagirovat na vneshnie vozdeistvia. -M.: OOO "ID. Vilyams", 2007. — 768 p.

NEWS OF THE NATIONAL ACADEMY OF SCIENCES OF THE REPUBLIC OF
KAZAKHSTAN
PHYSICO-MATHEMATICAL SERIES
ISSN 1991-346X
Volume 2. Number 346 (2023). 40–54
<https://doi.org/10.32014/2023.2518-1726.183>

UDC 004; 004.048

© **G.A. Anarbekova, N.N. Ospanova***, **D.Zh. Anarbekov, 2023**
S. Seifullin Kazakh Agrotechnical Research University, Astana, Kazakhstan.
E-mail: akerkegansaj@mail.ru

NORMALIZED INPUT VECTORS: THE PRIMARY STAGE OF DATA PREPARATION

Anarbekova G.A. — doctoral student, department of Computer Systems and Vocational Trai. S.Seifullin KATRU . 010000. Astana, Kazakhstan

E-mail: galiya2810@mail.ru. ORCID: 0000-0001-8960-8036;

Ospanova N.N. — c.p.s, facultyComputer science факультеті. Toraigurov University. 140000 Pavlodar, Kazakhstan

E-mail: nazira_n@mail.ru. ORCID: 0000-0003-0100-1008;

Anarbekov D.Zh. — Senior Lecturer, department of Computer Systems and Vocational Trai. S.Seifullin KATRU . 010000. Astana, Kazakhstan

E-mail: anarbekovd@gmail.com. ORCID: 0000-0002-9330-573X.

Abstract. The protection of agricultural crops from pests in Northern Kazakhstan, such as the striped bread flea (*Phyllotreta vittula*) occupies an important place in crop production, allowing not only to preserve the harvest, but also to strengthen food security. Early detection of the pest to find solutions remains a challenge in the field of sustainable agriculture. The most promising way to protect plants from pests is the use of new technologies, such as machine learning, neural network. We investigate the factors for the identification of the striped bread flea, in order to obtain versatile information: the collection and processing of data characterizing the species composition, distribution and development of the pest is organized, we collect data for the last ten years in Northern Kazakhstan. Machine learning methods will use numerical data to identify the pest of plants. The use of machine learning in agriculture makes it possible to increase the efficiency and accuracy of farming with less labor and high quality products. The next stage of the research is to normalize all the collected data using data normalization in the Python program and further use them in neural network training. The normalization process enters the initial stage, after data collection during forecasting in machine learning.

Keywords: data normalization, machine learning, forecasting, agriculture, bread striped flea

Financing: The research was conducted from the authors' personal funds.

Conflict of interest: The authors declare no conflict of interest.

© Г.А. Анарбекова, Н.Н. Оспанова*, Д.Ж. Анарбеков, 2023
С. Сейфуллин атындағы Қазақ агротехникалық зерттеу университеті,
Астана, Қазақстан.
E-mail: akerkegansaj@mail.ru

НОРМАЛАНҒАН КІРІС ВЕКТОРЛАРЫ: ДЕРЕКТЕРДІ ДАЙЫНДАУДЫҢ БАСТАПҚЫ КЕЗЕҢІ

Анарбекова Г.А. — докторант, Компьютерлік жүйелер және кәсіптік білім беру факультеті.

С. Сейфуллин атындағы ҚАТЗУ. 010000. Астана, Қазақстан

E-mail: galiya2810@mail.ru. ORCID: 0000-0001-8960-8036;

Оспанова Н.Н. — к.п.н, Computer science факультеті. Торайғыров Университеті. 140000
Павлодар, Қазақстан

E-mail: nazira_n@mail.ru. ORCID: 0000-0003-0100-1008;

Анарбеков Д.Ж. — аға оқытушы, Компьютерлік жүйелер және кәсіптік білім беру факультеті.

С.Сейфуллин атындағы ҚАТЗУ. 010000. Астана, Қазақстан

E-mail: anarbekovd@gmail.com. ORCID: 0000-0002-9330-573X.

Аңдатпа. Солтүстік Қазақстандағы ауыл шаруашылығы дақылдарынан жолақты бүрге (*Phyllotreta vittula*) сияқты зиянкестерден қорғау өсімдік шаруашылығында маңызды орын алады, бұл егінді сақтап қана қоймай, азық-түлік қауіпсіздігін нығайтуға мүмкіндік береді. Зиянкестерді ерте анықтау ауыл шаруашылығы саласында шешімдерді табу үшін әлі де проблема болып табылады. Өсімдіктерді зиянкестерден қорғаудың ең перспективалы жолы машиналық оқыту, нейрондық желі сияқты жаңа технологияларды пайдалану. Нанның жолақты бүргесін анықтау факторларын зерттейміз, жан-жақты ақпарат алу үшін: зиянкестердің түрлік құрамын, таралуы мен дамуын сипаттайтын деректерді жинау және өңдеу ұйымдастырылады, Солтүстік Қазақстанда соңғы он жылдағы деректерді жинаймыз. Машиналық оқыту әдістері өсімдік зиянкестерін анықтауға арналған сандық деректерді пайдаланады. Ауыл шаруашылығында машиналық оқытуды қолдану еңбек шығындарын азайтып және өнім сапасын арттыруға мүмкіндік береді. Зерттеудің келесі кезеңі-Python бағдарламасындағы деректерді қалыпқа келтіруді қолдана отырып, барлық жиналған деректерді қалыпқа келтіру және оны нейрондық желіні оқытуда одан әрі қолдану. Қалыпқа келтіру процесі машиналық оқытуда болжау кезінде деректерді жинағаннан кейін бастапқы кезеңге енеді.

Түйін сөздер: деректерді қалыпқа келтіру, машиналық оқыту, болжау, ауыл шаруашылығы, жолақты нан бүргесі

Қаржыландыру: Зерттеу авторлардың жеке қаржысынан жүргізілді.

Мүдделер қақтығысы: Авторлар осы мақалада мүдделер қақтығысы жоқ деп мәлімдейді.

© Г.А. Анарбекова, Н.Н. Оспанова*, Д.Ж. Анарбеков, 2023

Казахский агротехнический исследовательский университет

им. С. Сейфуллина, Астана, Казахстан.

E-mail: akerkegansaj@mail.ru

НОРМАЛИЗОВАННЫЕ ВХОДНЫЕ ВЕКТОРЫ: ПЕРВИЧНЫЙ ЭТАП ПОДГОТОВКИ ДАННЫХ

Анарбекова Г.А. — докторант, факультет Компьютерные системы и профессиональное образование. КАТИУ имени С.Сейфуллина. 010000. Астана, Казахстан

E-mail: galiya2810@mail.ru. ORCID: 0000-0001-8960-8036;

Оспанова Н.Н. — к.п.н, Computer science факультеті. Торайгыров Университеті. 140000. Павлодар, Қазақстан

E-mail: nazira_n@mail.ru. ORCID: 0000-0003-0100-1008;

Анарбеков Д.Ж. — старший преподаватель, факультет Компьютерные системы и профессиональное образование. КАТИУ имени С.Сейфуллина. 010000. Астана, Казахстан

E-mail: anarbekovd@gmail.com. ORCID: 0000-0002-9330-573X.

Аннотация. Защита сельскохозяйственных культур от вредителей в Северном Казахстане, таких как хлебная полосатая блошка (*Phyllotreta vittula*) занимает важнейшее место в растениеводстве, позволяя не только сохранить урожай, но и укрепить продовольственную безопасность. Раннее выявление вредителя для поиска решений по-прежнему остается проблемой в области устойчивого сельского хозяйства. Наиболее перспективным способом защиты растений от вредителя является использование новых технологий, таких как машинное обучение, нейронная сеть. Исследуем факторы по выявлению хлебной полосатой блошки, с целью получения разносторонней информации: организуется сбор и обработка данных, характеризующих видовой состав, распространение и развитие вредителя, собираем данные за последние десять лет в Северном Казахстане. Методы машинного обучения будут использовать числовые данные по идентификации вредителя растений. Применение машинного обучения в сельском хозяйстве позволяет повысить эффективность и точность ведения сельского хозяйства с меньшими затратами труда и высоким качеством продукции. Следующим этапом исследования является нормализовать все собранные данные, используя нормализацию данных в программе Python и в дальнейшем применение их в обучении нейронной сети. Процесс нормализации входит в начальный этап, после сбора данных при прогнозировании в машинном обучении.

Ключевые слова: нормализация данных, машинное обучение, прогнозирование, сельское хозяйство, хлебная полосатая блошка

Финансирование: Исследование проводилось из личных средств авторов.

Конфликт интересов: Авторы заявляют об отсутствии конфликта интересов.

Introduction

Agricultural productivity makes a significant contribution to the economy of Kazakhstan. Pests of grain crops seriously affect the main production and cause great damage to the economy. One of such permanent pests of grain crops is the bread striped flea.

Health monitoring and early diagnosis of pests of grain crops is the most important task of sustainable agriculture. Information about the early diagnosis of certain plant diseases can facilitate pest control by choosing the right methods of combating them to increase grain yield. Manual identification of disturbances in grain crops can lead to inaccurate measurements of pesticides.

This work is carried out manually by agronomists, the above problem requires the intervention of the latest technologies for predicting the appearance of plant pests. Therefore, a method for recognizing pests of agricultural crops based on a modified capsule network has been proposed in world science (Zhang, 2022). The modified capsule network is used to improve the traditional convolutional neural network, and the attention module is introduced to capture the most important classification features and accelerate network learning. The results of experiments with a set of images of pests confirm that the proposed method is effective and feasible in classifying various insect species in field crops and can be applied in the agricultural sector for plant protection.

Currently, machine learning is very widely used in agronomic research. The use of machine learning in agricultural data processing makes it possible to increase the efficiency and accuracy of farming with less labor and high product quality. Machine learning has emerged together with big data technologies and high-performance computing to create new opportunities for analyzing, quantifying and understanding processes that require large amounts of data in agricultural operating environments.

A comprehensive review of the research of foreign scientists on the application of machine learning in agricultural production systems is presented. The filtering and classification of the submitted articles demonstrate how agriculture will benefit from machine learning technologies. By applying machine learning to sensor data, farm management systems are transformed into real-time AI-enabled programs that provide detailed recommendations and analytical information to support farmers' decision-making and actions (Liakos, 2018).

Modern technologies, such as the Internet of Things (IoT), lay the foundation for precision farming, which minimizes human labor and costs, as well as increases agricultural productivity. The Internet of Things generates large amounts of data that can be used for practices such as crop monitoring or disease detection. The analysis and interpretation of these data make it possible to understand the relationships between various agricultural factors, such as soil characteristics and climatic variables. This contributes to timely and informed decision-making and planning. Machine learning plays a central role in these decision support systems by modeling complex patterns that may exist in the data (Condran, 2022).

The data generated in modern agricultural operations is provided by a variety of different sensors that allow a better understanding of the working environment (interaction of dynamic crops, soil, and weather conditions) and the operation itself (equipment data), which leads to more accurate and faster decision-making.

Tasks are usually divided into various broad categories depending on the type of training (supervised/unsupervised), learning models (classification, regression, clustering and dimensionality reduction) or learning models used to implement the selected task.

In general, the effectiveness of a machine learning solution depends on the nature and characteristics of the data and the performance of the learning algorithms. In the field of machine learning algorithms, classification analysis, regression, data clustering, foreign economic activity - for the effective construction of data-driven systems, there are methods of structure design and dimensionality reduction, the study of associative rules or reinforcement learning (Sarker, 2021).

Machine learning, including neural networks, prompted the processing and analysis of big data obtained in the research of the agricultural sector. To carry out machine learning, you need to perform a number of actions that include: preparation of reliable data, normalization of data suitable for machine learning algorithms, data transformation, method selection (neural network architecture), re-sampling, building a neural network model and training a neural network. The main and initial stage in data training is the normalization of the data itself.

Data normalization is a fundamental preprocessing step for data mining and data-based learning. However, finding a suitable method for normalizing time series is not an easy task. This is due to the fact that most traditional normalization methods make assumptions that are not valid for most time series. The first assumption is that all time series are stationary, i.e. their statistical properties, such as mean and standard deviation, do not change over time. The second assumption is that the time series is considered uniform. None of the methods currently available in the literature solves these problems. Therefore, a new method of normalization of time series is proposed in world science. The method, called Adaptive Normalization, was tested together with an artificial neural network in three prediction tasks. The results were compared with the other four traditional normalization methods and showed an increase in accuracy in both short- and long-term forecasts (Ogasawara, 2010).

The use of satellite images for monitoring drought in agriculture, both regionally and globally, is attracting increasing attention of researchers. Since agricultural drought is one of the most destructive agricultural threats worldwide, which can lead to significant losses in agriculture and water shortages. Foreign scientists first analyzed the correlation between the temperature of the earth's surface and the normalized difference vegetation index using time series under different vegetation growth conditions, the data obtained showed that the index of the temperature regime of vegetation can be used only in the warm season (late spring and summer

periods), when negative correlations are observed between the temperature of the earth's surface and the normalized difference vegetation index (Hu, 2019).

Thanks to the capabilities of nonlinear modeling, prediction networks with deep learning have become widely used for intelligent agriculture. Since sensor data contains noise and complex nonlinearity, and it was necessary to increase their performance, scientists proposed a reversible network for automatic normalization of selection. Integrating the level of normalization and renormalization for evaluation and selection of the normalization module of the forecasting model. Prediction accuracy has been effectively improved by scaling and transforming input data using trainable parameters. The results of the forecasting application showed that the model has a good forecasting ability and adaptability for greenhouses in the system of intelligent agriculture (Shi, 2022).

Plant diseases are one of the problems that can lead to losses in the production and economy of the agricultural sector. Early detection of this disease to find solutions and treatments remains a challenge in the field of sustainable agriculture. Currently, image processing methods and machine learning methods are used to successfully detect plant diseases. In order to increase efficiency in the multiclass classification of plant diseases, automatic diagnostics of leaf diseases is proposed. This is the development of a system for a convolutional neural network with a deep learning approach based on batch normalization for the classification of plant diseases. The importance of using deep learning technology is to make the system end-to-end, automatic, accurate, less expensive, and more convenient to detect plant diseases by their leaves (Fahad, 2021).

In the research conducted by, a total of 50 indicators from the categories of productivity, stability, efficiency, durability, compatibility and equity in agriculture are used to find out which normalization method is most suitable for further mathematical analysis to develop a final composite indicator. In order to understand the consistency and quality of normalization measurement methods and compare the advantages and disadvantages of various selected normalization processes, indicators of agricultural sustainability were considered (Talukder, 2017). Each of the different normalization methods had its advantages and disadvantages. This research showed that the rules of proportional normalization and hybrid aggregation of the arithmetic mean and geometric mean are suitable for the selected data set and that this method has wider application for the development of composite indicators for assessing the sustainability of agriculture.

In machine learning tasks, the source data is often set in different units of measurement and types of scales. Such data should be transformed into a single representation by normalizing or standardizing them. The paper shows the difference between these operations. The main types of scales, operations on the data presented in these scales, and the main options for normalization of functions are systematized. The rules for separating the features of tree classifiers are invariant to the scales of quantitative features. They only use the comparison operation.

Perhaps due to this property, the classifier of the "random forest" type as a result of numerous experiments is recognized as one of the best when analyzing data of different nature (Starovoitov, 2021).

Thus, it can be seen from the studied scientific works of world scientists that the following methods have been studied and proposed: the influence of adaptive normalization on the prediction of time series using artificial neural networks; the use of satellite images to monitor drought in agriculture; reversible automatic normalization network for greenhouses in the intelligent agriculture system; development of a system based on batch normalization for the classification of plant diseases; rules of proportional normalization and hybrid aggregation for assessing the sustainability of agriculture; a classifier of the "random forest" type transformed into a single representation by normalizing them. But the studied scientific sources do not provide complete information about the factors and normalization of input data on agricultural pests, including the striped bread flea, for training in a neural network. Analysis of literature sources has shown that the normalization of data regarding the processing of data on pests in the agricultural sector, such as the bread-striped flea, remains poorly studied. This area requires a more thorough analysis, taking into account statistical data on factors affecting the growth of agricultural pests, in particular the bread striped flea, and obtaining normalized data for further training in the neural network.

The purpose of this research is to determine the factors and indicators for the normalization of input data used in predicting the growth dynamics of the main pest based on an analytical system to improve yields. To achieve the goal, the following tasks were set: a) determination of factors for the identification of plant pests (numerical data of which will be data for machine learning) and data collection b) normalization of the collected data for further use in neural network training. The normalization process enters the initial stage, after data collection during forecasting in machine learning and affects the learning outcome. Therefore, it is important to consider this stage more carefully.

Methods and materials

Data normalization is an important step in data preparation before training the neural network. The result of training depends on correctly normalized data. And also, before data normalization, by analyzing previously researched data on pest population growth, factors were selected that reflect the dependence of pest population growth.

There are many ways to normalize and scale feature values to a common range for use in various machine learning models. They can be divided into two large groups: linear and nonlinear, depending on the function used. Nonlinear normalization uses logistic sigmoids or hyperbolic tangent functions with calculated ratios. With linear normalization, changes in variables are carried out proportionally according to a linear law.

The need to normalize data sampling is the very nature of the variables used

in neural network models. Because they differ in a physical sense, their absolute values are often very different from each other. For example, a sample may contain concentrations measured in tenths or hundredths of a percent, as well as a pressure of hundreds of thousands of pascals. Data normalization allows you to move all numeric values of the variables used to the same change area.

To perform data normalization, it is necessary to know exactly the limits of change (theoretically possible minimum and maximum values) of the values of the corresponding variables. Then they will correspond to the boundaries of the normalization interval. If the limits of variable variation cannot be set precisely, they are set taking into account the minimum and maximum available data samples.

Forecasting methods in machine learning, including: review and analysis of the state of the number of pests of grain crops in Kazakhstan; analysis of existing and promising methods of accounting and forecasting the current state of crops. The general methodology of the research was the application of a systematic approach. The following methods were used to solve the problems: analysis, scientific generalization, comparison, experimental research. During the analysis, a comprehensive analysis of the state of forecasting the yield of grain crops for the n th number of years. With the help of scientific generalization, the results of existing experiments in the field of research will be reviewed and summarized.

To obtain normalized data, first of all, factors for identifying plant pests were determined (numerical indicators of pest characteristics, the data of which will be data for machine learning) and data collection: exit from wintering sites on grasses, abundance, migration to crops, damage to the leaf surface, mating, hatching of larvae, the emergence of a new generation, going to wintering grounds.

The bread striped flea is a permanent pest of grain crops. Under the condition of a warm and humid spring, and also taking into account the accumulated stock of the population due to successful overwintering, a widespread increase in the pest population is expected. Its harmfulness is largely manifested in dry, hot weather in the initial period of growth and development. At this time, the first leaf suffers most from the bread flea. Young plants are noticeably depressed, turn yellow and dry up. Early crops are damaged more than later ones. Another factor contributing to the development of the pest is the severe damage of plants by root rot. To a greater extent, barley and spring wheat are damaged, to a lesser extent – oats, corn. The harm caused by the phytophagous can be reduced by using insecticidal mordants, observing agricultural cultivation techniques, and, if necessary (due to their peculiarities of populating fields) – carrying out local edge treatments (band width – 50-100 m) with insecticides.

For the future, the forecast of the number of growth of the bread striped flea is associated with the phenophase of culture development suitable for the insect (1–2 leaves), as well as with the weather conditions of the year. The authors have revealed a direct dependence of the bread flea on temperature and the reverse on precipitation in the year of sowing (Kozulina, 2021).

Key aspects of the development include knowledge of system dynamics based on data from several seasons. An alternative approach was to construct models parameterized on the basis of independent controlled experiments aimed at identifying the response of organisms to a number of environmental factors. The two most popular examples are the phenological models of insect pests and infection models (Welch, 1978, Magarey, 2005). Such models can be used to determine how climate change may affect the frequency of pesticide use. In some cases, it is possible to assess the impact on the crop by converting pest intensity forecasts into crop loss forecasts.

The development and spread of the pest of agricultural crops are influenced by the weather and climatic conditions of the growing season. Accounting of all data was carried out on warm windless days in the ear phases from germination to tillering, and for normalization, data obtained only on windless days were taken. Thus, objective data were obtained, which are presented in Table 1 since 2011.

Table 1. Factors influencing the number of development of the bread striped flea (Northern Kazakhstan)

North.Kaz.	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021
Exit from wintering grounds on grasses	14 April-5 May	10-21 April	16-22 April	21-25 April	20-29 April	11-19 April	15-20 April	20 April-5 May	10-25 April	15 April-5 May	17-23 April
Number (ex/m ²)	1-50	0,5-18	0,4-1,0	1-80	0,5-70	1-60	0,5-50	1,0-50	10-38,3	1,0-40,0	1,0-50
Migration to crops	I decade of June	May	28-31 May	24-28 May	25-29 May to 2-4 June	15-20 May	25-28 May	2-4 June	24-26 May	20 May	29 May
Number (ex/m ²)	1,0-300	2,0-300	2,0-300	2,0-25	1,0-150	1,0-300	1,0-150	2,0-300	1,0-100	1,0-100	2,0-200
D a m a g e to the leaf surface (%)	6-22	0,3-22,0	2-8 to 20-25	0,1-5,0	0,5-25,0	5-23	0,5-25,0	0,4-22,0	3-25,0	1,0-50,0	6-22
Mating	5-18 June	26 May-6 June	5-21 June	9-16 June	12-19 June	5-20 June	7-17 June	28 May-7 June	2-20 June	28 May	15-18 June
Hatching of larvae	from 28 June	28 June-29 July	25-26 June	late June	18-22 July	25 June	27 June	25 June	20 June	3-23 June	15-28 June
The emergence of a new generation	14-29 July	until mid-August	17-24 July	17-23 July	until mid-August	20-25 July	18-28 July	18-25 July	until mid-August	9-24 July	15-25 July
number (ex/m ²)	To 50	to 47	1-35	2-35	2-40	to 50	to 40	to 40	to 40	5,0-51	to 40

Going to wintering grounds	10 August-III decade September	3-24 August	17-20 August – 1st half of September	18-21 August-III decade of September	15-20 August	18-22 August	10-25 August	18-25 August	10-22 August	14-25 August-7-10 September	25 August-5 September
....

Thus, 10 factors were considered, the normalization results of which are shown in Table 1. To build a specialized normalization model, data from such factors as air temperature, relative humidity and leaf moisture, soil temperature, radiation, wind speed were included. This data is needed to reduce uncertainties during calibration and evaluation.

Other datasets at a lower level of detail may be collected from real fields to validate the model development and calibration performed using the detailed dataset.

Table 1 presents data in such units of measurement and ranges of values as in months, ex/m², percentages, degrees. The measurement data was used in a numerical format, since the numerical format allows us to compare and normalize the data. And in such a way that none of them has any advantages over others. the main condition for proper normalization is that all signs should be equal in the possibilities of their influence.

After normalization, all numeric values of input features will be reduced to the same area of their change – some narrow range. Data normalization will allow us to prepare input data for machine learning and ensure the correct operation of computational algorithms (Peshawa, 2014).

Table 2. Lower pest data (initial stage) from Table 1

Exit from wintering grounds on grasses	Number (ex/m ²)	Migration to crops	Number (ex/m ²)	Damage to the leaf surface (%)	Mating	Hatching of larvae	The emergence of a new generation	Number (ex/m ²)	Going to wintering grounds
14.04	20	01.06	100	6	05.06	28.06	14.07	20	10.08
10.04	9	10.05	100	1	26.05	28.06	10.08	15	03.08
16.04	0.4	28.05	100	8	05.06	25.06	17.07	7	17.08
21.04	30	24.05	2	1	09.06	30.06	17.07	7	18.08

20.04	15	25.05	50	5	12.06	18.07	10.08	8	15.08
11.04	20	15.05	90	5	05.06	25.06	20.07	10	18.08
15.04	20	25.05	45	2	07.06	27.06	18.07	8	10.08
20.04	18	02.06	35	3	28.05	25.06	18.07	8	18.08
10.04	10	24.05	10	3	02.06	20.06	10.08	8	10.08
15.04	15	20.05	10	7	28.05	03.06	09.07	11	14.08
17.04	17	29.05	33	6	15.06	15.06	15.07	9	25.08

Table 3. Upper data from Table 1

Exit from wintering grounds on grasses	Number (ex/m2)	Migration to crops	Number (ex/m2)	Damage to the leaf surface (%)	Matting	Hatching of larvae	The emergence of a new generation	Number (ex/m2)	Going to wintering grounds
05.05	50	10.06	300	22	18.06	30.06	29.07	50	20.09
21.04	18	20.05	300	22	06.06	29.07	15.08	47	24.08
22.04	1	31.05	300	20	21.06	26.06	24.07	35	15.09
25.04	80	28.05	25	5	16.06	30.06	23.07	35	20.09
29.04	70	02.06	150	25	19.06	22.07	15.08	40	20.08
19.04	60	20.05	300	23	20.06	25.06	25.07	50	22.08
20.04	50	28.05	150	25	17.06	27.06	28.07	40	25.08
05.05	50	04.06	300	22	07.06	25.06	25.07	40	25.08
25.04	38	26.05	100	25	20.06	20.06	15.08	40	22.08
05.05	40	20.05	100	50	28.05	23.06	24.07	51	10.09
23.04	50	29.05	200	22	18.06	28.06	25.07	40	05.09

Data normalization in Python is used to make the machine have to process a smaller range of data. The results depend on the choice of the correct normalization method. The popular *sklearn* method was chosen to normalize the data in this research. The work uses the *NumPy* and *sklearn* libraries for machine learning of the Python programming language. The *preprocessing* class is imported from the *sklearn* library. The *preprocessing* class allows access to the *normalize()* function. That is why this method is called the normalization method of the *sklearn* library. Next, a *NumPy* array is created with unique integer values, then the *normalize()* method from *preprocessing* is called and *numpy_array* is passed as a parameter. All integer data is normalized between zero and one.

Results

As a result of the research, more than 1000 input data were considered, taking into account the upper and lower bounds.

To begin with, we take the lower data in the form presented in Table 2 and normalize the data in the Python program, at the end we get the normalized data presented in Table 4.

Table 4. Normalized lower data from Table 2

0	1	2	3	4	5	6	7	8	9
0.0	0.127390	0.181468	0.009618	0.907339	0.054440	0.045911	0.254599	0.127663	0.181468
0.091460	1	0.091326	0.081866	0.091417	0.909622	0.009096	0.236957	0.255240	0.091690
0.028016	2	0.144079	0.003593	0.251958	0.898245	0.071860	0.045451	0.225100	0.153330
0.062877	3	0.351692	0.501462	0.402005	0.033431	0.016715	0.151441	0.502465	0.285332
0.117008	4	0.294068	0.220111	0.367585	0.733702	0.073370	0.176969	0.265160	0.147914
0.117392	5	0.108482	0.196526	0.147886	0.884366	0.049131	0.049721	0.246247	0.197214
0.098263	6	0.222798	0.296274	0.371084	0.666617	0.029627	0.104585	0.400859	0.267684
0.118510	7	0.313123	0.281248	0.032187	0.546871	0.046875	0.438278	0.391560	0.282342
0.124999	8	0.254030	0.253018	0.608507	0.253018	0.075905	0.052122	0.507553	0.255042
0.202414	9	0.320663	0.319810	0.427479	0.213206	0.149245	0.598044	0.065241	0.193378
0.234527	0.300195								

Next, we take the upper data from Table 3, we also run it through the Python program and get the normalized data in Table 5.

Table 5. Normalized upper data from Table 3

0	1	2	3	4	5	6	7	8	9
0.0	0.016124	0.159647	0.032121	0.957879	0.070244	0.057664	0.095980	0.092819	0.159647
0.064146	1	0.068066	0.058231	0.064863	0.970519	0.071171	0.019604	0.094043	0.048785
0.152048	2	0.071505	0.003244	0.100737	0.973300	0.064887	0.068326	0.084547	0.078091
0.113552	3	0.230647	0.736892	0.258373	0.230279	0.046056	0.147931	0.276887	0.212501
0.322390	4	0.162456	0.391596	0.011524	0.839135	0.139856	0.106626	0.123465	0.084361
0.223769	5	0.060349	0.190175	0.063550	0.950873	0.072900	0.063582	0.079430	0.079461
0.158479	6	0.114089	0.284654	0.159691	0.853961	0.142327	0.097124	0.154055	0.159805
0.227723	7	0.016251	0.160905	0.013065	0.965427	0.070798	0.022720	0.080645	0.080678
0.128724	8	0.194970	0.295880	0.202834	0.778632	0.194658	0.156194	0.156194	0.117418
0.311453	0.171922								

9 0.036514 0.289216 0.144970 0.723040 0.361520 0.202813 0.166733 0.174036 0.368751 0.072955
--

A specialized normalization model is presented in Table 6, where data from such factors as air temperature, relative humidity and leaf moisture, soil temperature, radiation, wind speed were included.

Table 6. Custom Model

0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	
0	0.016124	0.159647	0.032121	0.957879	0.070244	0.057664	0.095980	0.092819	0.159647	0.064146	0.018750	0.082031	0.017578	0.019922	0.996088	0.004687
1	0.068066	0.058231	0.064863	0.970519	0.071171	0.019604	0.094043	0.048785	0.152048	0.077900	0.013054	0.054520	0.012286	0.013822	0.998249	0.003839
2	0.071505	0.003244	0.100737	0.973300	0.064887	0.068326	0.084547	0.078091	0.113552	0.048957	0.023159	0.083372	0.017369	0.022001	0.995831	0.006948
3	0.230647	0.736892	0.258373	0.230279	0.046056	0.147931	0.276887	0.212501	0.322390	0.185052	0.021014	0.088482	0.017696	0.022121	0.995423	0.007742
4	0.162456	0.391596	0.011524	0.839135	0.139856	0.106626	0.123465	0.084361	0.223769	0.112332	0.023722	0.077097	0.016605	0.021350	0.996329	0.009489
5	0.060349	0.190175	0.063550	0.950873	0.072900	0.063582	0.079430	0.079461	0.158479	0.069984	0.019278	0.082466	0.016065	0.020349	0.996023	0.009639
6	0.114089	0.284654	0.159691	0.853961	0.142327	0.097124	0.154055	0.159805	0.227723	0.142782	0.021261	0.083926	0.014547	0.020142	0.995920	0.005595
7	0.016251	0.160905	0.013065	0.965427	0.070798	0.022720	0.080645	0.080678	0.128724	0.080710	0.023261	0.069784	0.015508	0.021046	0.996918	0.007754
8	0.194970	0.295880	0.202834	0.778632	0.194658	0.156194	0.156194	0.117418	0.311453	0.171922	0.018402	0.084431	0.017319	0.021649	0.995852	0.006495
9	0.036514	0.289216	0.144970	0.723040	0.361520	0.202813	0.166733	0.174036	0.368751	0.072955	0.019710	0.077745	0.016425	0.018615	0.996454	0.005475

Normalized data for the striped bread flea obtained by using the *normalize()* function. This data is already ready for training. The training of the obtained data will be carried out in the next research. Thus, we have obtained the results of normalization.

After normalization, all numeric values in the input features are moved to the same area of change (a certain narrow range). This combines them into one machine learning model and guarantees the correct operation of computational algorithms. For clarity, Figure 1 shows a graph of normalized upper and lower data of a bread striped flea.

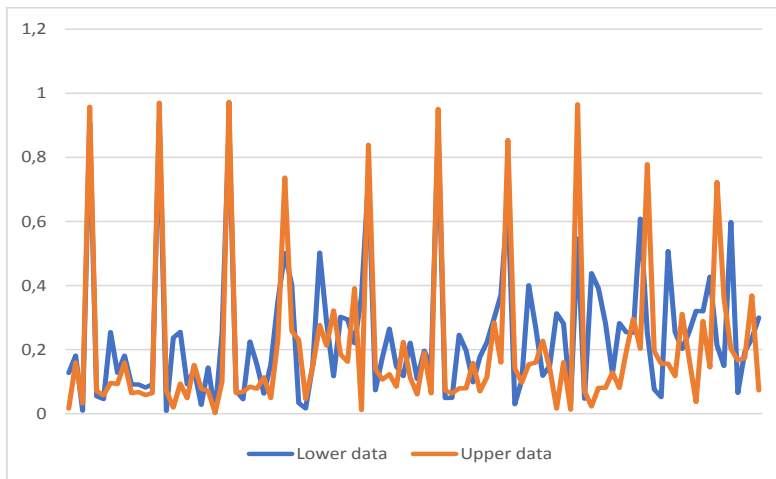


Fig.1. Normalized data in the range [0..1]

Discussion

When applying the lower bounds from Table 2, the normalized data from Table 4 is larger than from Table 3, where we took the upper data and got the normalized data from Table 5. That is, the lower the input data, we get a larger normalized data.

Normalization plays an important role for further research in network training. All pest detection data is ready, that is, it is normalized and ready for training in a neural network. Data normalization generally speeds up training and results in faster convergence. Normalizing the input data according to certain criteria before the training process is crucial for getting good results, as well as for significantly speeding up calculations. Agricultural research has profited from technological advances such as data mining, automation.

Conclusion

So, to achieve the goal is to determine the factors and indicators for the normalization of input data, for use in predicting the growth dynamics of the striped flea on the basis of an analytical system to increase productivity. Such tasks were implemented as in the first task: identifying factors and collecting numerical data to identify the pest, which will serve as data for machine learning. We collected data, identified the main factors for the detection of the striped bread flea. The fields of Northern Kazakhstan were studied from 2011 to 2021 and the results are shown in Table 1. The second task was to normalize the collected data for use in neural network training. Of all the collected data, the normalization indicated in Tables 3-6 was described and carried out.

As a result, the goal of the research was achieved and optimal normalized data were obtained for further training data in a neural network to predict the pest population. The continuation of this research will be work on creating a neural network model.

REFERENCES

- Condran S., Bewong M., Islam M. Z., Maphosa L., Zheng L., 2022 — *Condran S., Bewong M., Islam M. Z., Maphosa L., Zheng L.* Machine learning in precision agriculture: A survey on trends, applications and evaluations over two decades // *IEEE Access*, – 2022. – vol. 10. – Pp. 73786–73803. <https://doi.org/10.1109/ACCESS.2022.3188649>.
- Fahad R. Albogamy, 2021 — *Fahad R. Albogamy* A deep convolutional neural network with batch normalization approach for plant disease detection // *IJCSNS International Journal of Computer Science and Network Security*, – 2021. – Vol.21. – № 9. – Pp. 51–62. http://paper.ijcsns.org/07_book/202109/20210906.pdf.
- Hu X., Ren H., Tansey K., Zheng Y., Ghent D., Liu X., Yan L., 2019 — *Hu X., Ren H., Tansey K., Zheng Y., Ghent D., Liu X., Yan L.* Agricultural drought monitoring using european space agency sentinel 3A land surface temperature and normalized difference vegetation index imageries // *Agricultural and Forest Meteorology*, – 2019, – Vol. 279. – Pp. 1–9. <https://doi.org/10.1016/j.agrformet.2019.107707>.
- Kozulina N.S., Vasilenko A.V., Vasilenko A.A., Shmeleva Z.N., 2021 — *Kozulina N.S., Vasilenko A.V., Vasilenko A.A., Shmeleva Z. N.* Effective protection of grain crops from pests // *IOP Conference Series: Earth and Environmental Science*. – 2021. – Vol. 677. – № 4. Pp. 1–7. <https://doi.org/10.1088/1755-1315/677/4/042014>.
- Liakos K.G., Busato P., Moshou D., Pearson S., Bochtis D., 2018 — *Liakos K.G., Busato P., Moshou D., Pearson S., Bochtis D* Machine learning in agriculture: A review // *Sensors*. – 2018, – Vol. 18. – № 8. – Pp. 1–9. <https://doi.org/10.3390/s18082674>.
- Magarey R.D., Sutton T.B., Thayer C.L., 2005 — *Magarey R.D., Sutton T.B., Thayer C.L.* A simple generic infection model for foliar fungal plant pathogens // *Phytopathology*. – 2005. – Vol. 95. – № 1. – Pp. 92–100. <https://doi.org/10.1094/PHYTO-95-0092>
- Normalization of Data in Python. <https://linuxhint.com/normalization-of-data-in-python/>.
- Ogasawara E., Martinez L.C., De Oliveira D., Zimbrão G., Pappa G.L., Mattoso M., 2010 — *Ogasawara E., Martinez L.C., De Oliveira D., Zimbrão G., Pappa G. L., Mattoso M.* Adaptive normalization: A novel data normalization approach for non-stationary time series // *Proceedings of the International Joint Conference on Neural Networks*, – 2010. – Pp. 1–8. <https://doi.org/10.1109/IJCNN.2010.5596746>.
- Peshawa J. Muhammad Ali, Rezhna H. Faraj, 2014 — *Peshawa J. Muhammad Ali, Rezhna H. Faraj* Data Normalization and Standardization: A Technical Report // *Machine Learning Technical Reports*. – 2014, – Vol. 1. – № 1. – Pp 1–6. <https://doi.org/10.13140/RG.2.2.28948.04489>.
- Sarker I.H., 2021 — *Sarker I.H.* Machine learning: Algorithms, real-world applications and research directions // *SN Computer Science*. – 2021, – Vol. 160. – № 2. – Pp 1–21. <https://doi.org/10.1007/s42979-021-00592-x>.
- Shi Z., Bai Y., Jin X., Wang X., Su T., Kong J., 2022 — *Shi Z., Bai Y., Jin X., Wang X., Su T., Kong J.* Deep prediction model based on dual decomposition with entropy and frequency statistics for nonstationary time series // *Entropy*. – 2022. – Vol. 24. – № 3. – Pp. 1–16. <https://doi.org/10.3390/e24030360>.
- Starovoitov V.V., Golub Yu.I., 2021 — *Starovoitov V.V., Golub Yu.I.* Data normalization in machine learning // *Informatics*. – 2021. – Vol. 18. – №3. – Pp. 83–96. <https://doi.org/10.37661/1816-0301-2021-18-3-83-96>.
- Talukder B., Hipel K.W., vanLoon G.W., 2017 — *Talukder B., Hipel K.W., vanLoon G.W.* Developing Composite Indicators for Agricultural Sustainability Assessment: Effect of Normalization and Aggregation Techniques // *Resources*. – 2017. – Vol. 6. – № 4. – Pp. 1–27. <https://doi.org/10.3390/resources6040066>.
- Welch S.M., Croft B.A., Brunner J.F., Michels M.F., 1978 — *Welch S.M., Croft B.A., Brunner J.F., Michels M.F.* PETE: an extension phenology modeling system for management of multi-species pest complex // *Environ. Entomol.* – 1978. – Vol. 7. – № 4. – Pp. 487–494. <https://doi.org/10.1093/ee/7.4.487>.
- Zhang S., Jing R., Shi X., 2022 — *Zhang S., Jing R., Shi X.* Crop pest recognition based on a modified capsule network // *Systems Science and Control Engineering*. – 2022. – Vol. 10. – №1. – Pp. 552–561. <https://doi.org/10.1080/21642583.2022.2074168>.

NEWS OF THE NATIONAL ACADEMY OF SCIENCES OF THE REPUBLIC OF
KAZAKHSTAN
PHYSICO-MATHEMATICAL SERIES
ISSN 1991-346X
Volume 2. Number 346 (2023). 55–69
<https://doi.org/10.32014/2023.2518-1726.184>

UDK 681.518.3

© **A.E. Abzhanova^{1*}, A.I. Takuadina², S.K. Sagnaeva¹, S.K. Serikbayeva¹,
G.T. Azieva³, 2023**

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

²Karaganda Medical University, Karaganda, Казахстан;

³Esil University, Astana, Kazakhstan.

E-mail: alyoka.01@mail.ru

THE USE OF INFORMATION SYSTEMS IN THE METHODS OF TECHNICAL SOIL RECLAMATION

Abzhanova Ainagul Eralievna — master. Senior Lecturer of the Department of Information Systems of L.N. Gumilyov Eurasian National University, Astana, Kazakhstan

E-mail: abdygappar@mail.ru. ORCID: 681.518.3;

Takuadina Aliya Ibragimovna — Professor of the Department of Informatics and Biostatistics of the Medical University of Karaganda, Karaganda, Kazakhstan

E-mail: alyoka.01@mail.ru. ORCID: 681.518.3;

Saginayeva Saule Kayrollaevna — Candidate of Physical and Mathematical Sciences, Associate Professor of the Department of Control Automation, ENU named after L.N. Gumilyov, Astana, Kazakhstan

E-mail: sagnaeva_tar@mail.ru. ORCID: 681.518.3;

Serikbayeva Sandugash Kurmanbekovna — PHD Doctor of the Department of Information Systems of L.N. Gumilyov Eurasian National University, Astana, Kazakhstan

E-mail: Inf_8585@mail.ru. ORCID: 681.518.3;

Azieva Gulmi Tagibergenovna — Senior Lecturer, Department of Information Systems, Esil University, Astana, Kazakhstan

E-mail: gulmira_azieva@mail.ru. ORCID: 681.518.3.

Abstract. This article examines the integration of information systems into technical methods of soil improvement. It explores how modern technologies and data-based approaches can improve the efficiency of soil reclamation processes. The study contains an in-depth analysis of the various methodologies used, the benefits derived from their implementation, and potential problems that may arise. By studying the existing literature and case studies, this study aims to highlight the importance of information systems in improving soil conditions and to give an idea of future developments in this area.

Keywords: information systems, soil improvement, technical reclamation, data-based approaches, reclamation efficiency

© А.Е. Әбжанова^{1*}, А.И. Тақуадина², С.К. Сағнаева¹, С.К. Серикбаева¹,
Г.Т. Азиева³, 2023

¹Л.Н. Гумилев атындағы ЕҰУ, Астана, Қазақстан;

²Қарағанды Медициналық Университеті, Қарағанды, Қазақстан;

³Esil University, Астана, Қазақстан.

E-mail: alyoka.01@mail.ru

ТОПЫРАҚТЫ ТЕХНИКАЛЫҚ МЕЛИОРАЦИЯЛАУ ӘДІСТЕРІНДЕ АҚПАРАТТЫҚ ЖҮЙЕЛЕРДІ ПАЙДАЛАНУ

Әбжанова Айнагүл Ералықызы — магистр. Л.Н. Гумилева атындағы ЕҰУ ақпараттық жүйелер кафедрасының аға оқытушысы, Астана, Қазақстан

E-mail: abdygappar@mail.ru. ORCID: 681.518.3;

Тақуадина Алия Ибрагимовна — Қарағанды медицина университетінің информатика және биостатистика кафедрасының қауымдастырылған профессоры, Қарағанды, Қазақстан

E-mail: alyoka.01@mail.ru. ORCID: 681.518.3;

Сағнаева Сауле Кайроллаевна — ф-м.ғ.к., Л.Н. Гумилев атындағы ЕҰУ басқаруды автоматтандыру кафедрасының доценті, Астана, Қазақстан

E-mail: sagnaeva_tar@mail.ru. ORCID: 681.518.3;

Серикбаева Сандугаш Курманбековна — PhD доктор. Л.Н. Гумилева атындағы ЕҰУ ақпараттық жүйелер кафедрасы, Астана, Қазақстан

E-mail: Inf_8585@mail.ru. ORCID: 681.518.3;

Азиева Гүльми Тагирбегеновна — Esil University Ақпараттық жүйелер кафедрасының аға оқытушысы, Астана, Қазақстан

E-mail: gulmira_azieva@mail.ru. ORCID: 681.518.3.

Аннотация. Бұл мақалада ақпараттық жүйелердің топырақты жақсартудың техникалық әдістеріне интеграциясы зерттеледі. Ол заманауи технологиялар мен деректерге негізделген тәсілдер топырақты қалпына келтіру процестерінің тиімділігін қалай арттыра алатынын зерттейді. Зерттеуде қолданылатын әртүрлі әдістемелердің терең талдауы, оларды жүзеге асырудан алынған артықшылықтар және туындауы мүмкін мәселелер бар. Қолданыстағы әдебиеттер мен жағдайлық зерттеулерді зерттей отырып, бұл зерттеу ақпараттық жүйелердің топырақ жағдайын жақсартудағы маңыздылығын атап өтуге және осы саладағы болашақ әзірлемелер туралы түсінік беруге бағытталған.

Түйін сөздер: Ақпараттық жүйелер, топырақты жақсарту, техникалық мелиорация, деректерге негізделген тәсілдер, мелиорацияның тиімділігі

©А.Е. Абжанова^{1*}, А.И. Такуадина^{2*}, С.К. Сагнаева¹, С.К. Серикбаева¹,
Г.Т. Азиева³, 2023

¹ЕНУ имени Л.Н. Гумилева, Астана, Қазақстан;

²Медицинский Университет Караганды, Караганда, Қазақстан;

³Esil University, Астана, Қазақстан.

E-mail: alyoka.01@mail.ru

ИСПОЛЬЗОВАНИЕ ИНФОРМАЦИОННЫХ СИСТЕМ В МЕТОДАХ ТЕХНИЧЕСКИХ МЕЛИОРАЦИЙ ГРУНТОВ

Абжанова Айнагуль Ералиевна — магистр. Старший преподаватель кафедры информационных системы ЕНУ имени Л.Н.Гумилева, Астана, Қазақстан

E-mail: abdygappar@mail.ru. ORCID: 681.518.3;

Такуадина Алия Ибрагимовна — профессор кафедры информатики и биостатистики Медицинского университета Караганды, Караганда, Қазақстан

E-mail: alyoka.01@mail.ru. ORCID: 681.518.3;

Сагнаева Сауле Кайроллаевна — к.ф.-м.н., доцент кафедры автоматизации управления ЕНУ имени Л.Н. Гумилева, Астана, Қазақстан

E-mail: sagnaeva_tar@mail.ru. ORCID: 681.518.3;

Серикбаева Сандугаш Курманбековна — PhD доктор кафедры информационных системы ЕНУ имени Л.Н. Гумилева, Астана, Қазақстан

E-mail: Inf_8585@mail.ru. ORCID: 681.518.3;

Азиева Гульми Тагибереновна — Старший преподаватель кафедры информационных системы Esil University, Астана, Қазақстан

E-mail: gulmira_azieva@mail.ru. ORCID: 681.518.3.

Аннотация. В данной статье исследуется интеграция информационных систем в технические методы улучшения почвы. В нем исследуется, как современные технологии и подходы, основанные на данных, могут повысить эффективность процессов мелиорации почв. В исследовании содержится углубленный анализ различных используемых методологий, преимуществ, получаемых от их внедрения, и потенциальных проблем, которые могут возникнуть. Изучая существующую литературу и тематические исследования, данное исследование направлено на то, чтобы подчеркнуть важность информационных систем в улучшении состояния почв и дать представление о будущих разработках в этой области.

Ключевые слова: информационные системы, улучшение почв, техническая мелиорация, подходы, основанные на данных, эффективность мелиорации

Кіріспе

Топырақ сапасы көптеген салаларда, соның ішінде ауыл шаруашылығында, құрылыста және табиғатты ұтымды пайдалануда шешуші рөл атқарады. Топырақ жағдайын жақсарту немесе мелиорация деп аталатын топырақтың қасиеттері мен сипаттамаларын жақсарту оңтайлы өнімділік пен тұрақтылықты

қамтамасыз етудің маңызды міндеті болып табылады. Ақпараттық жүйелер мен озық технологиялардың пайда болуымен деректерге негізделген тәсілдерді енгізу топырақты жақсартудың дәстүрлі әдістерін айтарлықтай өзгертті. Мелиорацияның техникалық әдістерінде ақпараттық жүйелердің қолданылуын зерттей отырып, олардың артықшылықтарын бөліп көрсетуге және оларды енгізуге байланысты мәселелерді қарастыруға болады.

Жан-жақты талдау жүргізу үшін бұл зерттеу ғылыми журналдарды, конференция материалдарын және беделді онлайн-дерекқорларды қоса алғанда, тиісті әдебиеттерге жүйелі шолуға сүйенеді. Тандалған мақалалар топырақты жақсарту әдістерінде Ақпараттық жүйелер мен деректерге негізделген тәсілдерді қолдануға арналған. Ақпараттық жүйелердің мелиорация әдістерінің тиімділігіне әсерін түсіну үшін әртүрлі әдістемелер мен жағдайлық зерттеулер талданады. Шолу процесі жалпы тенденцияларды, артықшылықтар мен шектеулерді анықтау мақсатында әртүрлі көздерден алынған нәтижелерді жалпылау мен салыстыруды қамтиды.

Әдістер мен материалдар

Топырақты жақсартудың техникалық әдістеріне ақпараттық жүйелерді біріктіру, топырақты сынау және талдау, геокеңістіктік деректерді біріктіру, шешімдерді қолдау жүйелері, бақылау және бағалау сияқты әртүрлі аспектілерді қамтиды. Ақпараттық жүйелерді пайдалану топыраққа қатысты деректерді тиімді жинауды, сақтауды және талдауды қамтамасыз етеді, дәлелді шешімдер қабылдауды және нақты араласуды жеңілдетеді. Сонымен қатар, бұл жүйелер нақты уақыттағы кері байланыс пен адаптивті басқару стратегияларын қамтамасыз ету арқылы бақылау және бағалау процестерін жақсартады.

Ақпараттық жүйелерді біріктіруден алынған артықшылықтарға топырақ деректерінің дәлдігі мен сенімділігін арттыру, мелиорация әдістерінің тиімділігі мен рентабельділігін арттыру, өнімділік пен тұрақтылықты арттыру және ресурстарды бөлуді оңтайландыру кіреді. Дегенмен, деректердің құпиялылығы мен қауіпсіздігі, техникалық шектеулер және білікті қызметкерлер мен инфрақұрылымға деген қажеттілік сияқты мәселелер анықталып, талқыланады.

Топырақты жақсартудың техникалық әдістерінде ақпараттық жүйелерді қолдану өзін мелиорацияның дәстүрлі әдістерін өзгертетін трансформациялық тәсіл ретінде көрсетті. Осы зерттеуде жүргізілген талдау шешім қабылдау процесін жақсартуды, тиімділік пен тұрақтылықты арттыруды қоса алғанда, ақпараттық жүйелерді біріктірудің маңызды артықшылықтарын көрсетеді. Дегенмен, бұл жүйелерді енгізуге байланысты мәселелерді назардан тыс қалдыруға болмайды. Топырақты жақсарту саласындағы ақпараттық жүйелердің әлеуетін толық іске асыру үшін осы мәселелерді шешу және деректер қауіпсіздігін, функционалдық үйлесімділікті және ыңғайлы интерфейстерді қамтамасыз ететін сенімді құрылымдарды әзірлеу үшін қосымша зерттеулер қажет (Мустафаев, 2007).

Біріншіден, деректердің құпиялылығы мен қауіпсіздігіне қатысты мәселелерді шешуге күш салу керек. Ақпараттық жүйелердің интеграциясы топыраққа қатысты құпия деректерді жинауды және сақтауды қамтитындықтан, бұл деректердің құпиялылығы мен тұтастығын қамтамасыз ету өте маңызды. Сенімді шифрлау әдістерін әзірлеу, деректерді басқарудың қорғалған хаттамаларын енгізу және құпиялық ережелерін сақтау осы жүйелердің сеніміне ие болу және кеңінен енгізу үшін маңызды болады.

Екіншіден, топырақ жағдайын жақсарту кезінде ақпараттық жүйелердің тиімділігін арттыру үшін техникалық шектеулерді жою қажет. Бұл топырақ сенсорлары мен деректерді жинау құрылғыларының дәлдігі мен сенімділігін арттыруды, деректерді өңдеу алгоритмдерін жетілдіруді және әртүрлі деректер көздерін біркелкі біріктіруді қамтиды. Топырақтану, геокеңістіктік технологиялар және ақпараттық жүйелер саласындағы сарапшылар арасындағы бірлескен зерттеулер мен әзірлемелер осы техникалық қиындықтарды жеңуге ықпал етеді (Мустафаев, 2012).

Сонымен қатар, топырақ жағдайын жақсарту және басқару үшін ақпараттық жүйелерді тиімді пайдалануға қабілетті білікті персоналдың болуын қамтамасыз ету үшін әлеуетті арттыруға және оқыту бағдарламаларына инвестиция салу қажет. Бұған деректерді талдау, ГАЖ (географиялық ақпараттық жүйе) және топырақты басқаруға арналған шешімдерді қолдау жүйелерін қолдану бойынша мамандар даярлау кіреді. Жеке тұлғаларға қажетті дағдыларды бере отырып, ұйымдар мен мүдделі тараптар топырақты қалпына келтіру үшін ақпараттық жүйелердің әлеуетін толық пайдалана алады.

Операциялық жүйелер мен стандартталған деректер пішімдерін әзірлеу топыраққа қатысты деректерді үздіксіз біріктіру және оларды әртүрлі платформалар мен мүдделі тараптар арасында бөлісу үшін өте маңызды. Деректермен алмасу хаттамаларын құру, деректердің жалпы модельдерін құру және ашық деректер бастамаларын ілгерілету ынтымақтастыққа ықпал етеді және топырақ жағдайын жақсартуға бағытталған күш-жігер аясында ақпараттық жүйелердің ауқымдылығын жеңілдетеді.

Ақпараттық жүйелерді топырақты жақсартудың техникалық әдістеріне біріктіру топырақтың мелиорациясының тиімділігін арттыру үшін айтарлықтай әлеует ашатынын атап өткен жөн. Деректерге негізделген тәсілдерді қолдана отырып, шешім қабылдаушылар саналы таңдау жасай алады, ресурстарды бөлуді оңтайландырады және жерді басқарудың тұрақты әдістерін насихаттай алады. Алайда, осы артықшылықтарды толығымен жүзеге асыру үшін деректердің құпиялылығына, техникалық шектеулерге және білікті қызметкерлердің болуына байланысты мәселелерді шешу маңызды. Зерттеулерді жалғастыру, ынтымақтастық және топырақ жағдайларын жақсарту үшін ақпараттық жүйелерге Инвестициялар біздің топырақтарымыздың тұрақты дамуы мен сақталуына ықпал ететін инновациялық шешімдерге жол ашады (Кирейчева, 2017).

Топырақ жағдайын жақсартуға арналған ақпараттық жүйелерді зерттеу сонымен қатар жасанды интеллект, машиналық оқыту және қашықтықтан зондтау сияқты озық технологиялардың интеграциясын зерттеуі керек. Жасанды интеллект пен машиналық оқыту алгоритмдері үлкен деректер жиынтығын талдай алады және топырақтың мінез-құлқы мен қоректік заттарды басқарудың болжамды үлгілерін жасауға мүмкіндік беретін үлгілерді анықтай алады. Бұл модельдер топырақ пен тыңайтқыштарды Өзгертуді оңтайландыруға көмектеседі, бұл дақылдардың өнімділігін арттыруға және қоршаған ортаға әсерді азайтуға әкеледі.

Қашықтықтан зондтау технологиялары, соның ішінде спутниктік суреттер мен ұшқышсыз ұшу аппараттары үлкен аумақтардағы топырақ жағдайын бағалау үшін құнды құралдарды ұсынады. Жоғары ажыратымдылықтағы кескіндер мен спектрлік деректерді пайдалана отырып, ақпараттық жүйелер ылғал, органикалық заттар және қоректік заттар сияқты топырақ қасиеттерінің егжей-тегжейлі карталарын ұсына алады. Бұл Кеңістіктік егжей-тегжейлі деректер жиынтығы фермерлер мен жер менеджерлеріне белгілі бір проблемалық аймақтарды шешуге және ресурстарды бөлуді оңтайландыруға мүмкіндік беру арқылы топырақты жақсарту іс-шараларының мақсаттарын дәл анықтауға көмектеседі.

Сонымен қатар, ақпараттық жүйелерді дәл егіншілік әдістерімен біріктіру топырақ жағдайын жақсарту әдістерін жетілдіруге перспективалар ашады. Нақты уақыттағы сенсорлық деректерді, геокеңістіктік ақпаратты және машиналарды басқару жүйелерін біріктіре отырып, дәл егіншілік белгілі бір учаскеге байланысты топыраққа өзгерістер енгізуге, суаруға және басқарудың басқа әдістеріне мүмкіндік береді. Бұл тәсіл ресурстардың жоғалуын азайтады, қоршаған ортаға әсерді азайтады және топырақты жақсарту шараларының тиімділігін арттырады.

Болашақ зерттеулердің тағы бір саласы пайдаланушыға ыңғайлы интерфейстер мен шешімдерді қолдау құралдарын әзірлеуге бағытталуы мүмкін. Ақпараттық жүйелер қол жетімді және интуитивті болатындай етіп жасалуы керек, бұл әртүрлі деңгейдегі техникалық білімі бар мүдделі тараптарға жүйені оңай басқаруға және пайдалануға мүмкіндік береді. Шешімдерді қолдау құралдары пайдаланушыларға топырақ жағдайын жақсарту стратегиялары туралы негізделген шешімдер қабылдауға көмектесу арқылы топырақ деректеріне, жергілікті жағдайларға және алдын ала белгіленген мақсаттарға негізделген ұсыныстар бере алады.

Ақпараттық жүйелер арқылы жүзеге асырылатын топырақ жағдайын жақсарту жөніндегі іс-шараларды ұзақ мерзімді бақылау және бағалау олардың тиімділігін бағалау және жақсарту бағыттарын анықтау үшін өте маңызды. Нақты уақыттағы сенсорлық желілер мен деректерді тіркеу жүйелерін біріктіру адаптивті басқару стратегияларын жеңілдету және топырақты жақсарту әдістерінің тиімділігі туралы кері байланыс беру арқылы топырақ параметрлерін үздіксіз бақылауды қамтамасыз ете алады (Кирейчева, 2010).

Топырақ жағдайын жақсартуға арналған ақпараттық жүйелер жасанды интеллект, машиналық оқыту және қашықтықтан зондтау сияқты озық технологиялардың интеграциясын зерттеуге бағытталуы керек. Сонымен қатар, пайдаланушыға ыңғайлы интерфейстерді, дәл егіншілік әдістерін және шешімдерді қолдау құралдарын әзірлеу топырақты мелиорациялау саласындағы ақпараттық жүйелердің практикалық және тиімділігін одан әрі арттырады. Іс-шаралардың ұзақ мерзімді мониторингі мен бағалауы дәлелді әдістерді қолдануға және топырақты басқаруды үнемі жетілдіруге ықпал етеді. Осы салаларға сілтеме жасай отырып, зерттеушілер мен тәжірибешілер топырақ жағдайын жақсартудың тұрақты әдістерін қолдауда және топырақтың ұзақ мерзімді денсаулығы мен өнімділігін қамтамасыз етуде ақпараттық жүйелердің толық әлеуетін аша алады.

Ақпараттық жүйелер дамып, жетілдіріле бергенде, топырақтың болашақ жақсаруына уәде беретін бірнеше жаңа тенденциялар мен әлеуетті зерттеу бағыттары пайда болады (Кирейчева 2009).

Осындай тенденциялардың бірі-Интернет заттары технологиясын топырақты қалпына келтіруге арналған ақпараттық жүйелерге біріктіру. Топырақтың ылғалдылық датчиктері, метеостанциялар және автоматтандырылған суару жүйелері сияқты интернет заттары құрылғылары нақты уақыт режимінде топырақ жағдайы мен қоршаған орта факторлары туралы деректерді бере алады. Бұл құрылғыларды ақпараттық жүйелерге қосу арқылы мүдделі тараптар деректерге қашықтан қол жеткізе алады және оларды талдай алады, бұл белсенді шешімдер қабылдауға және уақтылы араласуға мүмкіндік береді. Заттар интернеті мен ақпараттық жүйелердің мұндай үздіксіз интеграциясы үздіксіз бақылауды, деректерді автоматтандырылған жинауды және икемді басқару стратегияларын қамтамасыз ету арқылы топырақ жағдайын жақсарту тәжірибесінде төңкеріс жасай алады.

Қызығушылықтың тағы бір жаңа бағыты-топырақ жағдайын жақсарту кезінде үлкен деректер аналитикасы мен болжамды модельдеуді қолдану. Үлкен деректер жиынының мүмкіндіктерін пайдалана отырып, аналитиканың озық әдістері топырақ деректеріндегі жасырын үлгілер мен корреляцияларды анықтай алады, бұл топырақтың мінез-құлқы мен Мелиорация әдістеріне реакцияның дәлірек болжамдарына әкеледі. Болжамдық модельдер топырақтың нақты түрлеріне, климаттық жағдайларға және ауылшаруашылық жүйелеріне арналған топырақ жағдайын жақсартудың оңтайлы стратегияларын анықтауға көмектеседі, осылайша іс-шаралардың тиімділігін арттырады.

Сонымен қатар, геокеңістіктік ақпарат пен картографиялық технологияларды енгізу топырақтың өзгергіштігін кеңістіктік түсінуді жақсарты алады және топырақ жағдайын жақсарту бойынша мақсатты іс-шараларды жүргізуге көмектеседі. Географиялық ақпараттық жүйелер топырақ деректерін егжей-тегжейлі топырақ карталарын жасау және дәл басқару тәсілдерін жеңілдету үшін топография, жерді пайдалану және

дақылдардың жарамдылығы сияқты басқа кеңістіктік қабаттармен біріктіре алады. Бұл Кеңістіктік егжей-тегжейлі ақпарат шешім қабылдау процестерін басқара алады, белгілі бір сайт үшін топыраққа өзгерістер енгізуді қамтамасыз етеді және жерді тұрақты пайдалануды жоспарлауды қолдайды.

Топырақты сауықтыру және қалпына келтіретін егіншілік тұжырымдамасы қарқын алуда және ақпараттық жүйелер осы тәжірибелерді қолдауда шешуші рөл атқара алады. Топырақ сапасының кешенді көрсеткіштеріне негізделген топырақ жағдайын бағалау құралдары топырақ жағдайы туралы тұтас түсініктерді қамтамасыз ету және топырақ жағдайын жақсарту стратегияларын басқару үшін ақпараттық жүйелерге біріктірілуі мүмкін. Ақпараттық жүйелер сонымен қатар тұрақты және тұрақты ауылшаруашылық жүйелеріне ықпал ете отырып, егіншілік, ауыспалы егіс және органикалық өзгерістер енгізу сияқты қалпына келтіру әдістерін құжаттауды және бақылауды жеңілдетуі мүмкін.

Пәнаралық ынтымақтастық пен серіктестік топырақ жағдайын жақсарту саласындағы ақпараттық жүйелердің болашағы үшін өте маңызды. Топырақтану, агрономия, ақпараттық жүйелерді және деректерді талдау сияқты салалардағы сарапшылардың бірігуі инновациялық тәсілдер мен идеялардың айқас тозаңдануына ықпал етуі мүмкін. Бірлескен зерттеу күштері деректердің үйлесімділігі, масштабталуы және стандарттау сияқты ақпараттық жүйелерге қатысты көп қырлы мәселелерді шешуге көмектеседі және әзірленген шешімдердің соңғы пайдаланушылардың практикалық қажеттіліктеріне сәйкес келуін қамтамасыз етеді (Шабанов, 1982).

Айта кету керек, топырақты жақсарту саласындағы ақпараттық жүйелердің болашағы үлкен әлеуетке ие. Интернет заттар технологияларын, үлкен деректерді талдауды, геокеністіктік ақпаратты және қалпына келтіретін егіншілік әдістерін біріктіру топырақ мелиорациясының тиімділігі мен тұрақтылығын одан әрі арттыра алады. Пәнаралық ынтымақтастық пен серіктестік инновацияларды ынталандырудың және ақпараттық жүйелерге қатысты күрделі мәселелерді шешудің кілті болады. Осы жаңа тенденциялар мен зерттеу бағыттарын пайдалана отырып, мүдделі тараптар топырақ жағдайын жақсартудың тиімдірек, деректерге негізделген және экологиялық саналы әдістеріне жол ашуы мүмкін.

Топырақты жақсарту саласындағы ақпараттық жүйелер, тағы бір зерттеу саласы блокчейн технологиясында жатыр. Блокчейн өзінің орталықтандырылмаған және өзгермейтін табиғатымен деректерді басқару мен топырақты жақсарту әдістерінің ашықтығын өзгерту мүмкіндігіне ие. Блокчейнді пайдалана отырып, топырақ сынағы нәтижелері, қоректік заттарды басқару жоспарлары және топыраққа өзгерістер енгізу жазбалары сияқты топыраққа қатысты деректерді сенімді түрде тіркеуге және мүдделі тараптарға беруге болады. Бұл технология сенім мен есептілікті арттыра алады, жеткізу тізбегін бақылауды жеңілдетеді және топыраққа қатысты ақпараттың тұтастығын қамтамасыз етеді.

Сонымен қатар, азаматтық ғылым мен кең қатысуға негізделген тәсілдерді ақпараттық жүйелерге біріктіру топырақ деректерін жинаудың дәлдігі мен көлемін одан әрі арттыра алады. Азаматтық ғылыми бастамалар жұртшылықты деректерді жинау, талдау және шешім қабылдау процестеріне тартуды көздейді. Мобильді қосымшаларды, краудсорсингтік платформаларды және азаматтық ғылыми желілерді қолдана отырып, ақпараттық жүйелер азаматтық ғалымдардың ұжымдық күшін қолдана алады, бұл оларға әртүрлі жерлерден және жерді пайдаланудан құнды топырақ деректерін беруге мүмкіндік береді. Азаматтардан алынған бұл деректер дәстүрлі деректер көздерін толықтыра алады және топырақ жағдайларын толық түсінуді қамтамасыз етеді, бұл топырақ жағдайын жақсартудың неғұрлым негізделген стратегияларын жасауға әкеледі (Духовный, 2015).

Қашықтықтан зондтау технологиялары мен машиналық оқыту алгоритмдерін дамыту сонымен қатар топырақ жағдайын жақсарту саласындағы Болашақ ақпараттық жүйелердің болашағын ашады. Жоғары ажыратымдылықтағы спутниктік суреттер, гиперспектрлік датчиктер және жарықты анықтау және диапазонды анықтау жүйесі топырақ пен өсімдік жамылғысының қасиеттері туралы егжей-тегжейлі және нақты ақпарат бере алады. Машиналық оқыту алгоритмдері осы деректер көздерін талдай алады және топырақтың жылдам және үнемді картографиясы мен мониторингін қамтамасыз ете отырып, құнды ақпаратты ала алады. Бұл жетістіктер, әсіресе кең ауқымды немесе жету қиын жерлерде, топырақ жағдайын жақсарту шараларының тиімділігі мен дәлдігін айтарлықтай жақсарта алады.

Климаттың өзгеруі топырақты басқаруда қиындықтар туғызуды жалғастыра отырып, ақпараттық жүйелер климаттың өзгеруін ескере отырып, топырақты жақсарту әдістерін қолдау үшін климаттық деректер мен модельдеуді біріктіре алады. Температура мен жауын-шашынның өзгеруі сияқты климаттық болжамдарды ескере отырып, ақпараттық жүйелер климаттың өзгеруін азайту үшін топырақты жақсарту стратегияларын бейімдеуге көмектеседі. Бұл суару кестелерін түзетуді, климатқа төзімді дақылдарды таңдауды және топырақтың ылғал сақтау қабілетін арттыру үшін топыраққа органикалық тыңайтқыштарды енгізуді қамтуы мүмкін.

Блокчейн технологиясынан, азаматтық ғылымнан, қашықтықтан зондтаудан, машиналық оқытудан, климаттық деректерді біріктіруден және ашық деректер принциптерінен туындайтын мүмкіндіктермен, осы жетістіктерді енгізу және пәнаралық ынтымақтастықты дамыту топырақты басқарудың тұрақты әдістеріне қол жеткізуде ақпараттық жүйелердің барлық әлеуетін пайдалану үшін маңызды болады. Осы технологиялар мен тәсілдерді қолдана отырып, мүдделі тараптар негізделген шешімдер қабылдай алады, ресурстарды бөлуді оңтайландырады және біздің топырақтың ұзақ мерзімді денсаулығы мен өнімділігіне үлес қоса алады.

Болашақта топырақ жағдайын жақсарту үшін ақпараттық жүйелерді

әзірлеу және енгізу кезінде туындауы мүмкін кейбір ықтимал мәселелер мен ойларды қарастырайық:

Деректердің сапасы мен стандартталуы: топыраққа қатысты деректердің дәлдігін, сенімділігін және дәйектілігін қамтамасыз ету тиімді ақпараттық жүйелер үшін өте маңызды. Топырақты зерттеу зертханалары, қашықтықтан зондтау технологиялары және азаматтық ғылыми бастамалар сияқты әртүрлі көздерден жиналған деректер сапасы мен пішімінде әртүрлі болуы мүмкін. Деректерді жинау, сақтау және бөлісу үшін стандартталған хаттамалар мен сапаны бақылау шараларын құру деректердің тұтастығын және олардың функционалдық үйлесімділігін сақтау үшін маңызды болады.

Масштабтау және инфрақұрылым талаптары: топырақ жағдайын жақсарту үшін ақпараттық жүйелер дамыған сайын масштабтау маңызды факторға айналады. Ауқымды енгізу деректерді сақтау, өңдеу және желілік қосылымды қоса алғанда, сенімді және ауқымды инфрақұрылымды қажет етеді. Топыраққа қатысты деректердің өсіп келе жатқан көлемін және ақпараттық жүйелердің өсіп келе жатқан пайдаланушы базасын қолдау үшін тиісті ресурстар мен инвестициялар қажет.

Қол жетімділік және цифрлық алшақтық: ақпараттық жүйелерге тең қол жетімділікті қамтамасыз ету бар теңсіздіктің нашарлауына жол бермеу үшін өте маңызды. Интернет байланысы немесе технологиялық ресурстары шектеулі аймақтарда ақпараттық жүйелерге қол жетімділік және оларды пайдалану шектеулі болуы мүмкін. Цифрлық алшақтықты жоюға және топырақ жағдайын жақсарту бастамаларына инклюзивті қатысуды қамтамасыз ету үшін аз қамтылған қауымдастықтарды оқытуға және қолдауға күш салу керек.

Деректердің құпиялылығы мен қауіпсіздігі, ақпараттық жүйелер жерді пайдалану схемаларын, топырақ сынақтарының нәтижелерін және агрономиялық әдістерді қоса алғанда, топыраққа қатысты құпия деректерді өңдейді. Бұл ақпараттың құпиялылығы мен тұтастығын қорғау үшін деректердің құпиялылығы мен қауіпсіздігін қамтамасыз ету үшін сенімді шаралар қабылдануы керек. Деректерді қорғау ережелерін сақтау және деректерді басқарудың қауіпсіз әдістерін енгізу мүдделі тараптардың сенімін сақтау үшін өте маңызды.

Интеграция және интероперабельділік: ақпараттық жүйелердің әлеуетін толық пайдалану үшін әртүрлі платформалар, мәліметтер базасы және деректер форматтары арасындағы интероперабельділік өте маңызды. Стандарттау әрекеттері әртүрлі жүйелер арасында үздіксіз деректер алмасуды және интеграцияны қамтамасыз ету үшін жалпы деректер үлгілерін, онтологияларды және қолданбалы бағдарламалау интерфейстерін әзірлеуге бағытталуы керек. Бағдарламалық жасақтама жасаушылар, зерттеушілер және топырақты басқару практиктері арасындағы ынтымақтастық функционалдық үйлесімділікке қол жеткізу үшін өте маңызды.

Экономикалық тиімділік және тұрақтылық: ақпараттық жүйелердің ұзақ мерзімді өміршеңдігі мен тұрақты дамуы олардың экономикалық тиімділігіне байланысты аппараттық және бағдарламалық жасақтаманың қол жетімділігі, техникалық қызмет көрсету және жаңарту шығындары, сондай-ақ техникалық қолдаудың болуы сияқты ойларды ескеру қажет. Қаржыландыру тетіктерін, мемлекеттік-жекеменшік әріптестікті және шығыстарды бірлесіп атқару модельдерін зерделеу топырақтың жай-күйін жақсарту үшін ақпараттық жүйелердің қаржылық тұрақтылығына ықпал етуі мүмкін (Духовный, 2015).

Ақпараттық жүйелер топырақ жағдайын жақсарту әдістерін ілгерілету үшін үлкен әлеуетке ие болғанымен, бірнеше мәселелерді шешу қажет. Деректер сапасы мен стандарттау, масштабтау, қолжетімділік, мүдделі тараптарды тарту, деректердің құпиялылығы мен қауіпсіздігі, функционалдық үйлесімділік, экономикалық тиімділік және тұрақтылық мәселелерін шешу, ақпараттық жүйелерді сәтті енгізу және топырақты жақсарту әрекеттеріне ұзақ мерзімді әсер ету үшін өте маңызды болады. Осы міндеттерді шеше отырып, мүдделі тараптар осы жүйелердің толық әлеуетін ашып, топырақты тұрақты басқаруға және өнімділікті арттыруға өз үлестерін қоса алады.

Нәтижелер және оларды талқылау

Топырақты жақсарту саласында ақпараттық жүйелерді сәтті енгізу мүдделі тараптардың жаңа технологияларды қолдануға және өздерінің дәстүрлі әдістерін өзгертуге дайындығына байланысты. Фермерлерді, жер менеджерлерін және шешім қабылдаушыларды топырақ жағдайын жақсарту және өнімділікті арттыру үшін ақпараттық жүйелердің әлеуетті артықшылықтары туралы тарту және оқыту өте маңызды. Пилоттық жобалар мен жағдайлық зерттеулер арқылы осы жүйелердің практикалық, пайдаланудың қарапайымдылығы мен оң нәтижелерін көрсету өзгерістерге төзімділікті жеңуге және кеңірек енгізуге ықпал етеді.

Сонымен қатар, ақпараттық жүйелер туралы ықтимал мәселелерді немесе қате түсініктерді жою өте маңызды, мысалы, технология адам тәжірибесін алмастырады деп қорқу немесе іске асырудың болжамды күрделілігі, ақпараттық жүйелерді пайдаланудың қосымша құндылығына, тиімділігіне және ұзақ мерзімді пайдасына баса назар аудару оларды енгізуге оң көзқарасты қалыптастырады және мінез-құлықтың өзгеруіне ықпал етеді.

Білімді тиімді беруді және әлеуетті арттыруды қамтамасыз ету, сондай-ақ мінез-құлықты енгізу мен өзгертуге байланысты мәселелерді шешу топырақты жақсарту саласындағы ақпараттық жүйелердің болашағы үшін маңызды ойлар болып табылады. Мүдделі тараптарға қажетті дағдылар мен білім беру және өзгерістерге қолайлы жағдай жасау арқылы осы жүйелерді кеңінен енгізуге және сәтті енгізуге қол жеткізуге болады. Бұл өз кезегінде топырақты тұрақты басқаруға, ауылшаруашылық өнімділігін арттыруға және қоршаған ортаны қорғауға ықпал етеді.

Шешім қабылдау процесін жақсарту: Ақпараттық жүйелер мүдделі

тараптарға топыраққа қатысты жан-жақты деректерге, талдау құралдарына және шешімдерді қолдау тетіктеріне қол жеткізуге мүмкіндік береді. Бұл оларға топырақты жақсарту стратегиялары, қоректік заттарды ұтымды пайдалану және ауылшаруашылық тәжірибелері туралы негізделген шешімдер қабылдауға мүмкіндік береді. Нақты уақыттағы деректерді және болжамды модельдерді қолдану арқылы ақпараттық жүйелер алдын-ала шешім қабылдауға мүмкіндік береді, бұл дәлірек және уақтылы араласуға әкеледі.

Топырақ жағдайы мен өнімділігін жақсарту, ақпараттық жүйелер топырақ денсаулығына қатысты мәселелерді анықтауды және топырақ жағдайын жақсарту бойынша мақсатты шараларды жүзеге асыруды жеңілдетеді. Топырақ деректері мен тарихи тенденцияларды талдай отырып, мүдделі тараптар топыраққа, тыңайтқыштарға және суару әдістеріне өзгерістер енгізуді оңтайландыруы мүмкін. Бұл топырақ құнарлылығының жоғарылауына, қоректік заттардың айналымының жақсаруына және дақылдардың өнімділігінің жоғарылауына әкеледі, бұл сайып келгенде ауылшаруашылық өндіріс жүйелеріне пайда әкеледі.

Ақпараттық жүйелер белгілі бір сайтқа байланысты топыраққа, суға және басқа ресурстарға нақты өзгерістер енгізуге мүмкіндік береді. Топырақ, ауарайы және егін жинау талаптары туралы деректерді біріктіру арқылы мүдделі тараптар қалдықтарды азайту және тиімділікті арттыру арқылы ресурстарды бөлуді оңтайландыра алады. Бұл шығындарды үнемдеуге, қоршаған ортаға әсерді азайтуға және топырақты басқару әдістерінің тұрақтылығын арттыруға әкеледі.

Қоршаған ортаны қорғау, ақпараттық жүйелер қоршаған ортаға теріс әсерді барынша азайтатын жер ресурстарын басқарудың тұрақты әдістерін енгізуге ықпал етеді. Қоректік заттарды қолдануды және суару әдістерін оңтайландыру арқылы Ақпараттық жүйелер қоректік заттардың ағуын, топырақ эрозиясын және жер асты суларының ластануын болдырмауға көмектеседі. Бұл су ресурстарын сақтауға, экожүйелерді қорғауға және агрохимикаттарды пайдалануды азайтуға ықпал етеді.

Климатқа төзімділік: Ақпараттық жүйелер ауылшаруашылық жүйелеріне климаттың өзгеруіне бейімделуге және оның әсерін азайтуға көмектесу арқылы климаттық жағдайларды ескере отырып, топырақты жақсарту әдістерін қолдай алады. Климаттық деректерді, болжамды модельдерді және климаттық жағдайларды ескере отырып, ауылшаруашылық тәжірибелерін біріктіре отырып, мүдделі тараптар топырақтың ылғал ұстау қабілетін арттыратын, топырақ эрозиясын азайтатын және көміртекті байланыстыруды жақсартатын топырақты басқару стратегияларын енгізе алады. Бұл климаттың өзгеруіне байланысты ауылшаруашылық жүйелерінің тұрақтылығын арттырады.

Ақпараттық жүйелер мүдделі тараптар арасында топыраққа, озық тәжірибеге және алынған сабақтарға қатысты деректермен алмасуды жеңілдетеді. Бұл

зерттеушілер, фермерлер, кеңейту қызметтері мен саясаткерлер арасындағы ынтымақтастық пен білім алмасуды ынталандырады. Ынтымақтастық атмосферасын құруға ықпал ете отырып, ақпараттық жүйелер үздіксіз оқытуға, инновацияларға және топырақ жағдайын жақсартудың тиімді әдістерін таратуға ықпал етеді (Маслов, 2009).

Топырақ жағдайын жақсарту саласында ақпараттық жүйелерді кеңінен енгізу және табысты іске асыру көптеген пайда мен оң салдарға әкелуі мүмкін екенін атап өткен жөн. Ақпараттық жүйелер топырақты басқару әдістерін өзгерту мүмкіндігіне ие - шешім қабылдау процесін жақсартудан және топырақ жағдайын жақсартудан ресурстарды оңтайландыруға және қоршаған ортаны сақтауға дейін. Деректер, технологиялар және ынтымақтастық мүмкіндіктерін пайдалана отырып, мүдделі тараптар біздің топырақтарымыздың болашақ ұрпақ үшін ұзақ мерзімді денсаулығы мен өнімділігін қамтамасыз ете отырып, тұрақты және өміршең ауылшаруашылық жүйелерін құру үшін жұмыс істей алады.

Ақпараттық жүйелер топырақты жақсарту тәжірибесіне қатысатын мүдделі тараптарға айтарлықтай экономикалық пайда әкелуі мүмкін. Нақты уақыттағы деректерді, болжамды модельдеуді және дәл басқару тәсілдерін қолдана отырып, мүдделі тараптар ресурстарды бөлуді оңтайландырады, кіріс шығындарын азайтады және жалпы кірістілікті арттырады. Ақпараттық жүйелер сонымен қатар фермерлерге тұрақты өндірілетін дақылдар мен өнімдерге нарықтық сұранысты қанағаттандыруға мүмкіндік беретін топырақты басқару тәжірибесінің қадағалануы мен құжатталуын қамтамасыз ету арқылы нарыққа қол жеткізу және сертификаттау бағдарламаларын қолдай алады. Сонымен қатар, топырақтың денсаулығы мен өнімділігін жақсарту арқылы ақпараттық жүйелер ауылшаруашылық жүйелерінің ұзақ мерзімді өміршеңдігі мен экономикалық тұрақтылығына ықпал етеді.

Қорытынды

Ұзақ мерзімді экологиялық тұрақтылық: топырақ жағдайын жақсартуда ақпараттық жүйелерді енгізу ұзақ мерзімді экологиялық тұрақтылыққа ықпал етеді. Жабық егіншілік, табиғатты қорғау және агроорман шаруашылығы сияқты мақсатты топырақты басқару стратегияларын енгізу арқылы мүдделі тараптар топырақтағы органикалық заттарды, биоәртүрлілікті және экожүйе қызметтерін арттыра алады. Салауатты топырақтар атмосферадан көмірқышқыл газын алу арқылы климаттың өзгеруін азайтуға көмектесетін көміртекті сіңіргіш ретінде әрекет етеді. Сонымен қатар, ақпараттық жүйелер топырақты жақсарту әдістерінің қоршаған ортаға әсерін бақылауға және бағалауға мүмкіндік береді, адаптивті басқаруға және топырақты тұрақты басқару бағытында үнемі жетілдіруге ықпал етеді.

Топырақ жағдайын жақсартудағы ақпараттық жүйелердің артықшылықтары мен әсері контекстке, аймаққа және нақты енгізу стратегияларына байланысты өзгеруі мүмкін екенін ескеру маңызды. Ақпараттық жүйелерді әзірлеу және

енгізу кезінде топырақ түрлері, климаттық жағдайлар және әлеуметтік-экономикалық ойлар сияқты жергілікті факторларды ескеру қажет. Сонымен қатар, уақыт өте келе осы жүйелердің тиімділігі мен бейімделуін бағалау үшін үнемі бақылау, бағалау және кері байланыс механизмдері қажет.

Қорытындылай келе, топырақ жағдайын жақсарту кезінде ақпараттық жүйелерді енгізу және пайдалану экономикалық, экологиялық және әлеуметтік пайда әкелуі мүмкін екенін атап өткен жөн. Деректерге негізделген тәсілдерді қолдана отырып, мүдделі тараптар негізделген шешімдер қабылдай алады, ресурстарды пайдалануды оңтайландырады және ауылшаруашылық жүйелерінің ұзақ мерзімді тұрақтылығын арттырады. Дегенмен, табысты іске асыру ынтымақтастықты, әлеуетті арттыруды және белгілі бір контекстке қатысты жергілікті білім мен ойларды ескеруді қажет ететінін мойындау маңызды. Осы күш жігердің арқасында Ақпараттық жүйелер топырақты басқару әдістерін өзгертуге үлес қоса алады және ауыл шаруашылығының тұрақты және тұрақты болашағына жол ашады.

ӘДЕБИЕТТЕР

Мустафаев Ж.С., 2007. Почвенно-экологическое обоснование мелиорации сельскохозяйственных земель.-Алматы:Ғылым. 2007.-358 с.

Мустафаев Ж.С., Рябцев А.Д., 2012. Адаптивно-ландшафтные мелиорации земель в Казахстане. -Тараз:«TOOBIGNEJSERSERVICE», 2012. -528 с.

Научные основы создания и управления мелиоративными системами в России. / Под научной редакцией д-ра техн. наук, профессора Кирейчевой И.В. Монография. - М.: ФГБНУ "ВНИИ агрохимии", 2017. – 296.

Новые технологии проектирования, обоснования строительства, эксплуатации и управления мелиоративными/Под научной редакцией д-ра техн. наук, проф. И.В. Кирейчевой. - М.: ВНИИА, 2010. - 240 с.

Эколого-экономическая эффективность комплексных мелиораций Барабинской низменности/ под ред. И. В. Кирейчевой. -М.: ВНИИА, 2009. -312 с.

Shabanov V.V., 1982. Avtomatizatsiya kompleksnogo regulirovaniya faktorov zhizni rastenii / Shabanov V. V. // Gidrotekhnika i melioratsiya. - 1982. - № 1. - с. 60–75. - [Elektronnyi resurs]. - Rezhim dostupa: http://www.ieek.timacad.ru/kmirz/Htmls/works/-1982_41.pdf.

Духовный В.А., Нерозин С.А., Стулина Г.В., Солодкий Г.Ф., 2015. Программирование урожая сельскохозяйственных культур (системный подход в приложении к мелиорации). 2015. Ташкент.

Maslov B.S., 2009. Agricultural Land Improvement: Amelioration and Reclamation. Volume I // EOLSS Publications, 2009. - 444 p.

REFERENCES

Mustafaev Zh.S., 2007. Soil-ecological substantiation of agricultural land reclamation.-Almaty: Gylm. 2007.-358 p.

Mustafaev Zh.S., Ryabtsev A.D., 2012. Adaptive landscape land reclamation in Kazakhstan. - Taraz: "TOOBIGNEJSERSERVICE", 2012. -528 p.

Scientific foundations for the creation and management of ameliorative systems in Russia. / Under the scientific editorship of Dr. tech. Sciences, Professor Kireycheva I.V. Monograph. - M.: FGBNU "VNI agrochemistry", 2017. – 296.

New technologies for design, justification of construction, operation and management of reclamation / Under the scientific editorship of Dr. tech. sciences, prof. I.V. Kireycheva. - M.: VNIIA, 2010. - 240 p.

Ecological and economic efficiency of complex reclamation of the Baraba lowland / ed. J.T.V. Kireycheva. -M.: VNIIA, 2009. -312 p.

Shabanov V.V., 1982. Avtomatizatsiya kompleksnogo regulirovaniya faktorov zhizni rastenii / Shabanov V. V. // *Gidrotekhnika i melioratsiya*. - 1982. - № 1 - Pp. 60–75. - [Elektronnyi resurs]. - Rezhim dostupa: http://www.ieek.timacad.ru/kmirz/Htmls/works/-1982_41.pdf.

Dukhovny B.A., Nerozin C.A., Stulina G.V., Solodkiy G.F., 2015. Programming of agricultural crops (system approach in application to land reclamation). 2015. Tashkent.

Maslov B.S., 2009. Agricultural Land Improvement: Amelioration and Reclamation. Volume I // EOLSS Publications, 2009. - 444 p.

NEWS OF THE NATIONAL ACADEMY OF SCIENCES OF THE REPUBLIC OF
KAZAKHSTAN
PHYSICO-MATHEMATICAL SERIES
ISSN 1991-346X
Volume 2. Number 346 (2023). 70–80
<https://doi.org/10.32014/2023.2518-1726.185>

MPHTI 81.93.29
УДК 004.056.55

© **K. Alibekova^{1*}, Zh. Alimzhanova¹, S.S. Baizakova², 2023**

¹Al Farabi Kazakh National University, Almaty, Kazakhstan;

²Arkalyk State Pedagogical Institute named after Y. Alttynsarin,
Arkalyk, Kazakhstan.

E-mail: kamalibekova@gmail.com

RATING VALUATION OF BLOCK CIPHERS FOR WIRELESS SENSOR NETWORKS

Alibekova Kamila Nurdauletkyzy — Doctoral student of the Department of Information Systems of the Al-Farabi Kazakh National University, Almaty, Kazakhstan

E-mail: kamalibekova@gmail.com; <https://orcid.org/0000-0001-5039-7833>;

Alimzhanova Zhanna Muratbekovna — Candidate of Physical and Mathematical Sciences Al-Farabi Kazakh National University, Almaty, Kazakhstan

E-mail: zhannamen@mail.ru, <https://orcid.org/0000-0001-6282-5356>;

Baizakova Saule Svyazkhanovna — Senior lecturer, Master's degree, Arkalyk State Pedagogical Institute named after I. Alttynsarin

Abstract. Security is one of the extensive relations in assorted Wireless Sensor Network (WSN) employment. A no of cryptanalytic rules area been developed to acknowledge safe keeping utility in WSNs. However, in an energy-efficient and jackanapes nonbeing is a difficult engagement due care to imagination restricted disposition of device excrescences. Systematic exploration of cryptanalytic rules is, accordingly, requisite to acknowledge an admirable sympathetic of the trade-off between safekeeping about and operation charge. In this idea, we've contemplated cardinal shaft nonentities: Skipjack, Corrected Obstruct Tiny Encryption Algorithm (XXTEA, RC5, Advanced Encryption Standard (AES, and Chaotic-Map and Genetic-Operations supported Encryption Algorithm (CGEA. The about of these ciphers is assessed on Arduino Pro and Mica2 device particles. Then the memory practice, operation spell, and computational cost are compared. Finally, any exhortations are provided on assessed shaft ciphers and implementation platforms.

Key words: Block ciphers, memory and energy utilization efficiency, performance evaluation, wireless sensor networks

© К.Н. Әлібекова^{1*}, Ж.М. Алимжанова¹, С.С. Байзакова², 2023

¹Әл-Фараби атындағы Қазақ Ұлттық университеті, Алматы, Қазақстан;

²Ы.Алтынсарин атындағы Арқалық педагогикалық институты,
Арқалық, Қазақстан.

E-mail: kamalibekova@gmail.com

СЫМСЫЗ СЕНСОРЛЫҚ ЖЕЛІЛЕР ҮШІН БЛОКТЫҚ ШИФРЛАРДЫҢ ӨНІМДІЛІГІН БАҒАЛАУ

Әлібекова Камила Нұрдәулетқызы — PhD Докторанты, Әл-Фараби атындағы Қазақ Ұлттық университеті, Алматы, Қазақстан

E-mail: E-mail: kamalibekova@gmail.com. ORCID: <https://orcid.org/0000-0001-5039-7833>;

Алимжанова Жанна Муратбековна — Физика-математика ғылымдарының кандидаты, Әл-Фараби атындағы Қазақ Ұлттық университеті, Алматы, Қазақстан

E-mail: zhannamen@mail.ru ORCID: <https://orcid.org/0000-0001-6282-5356>;

Кошпанова Каламкас Есқатқызы — Магистр дәрежесі, аға оқытушы, Ы. Алтынсарин атындағы Арқалық мемлекеттік педагогикалық институты, Арқалық, Қазақстан

E-mail: saule_alikosh@mail.ru.

Аннотация. Қауіпсіздік көптеген сымсыз сенсорлық желі (WSN) қосымшаларындағы негізгі мәселелердің бірі болып табылады. WSNS-те қауіпсіздік қызметтерін ұсыну үшін бірқатар криптографиялық алгоритмдер жасалды. Дегенмен, энергияны үнемдейтін және жеңіл шифрды таңдау сенсорлық түйін ресурстарының шектеулі болуына байланысты қиын міндет болып табылады. Осылайша, криптографиялық алгоритмдерді жүйелі бағалау қауіпсіздік шаралары мен операциялық шығындар арасындағы романы жақсы түсінуді қамтамасыз ету үшін қажет. Бұл мақалада біз бес блоктық шифрды қарастырдық: Skipjack, түзетілген Block tiny шифрлау алгоритмі (ХТЕА), RC5, кеңейтілген шифрлау стандарты (AES) және хаотикалық карта мен генетикалық операцияларға негізделген шифрлау алгоритмі (CGEA). Бұл шифрлардың өнімділігі Arduino Pro және Mica2 сенсорларының көмегімен бағаланады. Содан кейін жадты пайдалану, жұмыс уақыты және есептеу шығындары салыстырылады. Соңында, бағаланатын блок шифрлары мен іске асыру платформалары бойынша кейбір ұсыныстар беріледі.

Түйін сөздер: Блоктық шифрлар, жад пен энергияны пайдалану тиімділігі, өнімділікті бағалау, сымсыз сенсорлық желілер

© К.Н. Алибекова^{1*}, Ж.М. Алимжанова¹, С.С. Байзакова², 2023

¹Казахский национальный университет имени аль-Фараби,
Алматы, Казахстан;

²Аркалыкский государственный педагогический институт
им. И. Алтынсарина, Аркалык, Казахстан.
E-mail: kamalibekova@gmail.com

ОЦЕНКА ПРОИЗВОДИТЕЛЬНОСТИ БЛОЧНЫХ ШИФРОВ ДЛЯ БЕСПРОВОДНЫХ СЕНСОРНЫХ СЕТЕЙ

Әлібекова Камила Нұрдәулетқызы — Докторант PhD, Казахский национальный университет им. Аль-Фараби, Алматы, Казахстан

E-mail: kamalibekova@gmail.com. ORCID: <https://orcid.org/0000-0001-5039-7833>;

Алимжанова Жанна Муратбековна — Кандидат физико-математических наук, Казахский национальный университет им. Аль-Фараби, Алматы, Казахстан

E-mail: zhannamen@mail.ru. ORCID: <https://orcid.org/0000-0001-6282-5356>;

Кошпанова Каламкас Ескалқызы — Степень магистра, старший преподаватель, Аркалыкский государственный педагогический институт им. И. Алтынсарина, Аркалык, Казахстан

E-mail: saule_alikosh@mail.ru.

Аннотация. Безопасность является одной из основных проблем во многих приложениях беспроводной сенсорной сети (WSN). Для предоставления услуг безопасности в WSNS был разработан ряд криптографических алгоритмов. Однако выбор энергоэффективного и легкого шифра является сложной задачей из-за ограниченности ресурсов сенсорных узлов. Таким образом, систематическая оценка криптографических алгоритмов необходима для обеспечения хорошего понимания компромисса между показателями безопасности и эксплуатационными расходами. В этой статье мы рассмотрели пять блочных шифров: Skipjack, исправленный алгоритм шифрования Block Tiny (XTEA), RC5, расширенный стандарт шифрования (AES) и алгоритм шифрования на основе хаотической карты и генетических операций (CGEA). Производительность этих шифров оценивается с помощью датчиков Arduino Pro и Mica2. Затем сравниваются использование памяти, время работы и вычислительные затраты. Наконец, даются некоторые рекомендации по оцениваемым блочным шифрам и платформам реализации.

Ключевые слова: блочные шифры, эффективность использования памяти и энергии, оценка производительности, беспроводные сенсорные сети

Кіріспе

Сымсыз сенсорлық желілер дұшпандық ортада кездейсоқ орналас-тырылатын көптеген арзан сенсорлық түйіндерден (SNS) тұрады (Mojisola, 2022). Батареямен жұмыс істейтін бұл SNS-тердің жадысы аз, процессорлары әлсіз және байланыс мүмкіндіктері шектеулі болып келеді. Сондықтан соңғы бірнеше жылда WSNs-те энергия тиімділігіне қол жеткізу үшін бірқатар

қауіпсіз және жеңіл блоктық шифрлар ұсынылды. Алайда, эксперименттік нәтижелер бұл шифрлардың көпшілігінің әдеттегі шифрлармен салыстырғанда өнімділігі төмен екенін көрсетеді (Soni, 2019) Осылайша, өнімділікті бағалау әртүрлі криптографиялық схемалардың эталонын қамтамасыз ету үшін маңызды.

Қауіпсіздік саясатын енгізу шығындар мен өнімділіктің сақталуын қамтамасыз етуі керек. Мысалы, көптеген WSN қосымшалары қауіпсіздіктің жоғарылауын қамтамасыз ету үшін күрделі криптографиялық алгоритмдерді қажет етеді. Алайда, криптожүйені іске асыру үшін қуатты SNS қажет болғандықтан, шығындар артады. Сондықтан іске асыру шығындары мен тиімділік арасындағы байланысты нақты түсіну қажет. 1-кестеде кеңінен қолданылатын сенсорлық құрылғылардың бірқатар жабдықтарының құны мен техникалық сипаттамалары туралы салыстырмалы мәліметтер келтірілген. Кестеден EZ430-RF2500 және Arduino Pro motes арзанырақ екенін көруге болады, бірақ олардың жады да аз. Демек, арзан SNS-те криптографиялық схемалардың өнімділігін бағалау экономикалық тиімді платформалар құру мүмкіндігін зерттеу үшін қажет.

Бұл мақалада нақты сенсорлық жабдыққа негізделген криптографиялық алгоритмдерді эксперименттік бағалау ұсынылған. Жад тиімділігін, есептеу шығындарын және жұмыс уақытын салыстыру үшін Mica2 және Arduino Pro mote платформаларында бірқатар блоктық шифрлар енгізілген. Соңында, эксперимент нәтижелеріне сүйене отырып, ең жақсы криптографиялық алгоритм мен іске асыру платформасын таңдау үшін пайдалы болатын кейбір маңызды тұжырымдар беріледі. Құжаттың қалған бөлігі келесідей ұйымдастырылған: 2-бөлімде сәйкесінше сипатталған жұмыс істейді. 3-бөлімде WSNs-те бағаланатын блоктық шифрларға шолу берілген. 4-бөлімде енгізу платформалары егжей-тегжейлі сипатталған. Тиімділікті бағалау және талдау 5-бөлімде ұсынылған. 6-бөлімде құжат талқыланады және аяқталады.

Атауы	Бағасы (USD)	Пакеті (incl.)	Микро-контроллер	Жолы (bits)	Уақыты (MHz)	RAM (KB)	Жылдамдығы (KB)	EEPROM (KB)
SHIMMER	226	2 boards	MSP430F1611	16	4-8	10	48	None
Waspnote	168	4 sensors	ATmega1281	8	14	8	128	4
TelosB	102	3 sensors	MSP430F1611	16	4-8	10	48	1024
Mica2	99	N/Av.	ATmega128L	8	8	4	128	512
EZ430-RF2500	40	Board only	MSP430F2274	16	16	1	32	None
Arduino Pro (328)	25	nrf24L01 radio	ATmega328	8	8-16	2	32	1

Кесте 1. Жабдықтың құны және техникалық сипаттамалары

Әдебиеттерге шолу

WSNs үшін қауіпсіздік қасиеттерін де, жад пен блоктық Шифр энергиясын пайдалану тиімділігін де ескеретін жүйелі бағалау жүйесін ұсынады. Авторлар Rijndael cipher-ді қауіпсіздік пен энергия тиімділігін қамтамасыз ету үшін пайдалануды ұсынады, ал misty1 деректерді сақтау және энергия тиімділігін арттыру үшін ұсынылады

Криптографиялық хэш функцияларының (MD5 және SHA1) өнімділігін, сондай-ақ шифрлау алгоритмін (AES) талдау үшін сынақ стендін пайдалану қажет. Нәтижелер криптографиялық алгоритм 1 КБ массиві бар бір AES операциясы үшін 1,67 с сияқты ұзақ жұмыс уақытын қажет ететінін көрсетеді.

Салыстырмалы өнімділікті талдау (Sun, 2007) гс6, AES және масштабталатын шифрлау алгоритмі (SEA) sea AES және RC6 шифрларымен салыстырғанда аз жадты қажет ететінін көрсетеді, ал AES және RC6 сәйкесінше жұмыс уақыты мен өткізу қабілеттілігін пайдалану тұрғысынан ең жақсы өнімділікке қол жеткізеді.

Симметриялы кілттері бар шифрларға арналған энергияның есептеу шығындары әр түрлі блок өлшемдері мен пайдалы жүктемелерді ескере отырып есептеледі және салыстырылады (Segar, 2013) Сонымен қатар, авторлар қауіпсіздікті де, энергия тиімділігін де қамтамасыз ету үшін байт-бағдарланған ауыстыру-ауыстыру желілік шифрын (SPN) пайдалануды ұсынады. Симметриялық блок шифрларының WSN өнімділігі мен мінез-құлқына әсері желінің маңызды параметрлерін анықтау үшін талданады (Babenko, 1999). AES, RC5 және Skipjack шифрлары MicaZ және TelosB чиптерінде жүзеге асырылады, сонымен қатар сапалық және сандық жағынан маңызды компаға келеді

Mica2 сенсорларының шаңындағы AES, RC5 және RC6 алгоритмдерінің жад тиімділігін, жұмыс уақытын және қуат тұтынуын өлшеп, салыстырды. Эксперименттердің нәтижелері RC5 уақыт пен энергия тиімділігі тұрғысынан ең қолайлы блоктық Шифр екенін көрсетеді. Кәдімгі криптографиялық алгоритмдерден басқа, Mica2 motes-те HIGH (Schneier, 2002) Simple Lightweight Encryption (Scheme Koo, 2008) және Lightweight Security Protocol (Biswas, 2014) сияқты бірнеше жеңіл блоктық шифрлар енгізілген. Бұл алгоритмдер энергияны үнемдейді және WSNS-те қауіпсіздіктің жақсы деңгейін қамтамасыз етеді. Бұл жұмыс екі түрлі аппараттық платформада бірқатар блоктық шифрларды жүзеге асырады және қауіпсіздік көрсеткіштері мен пайдалану шығындарын зерттейді. Эксперименттік нәтижелер бағаланған блоктық шифрлардың тиімділігін көрсетеді.

Блоктық шифрлар	Ұзындық кілті (биттер)	Раундтар	Блоктың ұзындығы (биттер)
Skipjack	80	32	64
XXTEA	128	14	64
RC5	128	14	64
AES-128	128	10	128

AES-256	256	14	128
CGEA	256	N/Av.	128

Кесте 2. Эксперименттерде қолданылатын шифр параметрлері

Материалдар мен әдістер

Бұл бөлімде skipjack, XXTEA, RC5, AES және CGEA шифрлары сипатталған. 2-кестеде біздің тәжірибелеріміздегі әрбір блок шифры үшін қабылданған параметрлер келтірілген.

Skipjack 64 биттік деректер блоктары үшін 80 биттік кілтті пайдаланады. Онда 32 раундтық теңгерімсіз фейстель желісі бар. Бихам және басқалар 32-ден 31-ге қарсы шабуыл жасады мүмкін емес дифференциалды криптоанализді қолданатын раундтар (Biswas, 2014). Сонымен қатар, кілттің қысқа ұзындығы Skipjack-ті толық кілт іздеумен шабуылдарға осал етеді.

XXTEA 64 биттік блок ұзындығына және 128 биттік кілт ұзындығына ие. Онда раундтардың ауыспалы саны бар (6-дан 32-ге дейін толық цикл) теңгерімсіз Фейстель желісі жүзеге асырылады. Толық функционалды xxtea-ға соңғы тіркелген шабуыл-бұл 259 сұранысты қолдана отырып, дифференциалды криптоанализге негізделген тандалған ашық мәтіндік Шабуыл (Biham, 2005).

RC5-айнымалы параметрлері бар икемді блок шифры: блок өлшемі (32, 64 немесе 128 бит), кілт өлшемі (0-ден 2040 битке дейін) және раундтар саны (0-ден 255-ке дейін). Бұл WSNs-те кеңінен қолданылатын блоктық шифр. Алайда, 64 биттік блоктары бар 12 раундтық RC5 тандалған 244 ашық мәтінді қолдана отырып, дифференциалды шабуылға осал (Yarkov, 1998).

AES-бұл ауыстыру-ауыстыру желісіне негізделген және 128 биттік блоктың бекітілген өлшемі бар қайталанатын блок шифры. Ол 4×4 байт массивімен жұмыс істейді. Сәйкесінше 128, 192 және 256 биттік кілттер үшін 10, 12 және 14 раундтарда жұмыс істейтін AES осал (Polley, 2004).

CGIAR-жалған кездейсоқ биттер тізбегін жасау үшін хаотикалық картаны пайдаланатын жеңіл блоктық шифр (Wu) 256 биттік реттілік блоктары 128 биттік деректер блоктарын шифрлау немесе шифрын ашу үшін кілт ретінде пайдаланылады. Алгоритм шифрланған мәтінді құру үшін XOR, мутация және ашық мәтінді қиылысу операцияларын жүзеге асырады. Раундтарды пайдаланудың орнына ол әрбір деректер байты үшін қиылысу әрекетін ашық мәтін ретінде орындайды.

Жабдықтың сипаттамасы

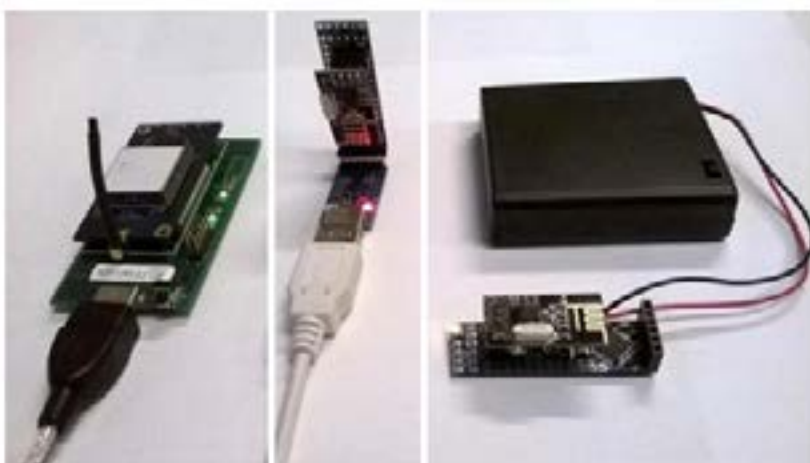
Arduino Pro-atmega168/328 негізіндегі микроконтроллер тақтасы (Biswas, 2013) Тәжірибелер келесі конфигурацияларда USB қуатымен жұмыс істейтін Arduino Pro (328) чиптерін қолданды: жұмыс кернеуі - 3,3 В, сағат жиілігі - МГц, жедел жады - 2 Кб, ФЛЭШ-жады - 32 Кб, EEPROM - 1 Кб, радио құрылғысы - nrf24L01 және деректер жылдамдығы - 19,2 Кбит/с Mica2 - бұл atmega128l процессорына негізделген төмен қуатты Mote сенсоры (Cazorla,

2013). Тәжірибелер келесі конфигурацияларда USB қуаты бар Mica2 чиптерін пайдаланды: жұмыс кернеуі - 3,3 В, сағат жиілігі - 8 МГц, жедел жады - 4 КБ, флэш-жады - 128 Кб, EEPROM - 512 Кб, радио құрылғысы - CC1000 және деректер жылдамдығы - 19,2 Кбит/с.

Бағдарламалық жасақтаманың сипаттамасы

Әр блоктық шифрдың бастапқы коды Arduino IDE-де Arduino Promotions-қа құрастыру және жүктеу үшін жазылады. Біздің тәжірибелерімізде екі кіріктірілген кітапхана функциясы (microsecondsToClockCycles() және Serial қолданылды. print()) процессор циклдары мен шифрлау уақытын алу және басып шығару үшін қолданылады. Mica2 чиптерінде шифрларды енгізу үшін компоненттерге негізделген жоғары деңгейлі бағдарламалау тілі (nesC) қолданылады (Young Adam and Moti Yung, 2005)..Жергілікті Уақыт Функциялары.get() және printf () жұмыс уақытын алу үшін пайдаланылады, ал процессор циклдары АТЕМУ [20] арқылы алынады.

Сурет. 1. Эксперименттік орнату: (а) USB қуаты бар Mica2 mote, (б) Arduino Pro бағдарламалаушы тақтасы бар USB қуаты бар mote, (с) батареямен жұмыс істейтін Arduino Promote



Соңында, AVR-size және AVR-objdump утилиталары сәйкесінше Arduino Pro және Mica2 motes құрылғыларында жадты пайдалануды өлшеу үшін қолданылады. Бұл екі утилита көрсетеді

нысан файлдарының тақырыптары туралы ақпарат. Ақпарат мәтінге, деректерге және bss бөліміне қатысты жедел жад пен ROM өлшемін қамтиды.

Тиімділікті бағалау және талдау

Блоктық шифрлар	Mica2	Arduino Pro
Skipjack	9820	12672
XXTEA	24064	30464

RC5	53014	61504
AES-128	37525	43200
AES-256	80344	88896
CGEA	67786	76212

Бұл бөлімде Mica2 және Arduino Promotions бағдарламаларында іске асырылған оңтайландырылған skipjack, XXTEA, RC5, AES және CGEA блок шифрларының салыстырмалы өнімділігі талдауы берілген. Салыстыру үшін үш маңызды параметр таңдалды: жадты тұтыну, есептеу шығындары және жұмыс уақыты.

Блоктық шифрлар	Mica2		Arduino Pro	
	RAM	ROM	RAM	ROM
Skipjack	3096	8658	398	4952
XXTEA	542	6312	226	4112
RC5	682	6110	350	3184
AES-128	1074	6296	814	3692
AES-256	1822	7932	1014	4190
CGEA	664	6268	548	3228

Жадты тұтыну

Жадты тұтыну-бұл аз жадты шифрлау алгоритмдерін таңдау үшін қолдануға болатын маңызды өнімділік көрсеткіші. 3-кестеде Mica2 және Arduino Pro платформаларында әрбір блок шифры тұтынатын жад көлемі көрсетілген. Skipjack және AES-256 басқа алгоритмдермен салыстырғанда көбірек жадты қажет ететінін көруге болады. AES-128 жадына қойылатын талаптар AES-256-мен салыстырғанда сәл төмен, ал RC5 барлық алгоритмдердің ішіндегі ең жеңілі.

Сындар - Skipjack және AES екеуі де 256 байтты үлкен s-блокты пайдаланады, нәтижесінде Алгоритмдер айтарлықтай жадты алады. Xxtea, RC5 және CGEA орындау үшін аз жадты қажет етеді, сондықтан Arduino Pro сияқты шектеулі жады бар SNS үшін қолайлы. Бір маңызды байқау-Arduino Promote-те AES-256 енгізу қол жетімді жадтың жетіспеушілігі туралы хабарламаны көрсетеді. Демек, артық жадты қолданатын криптографиялық Алгоритмдер Arduino Pro платформасында тұрақтылық мәселелеріне тап болуы мүмкін.

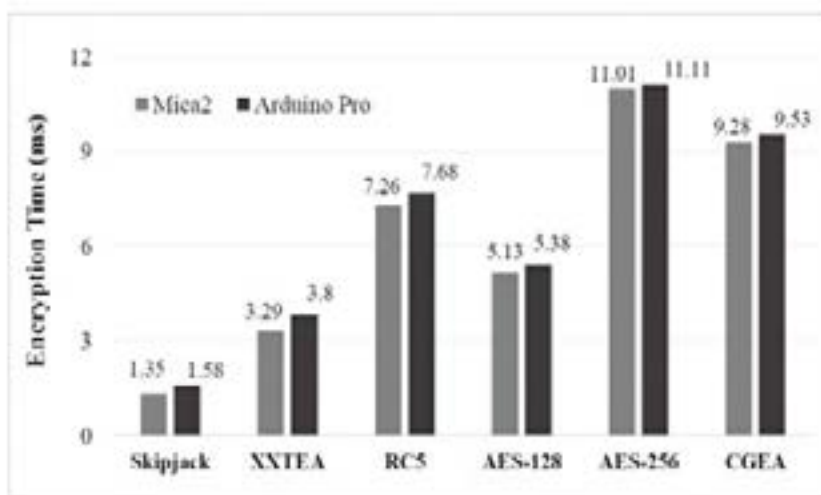
Есептеу шығындары

Алгоритмнің энергетикалық тиімділігін оның есептеу күрделілігіне қарай есептеуге болады. Процессор циклі үшін энергияны тұтыну тұрақты деп есептесек, оның мөлшері бір Байт үшін тұтынылатын энергияны бір Байт ашық мәтінді өңдеуге қажетті процессор циклдарының санын өлшеу арқылы есептеуге болады. Алайда, 4-кестеде жалпы саны көрсетілген 32 байт деректерді шифрлау үшін әр алгоритм талап ететін процессор циклдарының саны. Skipjack-бұл энергияны үнемдейтін блок-Шифр, ал AES-256 өнімділігі

барлық алгоритмдердің ішіндегі ең нашар екенін көруге болады. Сондай-ақ, AES 128 AES-256-дан екі есе жақсы жұмыс істейтіні атап өтілді.

Сын-кілттің өлшемі мен раундтар саны есептеу күрделілігінде маңызды рөл атқарады. AES-128 блок шифрын енгізу кілттің кішігірім өлшемі мен раундтардың аздығына байланысты AES-256 талап ететін есептеу шығындарының жартысынан көбін азайтады. Сондай-ақ, RC5 бірдей өлшемді кілттің болуына қарамастан, AES-128-мен салыстырғанда процессордың көп циклін тұтынады. Себебі, RC5 14 раундты орындайды, ал AES-128 тек 10 раундты пайдаланады.

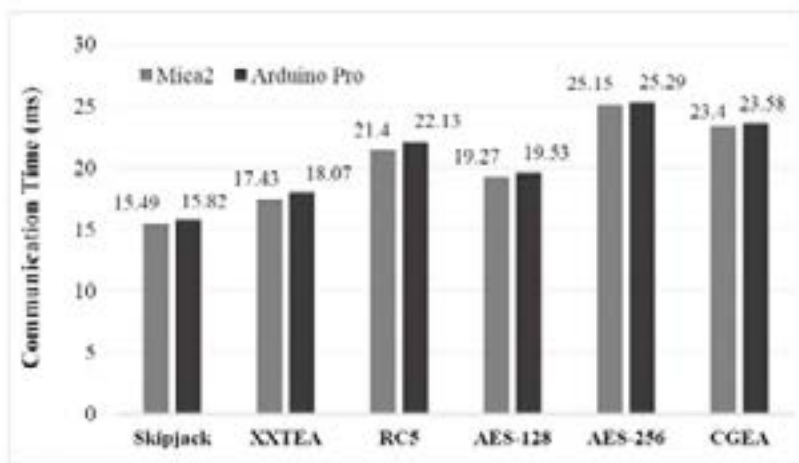
Сурет. 2. Шифрлау уақыты



Пайдалану шығындары

Жұмыс жылдамдығы уақытты пайдалану тиімділігін көрсетеді және шифрлау уақыты мен байланыс уақыты бойынша анықталады. Шифрлау уақыты-ашық мәтінді шифрлауға кететін уақыт мөлшері, ал шифрланған мәтінді шифрлауға және сәтті жіберуге кететін уақыт Байланыс уақыты ретінде анықталады. Інжір. 2 32 байтты деректерді шифрлау үшін қажетті жұмыс уақыты көрсетілген. Skipjack сәйкесінше AES және CGEA шифрларымен салыстырғанда 7 және 6 есе жылдам жұмыс істейтінін көруге болады. Сонымен қатар, AES-128 шифры AES-256-мен салыстырғанда шифрлау уақытын жартысынан астамға қысқартады. Дәл осындай нәтижелер суретте көрсетілгендей байланыс уақыты эксперименті үшін алынған.

Сурет. 3. Байланыс уақыты



Қорытынды

Эксперимент нәтижелеріне сәйкес, RC5 жад тұрғысынан ең тиімді блоктық Шифр болып табылады. Ххtea және CGEA сонымен қатар Arduino Pro сияқты жады шектеулі SNS үшін әлеуетті үміткерлер болып табылады. Екінші жағынан, Skipjack жұмыс уақыты мен есептеу шығындары бойынша ең жақсы өнімділікті көрсетеді. Ххtea және AES-128 шифрлары да аз қуат тұтынады. Алайда, қауіпсіздік тұрғысынан Skipjack-бұл кілттің қысқа ұзындығына байланысты жоғары тәуекел алгоритмі. Сол сияқты, XXTEA және RC5 уақыт шабуылы және тандалған ашық мәтінді қолданатын шабуыл сияқты бірқатар қауіпсіздік шабуылдарына осал. Сонымен қатар, 128 биттік кілт кванттық шабуылдан қорғалмайды. Кванттық есептеу жүйелері 264 секундта 128 биттік кілтті бұзуға қабілетті (Somsuk Kritsanapong, 2021). Дегенмен, AES256 және CGEA 256 биттік кілттің арқасында әлі де толық іздеуден қорғалған болады ұзындығы. Сондықтан қауіпсіздік басымдыққа ие болған кезде AES-256 немесе CGEA блок шифрларын пайдалануды ұсынамыз. RC5, XXTEA және AES-128 шифрларын қауіпсіздіктің минималды деңгейін қажет ететін қосымшалар үшін пайдалануға болады.

Бұл мақалада skipjack, XXTEA, RC5, AES және CGEA блоктық шифрларының салыстырмалы өнімділігі талдауы берілген. Айта кету керек, Arduino Pro mica2-ге қарағанда сәл ұзағырақ жұмыс уақыты мен байланысты қажет етеді, бұл шамалы. Осылайша, көп уақыт пен жадты қажет ететін қолданбалардың орнына қоршаған ортаны бақылау сияқты жалпы WSN қолданбалары үшін Arduino Pro пайдалану үнемді болады. Болашақ жұмысымызда біз ағындық шифрлардың өнімділігін бағалаймыз және нәтижелерді блоктық шифрлармен салыстырамыз.

REFERENCES

- Babenko J.I.K., 1999. Methodical manual: Introduction to the specialty "Organization and technology of information protection" Taganrog: TRTU, 1999. № 2800 E. Biham, A. Shamir: "Differential Cryptanalysis of the Full 16-round DES", *Crypto'92*, Springer-Verlag, 1998, p. 487.
- Biswas K., Muthukkumarasamy V., Sithirasanen E., Singh K., 2014. A simple lightweight encryption scheme for WSNs. In: Chatterjee, M., Cao, J., Kothapalli, K., Rajsbaum, S. (eds.) *ICDCN. LNCS. Vol. 8314. Pp. 499–504. Springer, Heidelberg (2014).*
- Biswas K., 2014. Lightweight security protocol for WSNs. In: *International Symposium on WoWMoM. Pp. 1–2, Sydney (2014).*
- Biham E., Biryukov A., Shamir A., 2005. Cryptanalysis of Skipjack reduced to 31 rounds using impossible differentials. In: *J. of crypt. 18(4). Pp. 291–311 (2005).*
- Biswas K., Muthukkumarasamy V., Sithirasanen E., 2013. Maximal clique based clustering scheme for WSNs. In: *IEEE ISSNIP. Pp. 237–241, Melbourne (2013).*
- Cazorla M., Marquet K., Minier M., 2013. Survey and benchmark of lightweight block ciphers for WSNs. In: *SECRYPT, (2013)*
- Intila C., Gerardo B., Medina R., 2019. A study of public key 'e' in RSA algorithm // *IOP Conference Series: Materials Science and Engineering. – IOP Publishing, 2019. – T. 482. – №. 1. – P. 012016.*
- Koo W.K., Lee H., Kim Y.H., Lee D.H., 2008. Implementation and analysis of new lightweight cryptographic algorithm suitable for WSNs. In: *ISA. Pp. 73–76, (2008).*
- Lee G. et al., 2019. PyWavelets: A Python package for wavelet analysis // *Journal of Open Source Software. – 2019. – T. 4. – №. 36. – P. 1237.*
- Mojisola F.O. et al., 2022. An improved random bit-stuffing technique with a modified RSA algorithm for resisting attacks in information security (RBMRSA) // *Egyptian Informatics Journal. – 2022.*
- Polley J., Blazakis D., Mcgee J., Rusk D., Baras J.S., 2004. ATEMU: a fine grained sensor network simulator. In: *IEEE SECON. Pp. 145–152, (2004).*
- Soni G.K., Arora H., Jain B., 2019. A novel image encryption technique using Arnold transform and asymmetric RSA algorithm // *International Conference on Artificial Intelligence: Advances and Applications 2019. – Springer, Singapore, 2020. – Pp. 83–90.*
- Somsuk Kritsanapong, 2021. "A New Methodology to Find Private Key of RSA Based on Euler Totient Function." *Baghdad Science Journal 18.2 (2021). Pp. 338–348.*
- Segar T.C., Vijayaragavan R., 2013. Pell's RSA key generation and its security analysis // *2013 Fourth International Conference on Computing, Communications and Networking Technologies (ICCCNT). – IEEE, 2013. – Pp. 1–5.*
- Schneier B., 2002. *Applied Cryptography. Protocols, algorithms, source texts in C. M.: TRIUMPH Publishing House, 2002.*
- Sun H.M. et al., 2007. Dual RSA and its security analysis // *IEEE Transactions on Information Theory. – 2007. – T. 53. – №. 8. – Pp. 2922–2933.*
- Thangavel M. et al., 2015. An enhanced and secured RSA key generation scheme (ESRKGS) // *Journal of information security and applications. – 2015. – T. 20. – Pp. 3–10.*
- Yarrkov E., 1998. Cryptanalysis of XXTEA, <http://eprint.iacr.org/2010/254.pdf>. Biryukov, A., Kushilevitz, E.: Improved cryptanalysis of RC5. In: N., Kaisa (eds.) *EUROCRYPT'98. LNCS. Vol. 1403. Pp. 85–99. Springer, Heidelberg (1998).*
- Young Adam and Moti Yung, 2005. "A space efficient backdoor in RSA and its applications." *International Workshop on Selected Areas in Cryptography. Springer, Berlin, Heidelberg, 2005.*
- Wu D. Introduction to cryptography. In: *Lecture Notes, Stanford University, <http://crypto.stanford.edu/~dwu4/notes/CS255LectureNotes.pdf>*

NEWS OF THE NATIONAL ACADEMY OF SCIENCES OF THE REPUBLIC OF KAZAKHSTAN
PHYSICO-MATHEMATICAL SERIES

ISSN 1991-346X

Volume 2. Number 346 (2023). 81–95

<https://doi.org/10.32014/2023.2518-1726.186>

УДК 004.056.53

© **K.B. Bagitova***, **Sh.Zh. Mussiraliyeva**, **M.A. Bolatbek**, **R.K. Ospanov**, 2023

Kazakh National University named after al-Farabi, Almaty, Kazakhstan.

E-mail: kbbagitova@gmail.com

DEVELOPMENT OF EXWEB SOFTWARE FOR DETECTING EXTREMIST CONTENT ON THE INTERNET

K. Bagitova — Lecturer of the Department of Information Systems of the Al-Farabi Kazakh National University, Almaty, Kazakhstan

E-mail: kbbagitova@gmail.com, <https://orcid.org/0000-0003-1587-1995>;

Sh. Mussiraliyeva — Candidate of Physical and Mathematical Sciences. Head of the Department of Information Systems of the Al-Farabi Kazakh National University, Almaty, Kazakhstan

E-mail: mussiraliyevash@gmail.com, <https://orcid.org/0000-0001-5794-3649>;

M. Bolatbek — Senior Lecturer. Department of Information Systems. Al-Farabi Kazakh National University, PhD, Almaty, Kazakhstan

E-mail: bolatbek.milana@gmail.com, <https://orcid.org/0000-0002-2153-180X>;

R. Ospanov — Lecturer of the Department of Information Systems of the Al-Farabi Kazakh National University. Almaty, Kazakhstan

E-mail: ospanov.ruslan.k@gmail.com, <https://orcid.org/0000-0001-8164-7057>.

Abstract. With the rise of social media, cyberterrorism has become an international problem threatening world peace. At a time when the Internet is becoming more widespread in all areas, interested users or organizations can use anonymity to terrorize citizens, communities, specific groups and entire countries, without physical contact. The analysis of social networks plays a key role in identifying various groups in the cyberterrorist network. Community discovery is one of the most important and interesting tasks in social network analysis. In general, we use machine learning methods to solve the problem of countering terrorism and extremism using information from the Internet. This task involves extracting messages from social networks, web resources potentially containing information of a terrorist or extremist nature, identifying the structure of user groups and Internet communities distributing this information, monitoring and modeling information flows in these communities, as well as threat assessment, risk prediction based on monitoring results. Within the framework of this article, the development of ExWeb software was considered. The architecture design was designed, including the development of the application structure, an analysis of the appropriate infrastructure for software placement was carried out. The front

and back-end parts of the web application were implemented. UML schemes were considered, programming languages HTML, CSS, JavaScript, and web frameworks Bootstrap and Django, JavaScript libraries such as jQuery, Chart were used.js and d3.js An experimental environment was developed to study the effectiveness of the developed models and algorithms.

Keywords: social network, social network analysis, cyberterrorism, influencers, machine learning

© **К.Б. Багитова***, **Ш.Ж. Мүсіралиева**, **М.А. Болатбек**, **Р.Қ. Оспанов**, 2023
Әл-Фараби атындағы Қазақ ұлттық университеті, Алматы, Қазақстан.
E-mail: kbbagitova@gmail.com

ИНТЕРНЕТТЕ ЭКСТРЕМИСТІК МАЗМҰНДЫ АНЫҚТАУҒА АРНАЛҒАН EXWEB БАҒДАРЛАМАЛЫҚ ЖАБДЫҚТАМАСЫН ӘЗІРЛЕУ

Багитова К.Б. — Әл-Фараби атындағы Қазақ Ұлттық университетінің Ақпараттық жүйелер кафедрасының оқытушысы. Алматы, Қазақстан

E-mail: kbbagitova@gmail.com, <https://orcid.org/0000-0003-1587-1995>;

Мүсіралиева Ш.Ж. — физика – математика ғылымдарының кандидаты, Әл-Фараби атындағы Қазақ Ұлттық университетінің Ақпараттық жүйелер кафедрасының меңгерушісі, Алматы, Қазақстан

E-mail: mussiraliyevash@gmail.com, <https://orcid.org/0000-0001-5794-3649>;

Болатбек М.А. — Әл-Фараби атындағы Қазақ Ұлттық университетінің Ақпараттық жүйелер кафедрасының аға оқытушысы, PhD, Алматы, Қазақстан

E-mail: bolatbek.milana@gmail.com, <https://orcid.org/0000-0002-2153-180X>;

Оспанов Р.Қ. — Әл-Фараби атындағы Қазақ Ұлттық университетінің Ақпараттық жүйелер кафедрасының оқытушысы. Алматы, Қазақстан

E-mail: ospanov.ruslan.k@gmail.com, <https://orcid.org/0000-0001-8164-7057>.

Аннотация. Әлеуметтік медианың өсуімен кибертерроризм әлемдік бейбітшілікке қауіп төндіретін халықаралық проблемаға айналды. Интернет барлық аймақтарда кең тараған кезде мүдделі пайдаланушылар немесе ұйымдар физикалық байланыссыз азаматтарды, қауымдастықтарды, белгілі бір топтарды және бүкіл елдерді қорқыту үшін анонимділікті пайдалана алады. Әлеуметтік желіні талдау кибертеррорлық желідегі әртүрлі топтарды анықтауда шешуші рөл атқарады. Қауымдастықты табу — әлеуметтік желіні талдаудағы ең маңызды және қызықты тапсырмалардың бірі. Жалпы, біз өз жұмысымызда интернеттен алынған ақпаратты пайдалана отырып, терроризм мен экстремизмге қарсы тұру мәселесін шешу үшін машиналық оқыту әдістерін қолданамыз. Бұл міндет террористік немесе экстремистік сипаттағы ақпаратты қамтитын әлеуметтік желілерден, веб-ресурстардан хабарламаларды алуды, осы ақпаратты тарататын пайдаланушылар топтары мен интернет-қауымдастықтардың құрылымын анықтауды, осы қауымдастықтардағы ақпараттық ағындарды бақылауды және модельдеуді,

сондай-ақ қауіп-қатерлерді бағалауды қамтиды, мониторинг нәтижелері бойынша тәуекелдерді болжау. Осы мақала аясында ExWeb бағдарламалық қамтамасыз етуді әзірлеу қарастырылды. Архитектураның дизайны әзірленді, оның ішінде қолданбалы құрылымды әзірлеу, бағдарламалық қамтамасыз етуді орналастыру үшін қолайлы инфрақұрылымға талдау жасалды. Веб-қосымшаның алдыңғы және серверлік бөліктерін іске асыру жүзеге асырылды. UML диаграммалары қарастырылды, HTML, CSS, JavaScript бағдарламалау тілдері және Bootstrap және Django веб фреймворктері, jQuery, Chart.js және d3.js сияқты JavaScript кітапханалары пайдаланылды. Жасалған модельдер мен алгоритмдердің тиімділігін зерттеу үшін эксперименталды орта әзірленді.

Түйін сөздер: әлеуметтік желі, әлеуметтік желіні талдау, кибертерроризм, ықпал етушілер, машиналық оқыту

© К.Б. Багитова*, Ш.Ж. Мусиралиева, М.А. Болатбек,
Р.К. Оспанов, 2023

Казахский национальный университет имени аль-Фараби,
Алматы, Казахстан.
E-mail: kbbagitova@gmail.com

РАЗРАБОТКА ПРОГРАММНОГО ОБЕСПЕЧЕНИЯ EXWEB ДЛЯ ВЫЯВЛЕНИЯ ЭКСТРЕМИСТСКОГО КОНТЕНТА В СЕТИ ИНТЕРНЕТ

Багитова К.Б. — преподаватель кафедры информационных систем Казахского национального университета имени аль-Фараби, г. Алматы, Казахстан

E-mail: kbbagitova@gmail.com, <https://orcid.org/0000-0003-1587-1995>;

Мусиралиева Ш.Ж. — кандидат физико-математических наук, заведующий кафедрой информационных систем Казахского национального университета имени аль-Фараби, г. Алматы, Казахстан

Алматы, Казахстан

E-mail: mussiraliyevash@gmail.com, <https://orcid.org/0000-0001-5794-3649>;

Болатбек М.А. — старший преподаватель кафедры информационных систем Казахского национального университета имени аль-Фараби, PhD, Алматы, Казахстан

E-mail: bolatbek.milana@gmail.com, <https://orcid.org/0000-0002-2153-180X>;

Оспанов Р.К. — преподаватель кафедры информационных систем Казахского национального университета имени аль-Фараби, г. Алматы, Казахстан

E-mail: ospanov.ruslan.k@gmail.com, <https://orcid.org/0000-0001-8164-7057>.

Аннотация. С ростом социальных сетей кибертерроризм стал международной проблемой, угрожающей миру. В то время, когда Интернет становится все более распространенным во всех областях, заинтересованные пользователи или организации могут использовать анонимность для терроризирования граждан, сообществ, конкретных групп и целых стран без физического контакта. Анализ социальных сетей играет ключевую роль в выявлении различных групп в кибертеррористической сети. Обнаружение сообществ — одна из самых важных и интересных задач в анализе

социальных сетей. В целом в работе мы применяем методы машинного обучения для решения задачи противодействия терроризму и экстремизму с использованием информации из сети Интернет. Эта задача предполагает извлечение сообщений из социальных сетей, веб-ресурсов, потенциально содержащих информацию террористического или экстремистского характера, выявление структуры групп пользователей и интернет-сообществ, распространяющих эту информацию, мониторинг и моделирование информационных потоков в этих сообществах, а также оценку угроз, прогнозирование рисков на основе результатов мониторинга. В рамках данной статьи была рассмотрена разработка ПО ExWeb. Был спроектирован дизайн архитектуры, в том числе разработка структуры приложения, был проведен анализ подходящей инфраструктуры для размещения ПО. Была проведена реализация лицевой и серверной части веб-приложения. Рассмотрены схемы UML, были использованы языки программирования HTML, CSS, JavaScript, и веб-фреймворки Bootstrap и Django, библиотеки JavaScript, такие как jQuery, Chart.js и d3.js. Была разработана экспериментальная среда для исследования эффективности разработанных моделей и алгоритмов.

Ключевые слова: социальная сеть, анализ социальных сетей, кибертерроризм, инфлюенсеры, машинное обучение

Работа выполнена в рамках проекта AP15473408 «Разработка моделей и методов выявления экстремистского контента в социальных сетях».

Введение

Террористические атаки по всему миру поставили под сомнение способность понимать явление радикализации и, что более важно, приобретать знания, необходимые для понимания и предотвращения таких актов. Интернет и социальные сети могут играть определенную роль в процессе насильственной радикализации, главным образом посредством распространения информации и пропаганды, а также укрепления, идентификации и вовлечения выбранной аудитории, которая заинтересована в радикальных и насильственных сообщениях. Исследования стали мощным инструментом для понимания радикализации, создания коллективного разума и принятия практических мер для обмена опытом по этой теме. Одной из таких областей накопления знаний является изучение роли Интернета и, в частности, социальных сетей в процессах насильственной радикализации и терроризма.

Террористическая сеть представляет серьезную опасность для различных пользователей социальных сетей, в основном для организации террористических актов и добавления новых участников в свои сообщества, особенно привлекая молодежь. Таким образом, наиболее полезным используемым методом является анализ социальных сетей (SNA) для анализа социальных сетей и обнаружения скрытой структуры; разделения сети на группы. Обычно социальная сеть представлена в виде графика, где графики

дают интуитивное представление о структуре сети. SNA может использоваться по-разному в социальной сети, например:

- Анализ информационного потока.
- Отслеживание распространения информации и определение пройденного пути.
- Определение правильного разделения для социальной сети.
- Определение отношений между субъектами (прямых или косвенных).

Выявление проблемного контента является сложной и приоритетной задачей для любого государственного учреждения. Целью мог бы стать автоматизированный поиск и классификация экстремистского контента, с использованием машинного обучения и искусственного интеллекта (Chen, 2011). В статье анализируются исламистские пропагандистские материалы (тексты, видео), посты и общение в чате, на основе, которых были разработаны категории и алгоритмы.)

Акинборо и др. рассматривают проблему унижительных и оскорбительных постов на платформах социальных сетей и обсуждают методы, которые могут быть использованы для обнаружения этого контента (Akinboro и др., 2021). Подходы, описанные в этом исследовании, включают обработку естественного языка (NLP), подход глубокого обучения, многоуровневую классификацию, гибридные подходы и подходы для распознавания многоязычных контекстов, которые становятся все более важными.

Ошибочно полагать, что «террористы» психически больны, т.е. «сумасшедшие». Почти все исследования по радикализации и психопатологии показали, что террористы — за исключением так называемых "одиноких волков" — психически ненормальны не больше и не меньше, чем остальная часть населения. Почти всегда ошибочными являются предположения, объясняющие радикализацию одной-единственной причиной. «Не все террористы сироты, необразованны или бедны, и не всегда они происходят из многодетных семей, имеют докторскую степень или богатых родителей» (Neumann, 2016).

Существует множество форм насильственной радикализации (Types de Radicalisation, 2021), такие как ультраправые (связанные с фашистскими, расистскими/расистскими, ультранационалистическими мотивами), политико-религиозные (связанные с политическим прочтением религии и защитой религиозной идентичности, какой бы ни была религия), ультра-левые (артикулированные вокруг антикапитализма и трансформации политической системы, воспринимаемой как генератор социального неравенства), уникальные причины (например, экологические, права животных, противники абортов, гомофобии, антифеминистки и т.д.). Существуют также различные средства массовой информации, в которых радикальные люди выражают себя: социальные сети, веб 2.0, определенные газеты, чаты, форумы на разных языках или диалектах (каждое средство массовой информации также имеет

свою особую форму выражения, например, твиты или газеты). Интернет и социальные сети привлекательны для экстремистских групп по разным причинам (MDR.DE., 2021, Holbrook, 2015):

- Интернет и социальные сети представляют собой идеальную возможность для самовыражения и общения.

- Интернет и социальные сети могут за очень короткое время охватить миллионы адресатов по всему миру.

- Интернет и социальные сети предлагают идеальную возможность для общения с единомышленниками.

- В Интернете и социальных сетях социальный контроль не только со стороны органов безопасности, но и со стороны социальной среды значительно затруднен.

Цель данной работы — разработка ПО ExWeb, проведение комплексного исследования и разработка алгоритмов идентификации вовлеченных пользователей и методов графической визуализации связей. Для достижения поставленных целей были определены следующие задачи:

- анализ контента социальных сетей Вконтакте, Youtube, Twitter с помощью предварительно обученной модели машинного обучения,

- определение связи между пользователями сети и их связи с определенными группами. Диаграмма последовательности анализа контента приведена на рисунке 1,

- определение лидера сообщества.

Дизайн архитектуры

Архитектура системы описывает ее основные компоненты, их отношения (структуры) и то, как они взаимодействуют друг с другом. Архитектура и дизайн программного обеспечения включают в себя несколько факторов, таких как бизнес-стратегия, атрибуты качества, динамика человека, дизайн и ИТ-среда. По проекту нам нужно создать программное обеспечение, которое анализирует социальные сети. Разработка веб-приложения состоит из нескольких этапов:

- определение целей и задач проекта (создание технического задания),

- разработка структуры приложения,

- создание макетов дизайна,

- верстка,

- программирование и разработка логики приложения,

- тестирование и отладка.

На этапе проектирования формулируются бизнес-цели создаваемого проекта, определяются требования, которым он должен удовлетворять, разрабатывается общая концепция. Погрешности здесь недопустимы, так как они могут проявиться на любом этапе и в итоге отбросить проект далеко назад.

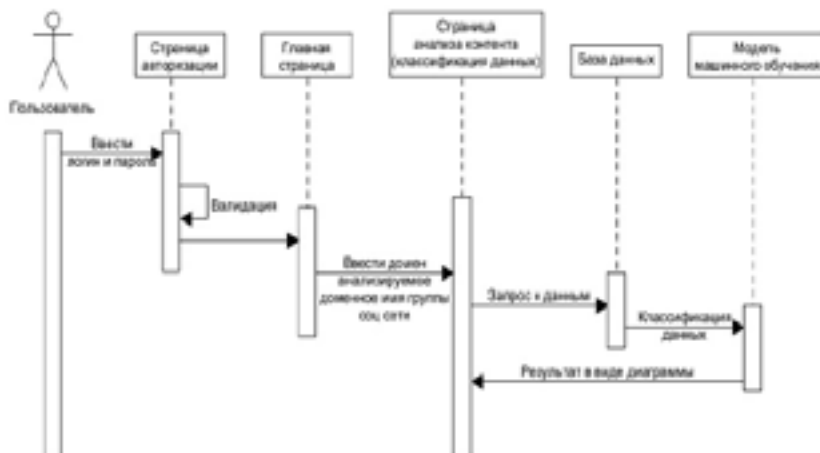


Рисунок 1 – Диаграмма последовательностей анализа контента

В UML схемах чаще всего встречаются следующие основные типы обозначений отношений между классами:

ассоциация показывает, что объекты одной сущности (класса) связаны с объектами другой сущности таким образом, что можно перемещаться от объектов одного класса к другому,

реализация — отношение между двумя элементами модели, в котором один элемент реализует поведение, заданное другим,

агрегация — это разновидность ассоциации при отношении между целым и его частями,

композиция имеет жёсткую зависимость времени существования экземпляров класса контейнера и экземпляров содержащихся классов. Если контейнер будет уничтожен, то всё его содержимое будет также уничтожено. Диаграмма классов социальной сети Vkontakte приведена на рисунке 2.

Разработка структуры приложения включает всё, что касается его содержания и информационной стратегии, определяющей, как должна быть организована подача информации, чтобы будущие посетители приложения могли быстро и легко её найти. Первоочередной задачей на данном этапе является создание карты приложения, отражающей взаимосвязи типовых страниц и их наиболее значимые функциональные возможности.

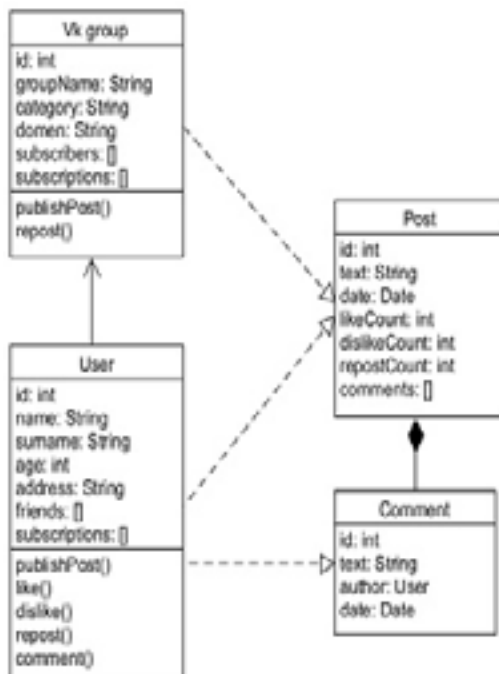


Рисунок 2 – Диаграмма классов социальной сети Vkontakte

Разработка структуры приложения

На главной странице размещена общая информация о запрещенных в стране религиозных организациях и казахстанцах, которые покинули страну и пошли на войну в Сирии и Ираке (рисунок 3).

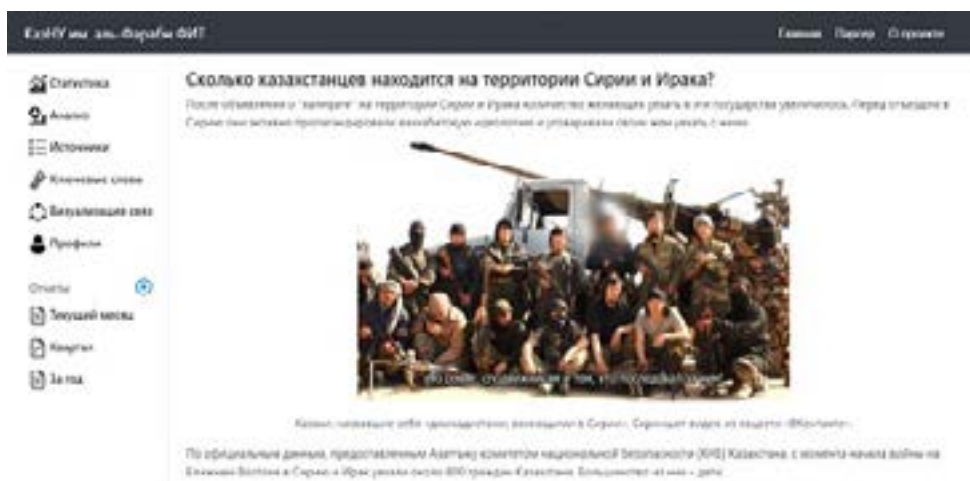


Рисунок 3 – Главная страница приложения

Мы можем анализировать сообщества и группы Вконтакте на странице анализа. Заполняем форму следующими данными: название группы и период, в течение которого были опубликованы записи. Отправляем форму запроса. Сервер принимает запрос клиента по указанному URL-адресу и передает его соответствующей функции-обработчику для обработки запроса. Функция-обработчик обращается к базе данных и пытается получить записи с параметрами (имя группы и дата публикации), указанными в запросе. Полученные данные предварительно обрабатываются для классификации текста. Используя предварительно обученную модель, функция классифицирует записи как экстремистские и неэкстремистские. И вернет процент экстремистского контента (рисунок 4).

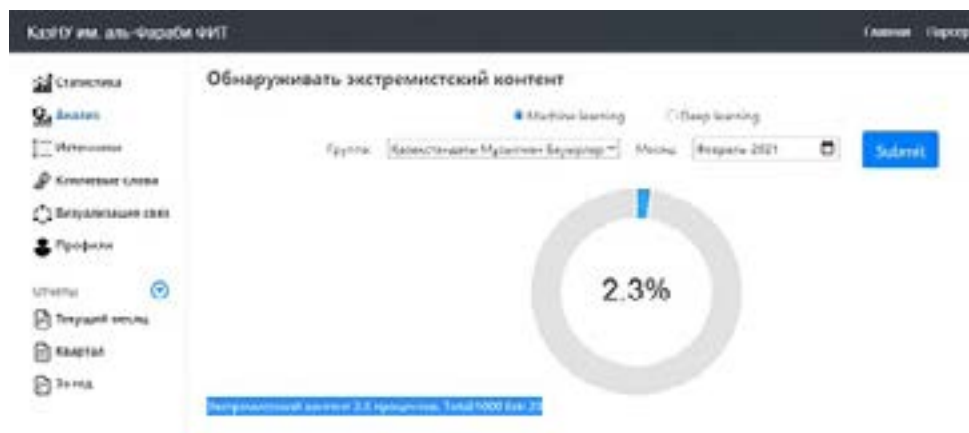


Рисунок 4 – Демонстрация анализа сообществ и групп Вконтакте

Визуализация графа была выполнена в формате веб-приложения. Веб-приложение использует html, css и javascript для отрисовки графа. Для визуализации была использована библиотека d3.js. В качестве формата отображения графа был выбран svg, так как он позволяет приближать и отдалять изображение во сколько угодно раз, при этом не теряя в качестве. В качестве бэкэнда веб-приложения используется фреймворк Django на языке Python.

На html-странице был размещен элемент svg, в котором затем отрисовывался граф с помощью скрипта javascript с использованием d3.js. Еще на этапе формирования html-страницы Django передает необходимые данные, которые мы сформировали после получения данных. Затем каждый элемент заполняется соответствующими ему данными и можно задать для элементов дополнительные параметры. Например, можно задать радиус вершин графа или толщину ребер графа. В результате этих этапов можно будет построить визуализацию связей пользователей (рисунок 5).

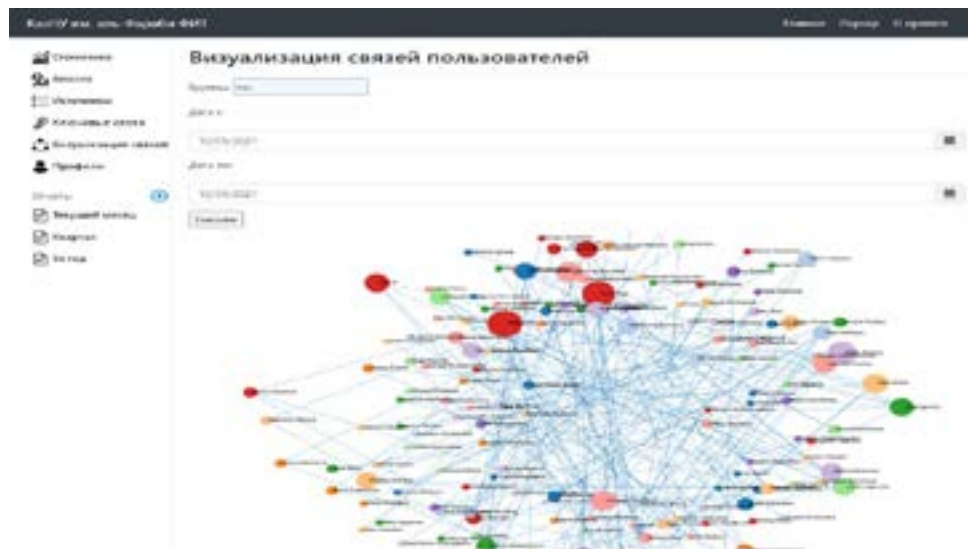


Рисунок 5 – Визуализация связей пользователей

Сбор данных. При вводе списка групп через запятую и указании временного отрезка, в котором содержатся нужные посты из этих групп, можно получить граф постов. Здесь при сборе данных собирались посты указанных группы в определенном временном промежутке, а также информация о пользователях, оставивших комментарии или лайки. В частности, для каждой группы в списке запускался метод `get_posts`. После получения всех постов, для каждого поста получают все его комментарии или лайки методом `get_post_weights`. Затем на основе этих данных строится граф. Планируется использование базы данных для того, чтобы запрос визуализации графа не занимал много времени при повторном вызове.

Составление графа из данных. Пост здесь является вершиной, а суммарное количество комментариев или лайков под ним его весом. Весом ребер между вершинами является количество одинаковых пользователей у постов, которые лайкнули запись или оставили комментарий. Соответственно, если общих пользователей у постов нету, то ребра не будет (рисунок 6).

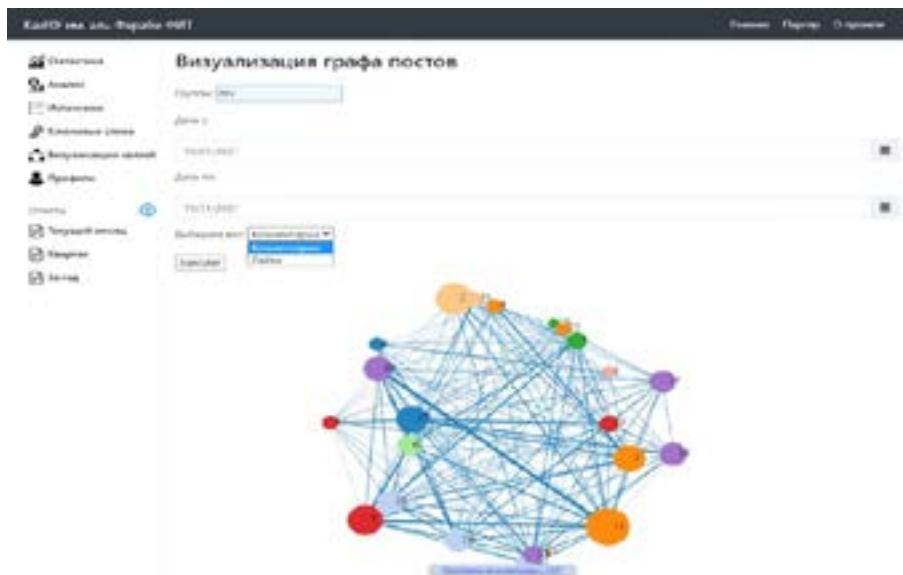


Рисунок 6 – Визуализация графа постов

В разделе статистики отображаются такие метрики, как WordCloud, биграммы, триграммы, tf-idf и liwc нашего корпуса. Word Cloud — это ресурс, позволяющий создать визуальный образ ключевых слов, текста в привлекательной форме. Наиболее часто используемые слова отображаются крупным шрифтом. Важность каждого ключевого слова обозначается размером шрифта или цветом. Такое представление удобно для быстрого восприятия (рисунок 7).

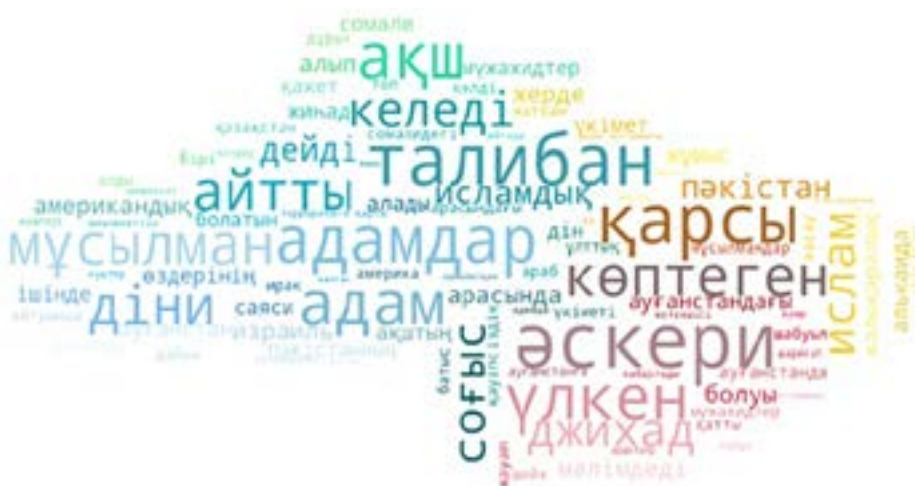


Рисунок 7 – Облако слов обучающего корпуса

N-грамма — это просто последовательность из n элементов, идущих в каком-то тексте подряд. На практике чаще имеют в виду ряд слов (реже — символов). Последовательность из двух элементов называют биграмма, из трёх элементов — триграмма. Вычислив частоту вхождения N-грамм в текстах корпуса, мы можем добавить биграммы или триграммы в качестве функций для представления наших документов в задачах классификации текста (рисунки 8–9).

#	Уникальность	Выявлено	Триграмма
1	карты	779	террористиче карты
2	море	509	болды море
3	ислам	495	длин исламиче
4	америк	440	ауланган ислам
5	болды	422	америка кулма
6	амер	380	Будариче америче
7	длин	358	болды миде
8	болды	347	ислам америче
9	болды	344	карты туре
10	болды	337	слова болды
11	таблиц	336	карты амер
12	амер	302	амер амер
13	таблиц	296	ислам америче
14	таблиц	282	ауланган болды
15	таблиц	271	америче террористиче

Рисунок 8 – Таблица биграмм и триграмм



Рисунок 9 – Демонстрация биграмм и триграмм

TF-IDF (от англ. term frequency и inverse document frequency — частота слова и обратная частота документа) – это показатель, который используется для оценки важности слова в документе.

TF (частота слов) характеризует отношение числа вхождений конкретного слова к общему набору слов в документе. Чем выше TF, тем весомее конкретное слово в рамках документа.

IDF (обратная частота документа) характеризует инверсию частотности, с которой конкретное слово используется в тексте. С помощью этой метрики можно снизить важность слов, например, союзов или предлогов (рисунок 10).

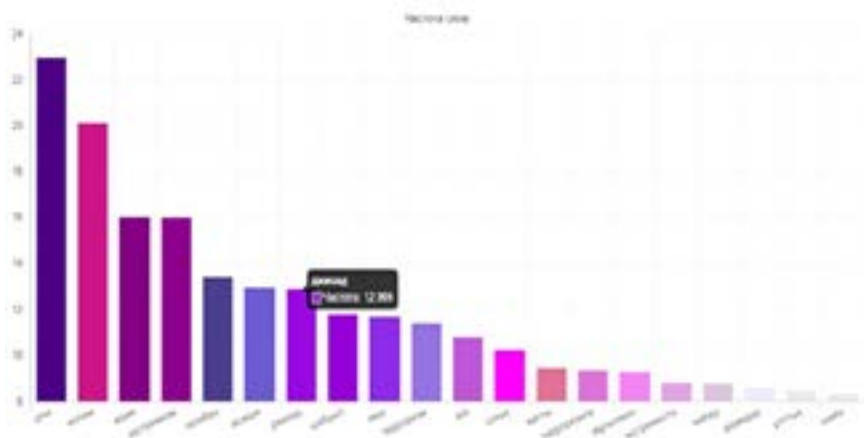


Рисунок 10 – TF-IDF показатель корпуса

Визуализация была выполнена в формате веб-приложения. Веб-приложение использует html, css и javascript для отрисовки графа. Для визуализации была использована библиотека d3.js. В качестве формата отображения графа был выбран svg, так как он позволяет приближать и отдалять изображение во сколько угодно раз, при этом не теряя в качестве (рисунки 11).

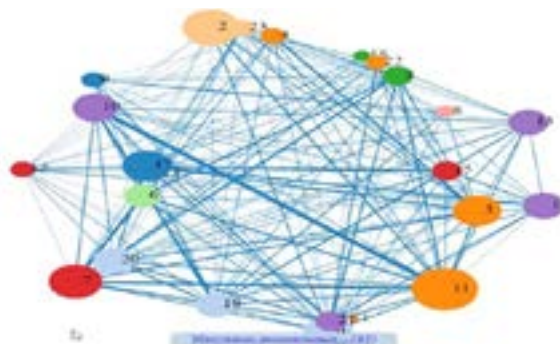


Рисунок 11 – Визуализация графа постов

Сервер и лицевая часть взаимодействуют друг с другом по протоколу HTTP. HTTP расшифровывается как протокол передачи гипертекста и используется для структурирования запросов и ответов через Интернет. Протокол HTTP требует передачи данных из одной точки в другую по сети. Передача ресурсов происходит с использованием TCP. TCP управляет каналами между браузером и сервером. Протокол TCP используется для управления многими типами интернет-соединений, при которых один компьютер или устройство хочет отправить что-либо другому.

Веб-приложению нужна база данных для управления всей информацией. База данных хранит контент веб-сайта в структуре, которая упрощает извлечение, организацию, редактирование и сохранение данных. Он работает на удаленном компьютере, который называется сервером. Существует множество широко используемых баз данных, таких как MySQL, SQL Server, PostgreSQL и Oracle. Мы используем базу данных PostgreSQL. Структура использованной базы данных (рисунок 12):

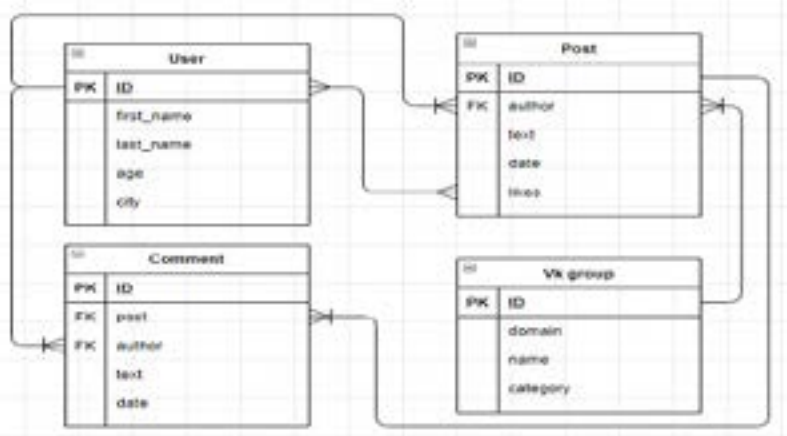


Рисунок 12 – Схема базы данных

где $\text{---}\{\} \text{---}$ обозначает отношения один ко многим, а $\text{---}\}\{\text{---}$ отношения многие ко многим.

Заключение

Рассмотрены этапы разработки ПО ExWeb, предназначенной для выполнения мониторинга социальных сетей, таких как Вконтакте, Youtube, Twitter с целью обнаружения экстремистского контента, был спроектирован дизайн архитектуры, в том числе разработка структуры приложения, показана реализация поставленной задачи. При этом используются методы машинного обучения для решения задачи противодействия терроризму и экстремизму с использованием информации из сети Интернет. В ПО реализованы алгоритмы графической визуализации связей в сообществах.

REFERENCES

Chen H., 2011. *Dark Web: Exploring and Data Mining the Dark Side of the Web*; Springer Science & Business Media: Berlin, Germany, 2011. Volume 30. [Google Scholar]

Akinboro S., Adebusoye O., Onamade A.. 2021. A Review on the Detection of Offensive Content in Social Media Platforms. *FUOYE J. Eng. Technol.* 2021, 6. [Google Scholar] [CrossRef]

Neumann P.R., 2016. *Der Terror ist unter uns: Dschihadismus, Radikalisierung und Terrorismus in Europa*; Ullstein eBooks: Berlin, Germany, 2016. [Google Scholar]

Types de Radicalisation. Available online: <https://info-radical.org/fr/types-de-radicalisation/> (accessed on 7 November 2021).

MDR.DE. Wie das Internet zur Radikalisierung Beiträgt|MDR.DE. Available online: <https://www.mdr.de/wissen/bildung/extremismus-internet-online-radikalisierung-100.html> (accessed on 7 November 2021).

Holbrook D., 2015. A critical analysis of the role of the internet in the preparation and planning of acts of terrorism. *Dyn. Asymmetric Confl.* 2015, 8. Pp. 121–133. [Google Scholar] [CrossRef] [Green Version]

NEWS OF THE NATIONAL ACADEMY OF SCIENCES OF THE REPUBLIC OF
KAZAKHSTAN
PHYSICO-MATHEMATICAL SERIES
ISSN 1991-346X
Volume 2. Number 346 (2023). 95–107
<https://doi.org/10.32014/2023.2518-1726.187>

UDC 543; 543.62

© **A.Sh. Barakova***, **O.A.Usatova**, **A.S.Orynbaeva**, 2023
Al-Farabi Kazakh National University, Almaty, Kazakhstan.
E-mail: balia_79@mail.ru

DIGITAL RESOURCES ON WEBSITES MODEL OF PROTECTION BY STEGANOGRAPHY

Barakova Aliya Sharizatovna — PhD student, Al-Farabi Kazakh National University. Almaty, Kazakhstan

E-mail: balia_79@mail.ru. ORCID: <https://orcid.org/0000-0002-0904-745X>;

Ussatova Olga Aleksandrovna — PhD, Gumarbek Daukeev Almaty University of energy and communications,
Almaty, Kazakhstan

E-mail: uoa_olga@mail.ru. ORCID: <https://orcid.org/0000-0002-5276-6118>;

Orynbaeva Ainur Serikovna — Senior Lecturer, Astana Medical University. Astana, Kazakhstan
E-mail: ainur_tas@mail.ru.

Abstract. This article provides an overview of ways and methods to prevent pirates from copying, processing, distributing and making digital resources public, and using digital creations for personal gain. To improve the quality of research and analysis, the steganography method is set as the goal. Finally, this review contains discussion, Research, Critical Analysis and a detailed summary, so that you can easily understand the development of steganography of media information such as image, video, audio. Steganography is a method of hiding data in a file, such as an image, audio, video, etc. and this data is reliably transmitted from the sender to the recipient. This method serves as a better way to protect the message than cryptography, which hides the content of the message. The original message is hidden so that the changes that occurred in the file are not noticed. To hide confidential information, you can use reliable steganography methods that are more complex than others, they all have their strengths and weaknesses. The absolute invisibility of confidential information is supported by various applications, while others require a large secret message to be hidden. This article discusses an overview of video, image, audio steganography, its use and methods to meet the need for privacy on the internet.

Key words: steganography, cryptography, confidential information, digital image, video, audio

© А.Ш. Баракова*, О.А. Усатова, А.С. Орынбаева, 2023
эл-Фараби атындағы Қазақ ұлттық университеті, Алматы, Қазақстан.
E-mail: balia_79@mail.ru

ВЕБ САЙТТАРДАҒЫ САНДЫҚ РЕСУРСТАРДЫ СТЕГАНОГРАФИЯ ӘДІСІМЕН ҚОРҒАУДЫҢ МОДЕЛІ

Баракова Алия Шаризатқызы — PhD докторанты, эл-Фараби атындағы Қазақ ұлттық университеті. Алматы, Қазақстан

E-mail: balia_79@mail.ru. ORCID: <https://orcid.org/0000-0002-0904-745X>;

Усатова Ольга Александрқызы — PhD, Ғұмарбек Дәукеев атындағы Алматы энергетика және байланыс университеті, Алматы, Қазақстан

E-mail: uoa_olga@mail.ru. ORCID: <https://orcid.org/0000-0002-5276-6118>;

Орынбаева Айнур Серикқызы — Аға оқытушы, Астана медицина университеті. Астана, Қазақстан

E-mail: ainur_tas@mail.ru.

Аннотация. Бұл мақалада қарақшылардың сандық ресурстарды көшіру, өңдеу, тарату және көпшілікке жария ету, сандық туындыларды жеке пайда табу үшін пайдаланудың алдын алу жолдары мен әдістеріне шолу жасалады. Зерттеу мен талдаудың сапасын жақсарту үшін стеганография әдісін алға мақсат етіп қойылады. Соңында, бұл шолуда талқылау, зерттеу, сыни талдау және егжей-тегжейлі түйіндеме бар, осылайша сурет, видео, аудио сияқты меди ақпараттардың стеганографиясының дамуын оңай түсінуге болады. Стеганография — бұл файлдағы деректерді жасыру әдісі, мысалы, сурет, аудио, видео және т.б. және бұл деректер жіберушіден алушыға сенімді түрде беріледі. Бұл әдіс хабарламаның мазмұнын жасыратын криптографияға қарағанда хабарламаны қорғаудың жақсы әдісі ретінде қызмет етеді. Бастапқы хабарлама файлда болған өзгерістер байқалмайтындай етіп жасырылады. Құпия ақпаратты жасыру үшін басқаларға қарағанда күрделі болып табылатын сенімді стеганография әдістерін қолдануға болады, олардың барлығының күшті және әлсіз жақтары бар. Құпия ақпараттың абсолютті көрінбейтіндігіне әртүрлі қосымшалар қолдау көрсетеді, ал басқалары үлкен құпия хабарламаны жасыруды талап етеді. Бұл мақалада видео, сурет, аудио стеганографиясына шолу, оны пайдалану және интернеттегі құпиялылық қажеттілігін қанағаттандыру әдістері қарастырылады.

Түйін сөздер: стеганография, криптография, құпия ақпарат, сандық сурет, видео, аудио

Мүдделер қақтығысы: Авторлар осы мақалада мүдделер қақтығысы жоқ деп мәлімдемейді.

© А.Ш. Баракова *, О.А. Усатова, А.С. Орынбаева, 2023

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

E-mail: balia_79@mail.ru

РАЗРАБОТКА МОДЕЛИ ЗАЩИТЫ ЦИФРОВЫХ WEB РЕСУРСОВ С ИСПОЛЬЗОВАНИЕМ МЕТОДОВ СТЕГАНОГРАФИИ

Баракова Алия Шаризатовна — докторант PhD, Казахский национальный университет им. Аль-Фараби. Алматы, Казахстан

E-mail: balia_79@mail.ru. ORCID: <https://orcid.org/0000-0002-0904-745X>;

Усатова Ольга Александровна — PhD, Алматинский университет энергетики и связи имени Гумарбека Даукеева, Алматы, Казахстан

E-mail: uoa_olga@mail.ru. ORCID: <https://orcid.org/0000-0002-5276-6118>;

Орынбаева Айнур Сериковна — старший преподаватель, медицинский университет Астана. Астана, Казахстан

E-mail: ainur_tas@mail.ru.

Аннотация. В данной статье представлен обзор способов и средств предотвращения от атак злоумышленника: копирования, редактирования, распространения и публикации цифровых ресурсов общественности, а также предотвращение использования цифрового контента в личных целях. Для повышения качества исследования и анализа в качестве цели выдвигается метод стеганографии. Данный обзор включает в себя обсуждение, исследование, критический анализ и подробный обзор, чтобы можно было легко понять развитие стеганографии для медиаинформации, такой как изображения, видео и аудио. Стеганография — это метод сокрытия данных в файле, таких как изображение, аудио-, видео-, которые безопасно передаются от отправителя к получателю. Данный метод является наиболее защищенным для передачи сообщения, чем криптография, которая кодирует информацию. Изменения в файле не доступны, так как исходное сообщение скрыто. Для сокрытия конфиденциальной информации могут использоваться разные методы стеганографии, которые имеют сильные и слабые стороны. Абсолютная невидимость конфиденциальной информации поддерживается различными приложениями, в то время как другие требуют сокрытия большого секретного сообщения. В этой статье представлен обзор методов стеганографии для защиты конфиденциальной информации в Интернете таких файлов, как видео, аудио и изображений.

Ключевые слова: стеганография, криптография, конфиденциальная информация, цифровое изображение, видео, аудио

Кіріспе

Цифрлық технологиялар мен интернеттің дамуымен мультимедиялық және цифрлық туындылар кескіндер, аудио, видео және басқа да цифрлық форматтар түрінде интернетте жарияланады және оларға авторлық құқықты

қорғау және ақпараттың тұтастығын қамтамасыз ету біртіндеп шешілуі керек өзекті мәселеге айналуға (Усатова, 2022). Сандық туындыларды көшіру, өңдеу, тарату және көпшілікке жария ету оңай болғандықтан, қарақшылар цифрлық туындылардың осы сипаттамаларын жеке пайда табу үшін авторлық құқық иелерінің заңды құқықтарын бұзу үшін пайдаланады.

Біз өз зерттеуімізде стеганографиялық әдістердің көбеюімен қатар, мақалаларда сурет, видео т.б. медиа ақпараттың кеңістіктік аймақтың стеганографиясына көбірек бағытталғанын байқадық. Әдебиеттік шолу жалпы стеганографияны жан-жақты талқылайтын әр әдісті енгізу процесіне бағытталған. Әдістерді ендіру процесінде ең көрнекті идея негізінде бірнеше санатқа бөлуге болады; сондықтан жаңа классификация ұсынылады. Ұсынылған жіктеу аудиостеганографияда ең көп қолданылатын тәсілдерді жалпылауға және түсінуге мүмкіндік береді. Содан кейін бұл зерттеу әрбір әдістің күшті және әлсіз жақтарын оның іске асыру мінез-құлқына негізделген сыни тұрғыдан қарастырады. Ендіру, медиа түрлері және бағалау орталары саласындағы басқа мәселелер де талқыланады.

Бірнеше жаңа әдістер маңызды аспект ретінде сенімділікті арттырады. Бұл стеганографиялық әдістердің дамуын әр түрлі етеді және әдетте әдістің мақсатын бір немесе екі аспектке бағыттайды. Қазіргі уақытта стеганография туралы мақалалар, әсіресе шолулар, стеганографияны оның мақсаттарына қарай сирек жіктейді. Осылайша, бұл шолулар стеганографиялық зерттеулердің мақсаттарына ықпал етеді. Зерттеу жұмысымызды деректерді жинаудан бастау алды. Ол әлемдік деңгейдегі бес дерекқорды қамтиды, атап айтқанда Web of Science, IEEE Xplore, ScienceDirect, Scopus және Springer, нәтижесінде барлығы 57 мақала қарастырылды. Әдебиеттік шолуда әр әдісті енгізу процесіне бағытталған яғни тәжірбиеден өткізілген. Әдістерді ендіру процесін бірнеше санатқа бөлуге болады. Стеганография әдісін зерттеу барысында сурет, видео, аудио т.б. медиа ақпараттарды қорғауға арналған бірнеше авторлардың жұмыстарын қарастырайық. Олар ұсынған ең көп қолданылатын тәсілдерді жалпылауға және түсінуге мүмкіндік береді. Содан кейін бұл зерттеу әрбір әдістің күшті және әлсіз жақтарын оның іске асыру мінез-құлқына негізделген сыни тұрғыдан қарастырады.

Әдебиеттік шолу

J. Bai тағы басқалары (Jiawang, 2022) өз зерттеу жұмыстарында авторлар видео деректердің қорғаудың жаңа әдісін ұсынған. Universal attack шабуылдарының әсерінен болатын зиянды видео деректерін жинауын болдырмау үшін видеоны терең хэштеу әдісін әзірледі. Бұл әдісті сценарийлерге деректерсіз таратады яғни видеоларды түпнұсқасыз да қорғауды қамтамасыз етеді. Әдістерін жүзеге асыруда эксперименттер жасап, екі деректер жиынтығы қолданған оның бірі ақ жәшік параметрлерінде де, қара жәшік параметрлерінде қарастырған. Бұл әдіс интернетте бөлісуге болатын видеоның құпиялылығын қорғайтындығына ерекше назар аударған.

Трапезников Е.В. (Трапезников, 2023) өз еңбегінде видео мазмұнын ашық желілерде қорғаудың бір әдісі ретінде логикалық жинақтауды қолданған, кездейсоқ бөліктерде видео деректеріне су таңбаларын енгізу әдісі ұсынылған. Ұсынылған әдіс видео ақпаратты бұрмалаудан қорғауға арналған бағдарламалық өнім деп сипаттаған. Бағдарламалық өнім C # тілінде жүзеге асырылады.

Ala Abu-Zahra (Ala Abu-Zahra, 2012) тағы басқалары өз жұмыстарында SAVLC және CABAC негізіндегі SVC-де әртүрлі рұқсаттарға тәуелсіз қол жеткізуге арналған жаңа платформа ұсынған. Ұсынылған схемада қабат аралық болжамды қолданған. Шифрлау схемасында нақты уақыттағы гетерогенді медиа ағынының сценарийлері үшін қолайлы деп есептейді. Авторлар эксперимент кезінде бірдей рұқсаттарды беру арқылы әрбір рұқсатты дербес қорғауды қамтамасыз ете алатынын көрсеткен.

Аудио-стеганография саласында көптеген шолу мақалалары жарияланған. А. Das өз әріптестерімен (Das, 2012) стеганографияны бұқаралық ақпарат құралдарының бірнеше түріне бағыттаған. Мысалы, мәтін, сурет, аудио және видео сияқты әртүрлі медиа түрлеріндегі стеганография әдістеріне шолу жасап, соның ішінде LSB кодтау, спектрді тарату және жаңғырықты жасыру, онда пакет аймағындағы деректерді жасырудың бірнеше әдістерін ұсынған. Зерттеуде стеганография және суреттерді, аудионы және HTML беттерін стеганоализдеу әдістері қарастырылған.

Trivikram M. Басқалары (Assaf Cohena, 2022) зерттеулерінде кескін, сурет файлдарының кең қолданылуына және көлемінің салыстырмалы айырмашылықтарына байланысты кескінді стеганографиялаудың заманауи әдістеріне бағытталған. Бұл мақалада сандық суреттерге арналған стеганография саласы мен стеганография әдістері қарастырылған. Соңғы жылдары стеганография басты назарда болды, өйткені стеганографиямен байланысты кибершпиондық шабуылдар көбейе бастады. Сурет стеганографиясының мазмұнын (стеганографияның кез келген түрін) анықтау үшін жанама арнаға шабуыл жасау әлдеқайда көп эксперименттер мен инновацияларды қажет етеді және зерттеудің тағы бір айқын бағыты болып табылады.

Әдістер мен материалдар

Адамзат тарихының барлық уақыттарында ақпаратты бұрмаланудан және рұқсатсыз қол жеткізуден қорғау мәселесі өзекті болып табылады (Коржик, 2017). Ежелгі әлемде бұл мәселені шешудің екі негізгі бағыты ерекшеленді, олар: криптография және стеганография. Стеганография әдетте криптографиямен қатар деректерді қорғау әдістері ретінде жіктеледі. Криптография құпия хабарламаны мағынасыз және оқылмайтын түрге айналдырады, сондықтан оны шабуылдаушылар түсіне алмайды, ал стеганография құпия хабарламалар мен олардың жазықсыз ортада болуын құпия түрде жасырады. Стеганография деректерді жасырын түрде жіберу

әдістері ғана емес (классикалық стеганография деп аталады), сонымен қатар шуылға төзімді аутентификация, ақпаратты рұқсатсыз көшіруден қорғау, байланыс желілері арқылы ақпараттың таралуын бақылау, мультимедиялық мәліметтер базасында ақпаратты іздеу және т.б. Жағдайларда дәстүрлі қолданыстағы ақпараттық ағындар немесе ақпараттық орта шеңберіне мүмкіндік береді және әр саланың р қолданбалы ақпаратын қорғаудың маңызды мәселелерін шешеді (Шелухин, 2018).

Мультимедиялық технологиялардың танымалдығы MP3, MPEG-4, JPEG2000 стандарттарында пайдалану үшін алгоритмдерін әзірлеуге, DVD дискілерін көшіруден қорғауға байланысты көптеген зерттеулер туындады. Стеганография технологиясын қолданудың негізгі бағыттарын төрт топқа біріктіруге болады: көшіруден қорғау (пайдалану), құжаттардың жасырын аннотациясы, ақпараттың түпнұсқалығын дәлелдеу және жасырын байланыс



Сур.1. Стеганографияны қолдану салалары (Fig.1. Areas of application of steganography)

Стегожүйенің математикалық моделін қарастырайық

Стегожүйені байланыс жүйесі ретінде ССБ (Сандық су белгісі).

Цифрлық стеганография алгоритмін ендіру үш негізгі кезеңнен тұрады:

- 1) ССБ генерацияланады,
- 2) кодерге ССБ ендіру
- 3) детекторда ССБ анықтау

ССБ мүмкін болатын W^* , K^* , I^* , B^* кілттер болсын, сәйкесінше контейнерлер мен жасырын хабарламаларымен. Онда ССБ генерациялау мына түрінде ұсынылуы мүмкін

$$F: I^*K \times B^* \rightarrow W^*, \quad W = F(I, K, B),$$

Мұндағы W , K , I , B - тиісті жиындардың өкілдері. Жалпы айтқанда, функция F ерікті болуы мүмкін, бірақ іс жүзінде қойлатын талаптарда шектеу бар.

Сонымен көп жағдайда $F(I, K, B) \approx F(I + \varepsilon, K, B)$ яғни шамалы

өзгертілген контейнер ССБ өзгеруіне әкелмейді. F функциясы әдетте құрама болып табылады:

$$F = T \circ G, \text{ мұнда } G: K^* \times B^* \rightarrow C^* \text{ және } T: C^* \times I^* \rightarrow W$$

яғни, ССБ контейнердің қасиеттеріне байланысты. Функция G бастапқы мәндер ретінде қауіпсіз жалған кездейсоқ реттілік генераторының көмегімен жүзеге асырылуы мүмкін.

T операторы C* кодтық сөзін өзгерте отырып, нәтижесінде W* ССБ алады. Бұл функцияға шектеу қоймасақта болады, себебі G кепіл береді F функциясы қайтымсыз екеніне. T функциясы таңдалуы керек толтырылмаған I₀ контейнер, толтырылған I_w контейнер және шамалы I_w толтырылған контейнерлер бірдей нәрсені тудырады.

$$T(C, I_0) = T(C, I_w) = T(C, I'_w)$$

Нәтижелер және оларды талқылау

Видео файлды өңдеу үшін avifil32.dll функциялар қолданылады. Бұл тәсілде C # қабықшалары қол жетімділік функциялары үшін қолданылады әрі avifil32.dll -де қол жетімді бола алады. Жүйені құру үшін 4 модуль қолданылады:

- Видео кадрларды бөлу және генерациялау технологиясы
- Байт-код фабрикасы
- Криптографиялық зауыт
- Стеганография фабрикасы



Сур.2. Стеганография фабрикасының сұлбасы
(Fig.2. outline of Steganography factory)

Зерттеу жүргізу үшін деректердің 4 түрі қажет болады:
Кілт файлы;
Құпия сөз;

Медиа;

Хабарлама.

Медиа файл өлшемі қолданылатын әдіске байланысты пропорцияда хабарлама өлшемінен асып кетуі керек.

Құпия сөз: пароль хабарламаны медиа файлға енгізер алдында оны шифрлау үшін қолданылған.

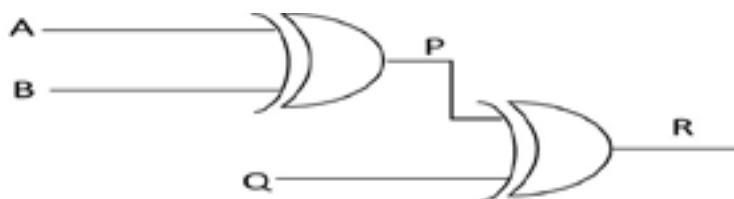
Кілт файлы: кілт файлы жасырын хабарламаның қауіпсіздігін жақсарту үшін қолданылады. Жеке парольдерді әртүрлі тәсілдермен табуға болады. Кілт файлы құпия сөзді болжауға жол бермейді, сондықтан әдіс қауіпсіздігін арттырады.

Тасымалдаушы Файл: бұл зерттеуде медиа файл ретінде видео файл қолданылады. Видео файл пішімі - AVI.

Хабарлама: Хабарлама медиа файл өлшемінен кішірек видео файл болады.

Деректерді жасыру процесінде визуалды файл тасымалдаушы файл болып табылады. Видео кадрларды бөлу және генерациялау фабрикасын қолдана отырып, кадрлар медиа файл үшін жасалады. Хабарды байт коды зауыты байт кодына түрлендіреді. Содан кейін хабарламаның байт коды сyptography factory көмегімен шифрланады. Syptography factory үшін 2 кіріс қажет: кілт файлы және құпия сөз. Бұл қауіпсіздіктің жоғарылауына әкеледі, өйткені құпия сөзді тек болжау арқылы жасауға болады. Бұл шифрланған хабарлама steganographic factory модуліндегі медиа файл жақтауларында жасырылған. Нәтижесінде кескін осы кадрлардың көмегімен қайта жасалады.

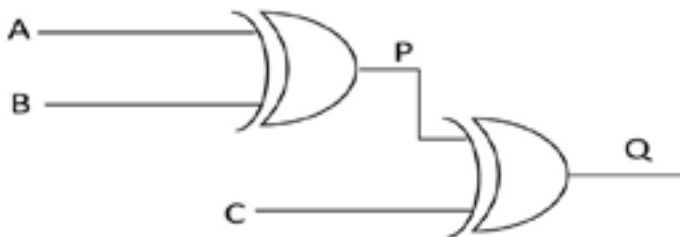
Бұл шифрлау әдісі XOR екі элементі пайдаланылады. 3 суретте осы шифрлау үшін логикалық элементтің қалай жұмыс істейтінін көрсетеді. А кілт файлы, ал В құпия сөзді білдіреді. Құпия сөздің өлшемі кілт файлының өлшемінен кіші болғандықтан, құпия сөз қайталанады. С хабарламаны білдіреді.



Сур.3. Шифрлау үшін логикалық элементтің жұмысы
(Fig.3. operation of the logical element for encryption)

4 суретте шифрды ашу үшін логикалық элементтің қалай жұмыс істейтінін көрсетеді. Q шифрланған хабарламаны білдіреді. R декодталған дегенді білдіреді

С жасырын хабарламасы шифрланған Q хабарламасына тең болатынын көре аламыз.



Сур.4. Шифрды ашу үшін логикалық элементтің жұмысы
(Fig.4. operation of the logical element to decrypt)

Тәжірбиені іске асыру

Бұл криптографияның орнына стеганографияны қолданудың басты артықшылығы. Осылайша, хабарламаны делдалдан жасырудың тиімділігін бағалау маңызды. Хабарламаны жасырудың тиімділігін өлшеу үшін кірістірілген видеолар зерттеушіге хабарламаның бар-жоғын анықтау үшін жіберіледі. Ол үшін 17 ендірілген және 17 ендірілмеген видео файл қолданылады. Зерттеушіге бұл видеоларды сынап көру және оларды кірістірілген хабарламалар мен кәдімгі видеолар ретінде жіктеу ұсынылады. Нәтижелерді келесідей жіктеуге болады. Зерттеуші видеоны келесідей жіктейді

- кірістірілген видео және ендірілген видео (TP)
- кірістірілген видео және кірістірілмеген видео (FP)
- видео кіріктірілмеген және видео ендірілген (FN)
- видео ендірілмеген және видео кірістірілмеген (TN)

Бұл санаттарды ұсынылған әдістің өнімділігін тексеру үшін пайдалануға болады. Дұрыс стеганографиялық әдіс үшін дұрыс жіктеу мен дұрыс емес жіктеу арасындағы байланыс маңызды болмауы керек. Сондықтан өнімділікті өлшеу үшін коэффициент коэффициентін қолдануға болады.

Стеганографияны бағалау кезінде медиа файлдан дұрыс хабарлама алу мүмкіндігін тексеру маңызды. Хабарлама видео файл, аудио файл, сурет немесе мәтіндік құжат болуы мүмкін. Өнімділікті тексеру үшін 5 видео файл, 4 аудио файл, 3 сурет және 3 мәтіндік файл таңдалды. Алынған хабарламамен бастапқы хабарламаны бағалау үшін Aliyaschool.kz веб курстың қолданушыларының ішінен 10 қатысушытаңдалады. Олардан хабарламаларды 0-ден 4-ке дейін бағалауды сұрады, мұнда 0 бастапқы хабарлама мен алынған хабарлама арасындағы айырмашылықтың жоқтығын көрсетеді, ал 4 екі хабарлама арасындағы 100 % айырмашылықты көрсетеді.

1 кесте. Әр жағдайға берілген дәрежелер көрсеткіштері

Дәреже саны	Сипаттамасы
0	Айырмашылығы жоқ
1	Аз ғана айырмашылық бар

2	Орташа айырмашылық бар
3	Үлкен айырмашылық бар
4	жалпы айырмашылық

Видео, аудио, сурет және мәтіннің орташа рейтингі барлық 10 қатысушы үшін есептеледі.

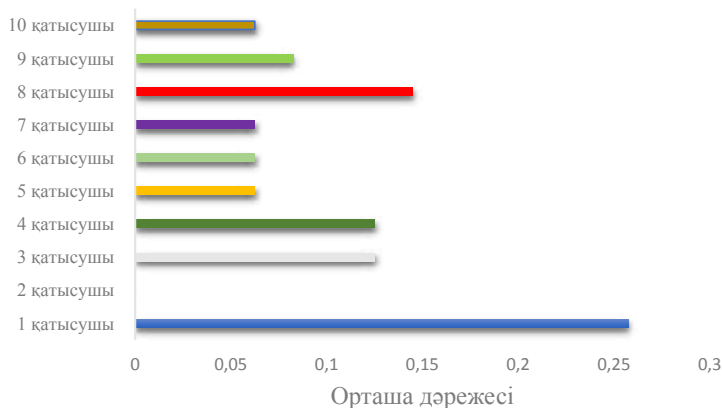
2 кесте. Бір ғана қатысушының нәтиже көрсеткіші

	Видео	Аудио	Сурет	Мәтін
1 файл	0	1	0	0
2 файл	1	1	0	
3 файл	0	0	1	
4 файл	0	0		
5 файл	0			
орташа дәрежесі	0,2	0,5	0,33	0

3 кесте. Барлық қатысушылардың орташа рейтингі көрсеткіштері

	Видео	Аудио	Сурет	Мәтін
1 қатысушы	0,2	0,5	0,33	0
2 қатысушы	0	0	0	0
3 қатысушы	0,25	0,25	0	0
4 қатысушы	0,5	0	0	0
5 қатысушы	0	0,25	0	0
6 қатысушы	0	0,25	0	0
7 қатысушы	0,25	0	0	0
8 қатысушы	0	0,25	0,33	0
9 қатысушы	0	0	0,33	0
10 қатысушы	0,25	0	0	0
	0,145	0,15	0,09	0

Осы нәтижелерден 0–4 дәрежелер диапазонында видео үшін орташа дәреже 0,145, аудио үшін орташа дәреже 0,15, видео үшін орташа дәрежеден аз, кескін үшін орташа дәреже 0,09, видео үшін де, аудио үшін де орташа дәрежеден аз, ал мәтін үшін орташа дәреже дәл 0. Осылайша, жүйенің тасымалдаушы файлдан деректерді дұрыс алудың жоғары өнімділігі бар екені анықталды.



Сур.5. Орташа дәрежелі график көрсеткіштері
(Fig.5. average graph indicators)

Ағымдағы зерттеу визуалды видео мазмұнын жасыру үшін жүргізілді. Видеоны жасырудың маңыздылығы мынада: видеода әдетте үлкен файл болады, өйткені ол көптеген мәліметтерден тұрады. Ұсынылған тәсілдің шегі-ол тек медиа файл ретінде пайдалану үшін қысылмаған, тағы MS AVI файлында қолдайды. Алайда, көрсетілген файл түрі үшін (қысылмаған AVI) ол жақсы тексерілген және жақсы жұмыс істейді. Деректердің қауіпсіздігін арттыру мақсатында хабарлама тасымалдаушы файлға енгізілмес бұрын шифрланған. Бұл шифрлау әдісі кілт файлын да, парольді де қолданады. Бұл жағдайда пароль тандалмайды, ал хабарламаны шифрлау үшін 2 логикалық XOR элементі қолданылады.

Қорытынды

Сонымен қорыта келгенде Стеганография - бұл ақпараттың болуы белгісіз болатындай етіп медиада ақпаратты жасыру өнері (Amirtharajan, 2013). Сандық кескін стеганографиясы қазіргі заманғы қосымшалардың көпшілігінде шешуші болып табылатын байланыс әлеуетін жүзеге асырады. Стеганографияның бірнеше пайдалы қолданылуы бар. Есептеу қуатының таңғажайып дамуы және қауіпсіздік туралы хабарлардың артуы арқылы ол заманауи қауіпсіздік жүйелері арасында көшбасшы болды (Ahmed, 2020). Стеганографиялық техниканы ұсынудағы басты мәселесі - оны криптография және су белгілері сияқты өзара байланысты жүйелерден бөлетін жоғары ендіру сыйымдылығы мен жасырын және қауіпсіздік арасындағы сәйкес тепе-теңдікті сақтау. Бұл мақалада стеганографияның кейбір жаңа әдістеріне кең шолу мен талдау ұсынылады. Сонымен қатар, біз танымал стеганография құралдарын егжей-тегжейлі талқыладық. Терең оқытуға негізделген соңғы стеганографиялық әдістерге қатысты мәселелер шешілді. Пәндік саланы зерттеуге арналған мақаланың соңында болашақ зерттеулердің кейбір бағыттары туралы айтылады.

ӘДЕБИЕТТЕР

- Усатова О.А., Баракова А.Ш. «Анализ современных систем защиты веб-ресурсов» // Известие НАН РК № 1 (341), Алматы, 2022 г. с. 88-95.
- Jiawang Bai, Bin Chen, Kuofeng Gao, Xuan Wang, Shu-Tao Xia. Practical protection against video data leakage via universal adversarial head//Pattern Recognition Volume 131, November 2022, 108834. <https://doi.org/10.1016/j.patcog.2022.108834>
- Трапезников Е.В. Способ защиты видео контента в открытых сетях // ОмГТУ. 2014. №4. <https://cyberleninka.ru/article/n/sposob-zaschity-video-kontenta-v-otkrytyh-setyah> (дата обращения: 25.05.2023).
- Ala Abu-Zahra, Zafar Shahid, Amjad Rattout, William Puech//Independent Protection of Different Layers in Spatially Scalable Video Coding. Procedia Computer Science Volume 10, 2012, Pages 240-246. <https://doi.org/10.1016/j.procs.2012.06.033>
- A. Das, Steganography: Secret data hiding in multimedia, in: Signal Conditioning, Springer, Berlin Heidelberg, 2012, pp. 275–295
- The infinite race between steganography and steganalysis in images Trivikram Muralidharan, Aviad Cohena, Assaf Cohena, Nir Nissima// Signal Processing Volume 201, December 2022, 108711. <https://doi.org/10.1016/j.sigpro.2022.108711>
- Коржик В. И. Цифровая стеганография и цифровые водяные знаки [Текст] / монография: в 2 ч В. И. Коржик // СПбГУТ, 2017.
- Шелухин О.И. Стеганография. Алгоритмы и программная реализация [Текст] / О. И. Шелухин, С. Д. Канаев // Горячая линия - Телеком, 2018.
- Amirtharajan, R., & Rayappan, J. B. B. (2013). Steganography-time to time: A review. Research Journal of the Information Technology, 5, 53–66. <http://dx.doi.org/10.3923/rjit.2013.53.66>
- Ahmed A. AlSabhany, Ahmed Hussain Ali, Farida Ridzuan, A.H. Azni, Mohd Rosmadi Mokhtar//Digital audio steganography: Systematic review, classification, and analysis of the current state of the art. Computer Science Review Volume 38, November 2020, 100316. <https://doi.org/10.1016/j.cosrev.2020.100316>

REFERENCES

- Ala Abu-Zahra, Zafar Shahid, Amjad Rattout, William Puech, 2012. //Independent protection of various levels in spatially scalable video encoding. Procedia Computer Science, volume 10, 2012. Pp. 240–246. <https://doi.org/10.1016/j.procs.2012.06.033>
- Amirtharajan R. and Rayappan J.B.B., 2013. Steganography - from time to time: an overview. Research Journal of Information Technology, 5. Pp. 53–66. <http://dx.doi.org/10.3923/rjit.2013.53.66>
- A. Das, Steganography, 2012. Hiding secret data in multimedia, in: Signal Generation, Springer, Berlin-Heidelberg, 2012. Pp. 275–295
- Jiawang Bai, Bin Chen, Kuofeng Gao, Xuan Wang, Shu-Tao Xia, 2022. Practical protection against video data leakage with universal adversarial head //Pattern Recognition. Volume 131. November 2022. 108834. <https://doi.org/10.1016/j.patcog.2022.108834>
- Korzhek V.I., 2017. Digital steganography and digital watermarks [Text] / monograph: in 2 hours V.I. Korzhik // SPbGUT, 2017.
- Usatova O.A., Barakova A.V., 2022. "Analysis of a modern method of combating HIV infection" // Izvestiya NAS RK No. 1 (341)Almaty, 2022. Pp. 88–95.
- Trapeznikov E.V., 2014. Method of protecting video content in open networks // OmSTU. 2014. № 4. <https://cyberleninka.ru/article/n/sposob-zaschity-video-kontenta-v-otkrytyh-setyah> (accessed: 05/25/2023).
- The endless race between steganography and steganalysis in images Trivikram Muralidharan, Aviad Cohen, Assaf Cohen, Nir Nissima // Signal Processing, volume 201, December 2022. 108711. <https://doi.org/10.1016/j.sigpro.2022.108711>
- Shelukhin O.I., 2018. Steganography. Algorithms and software implementation [Text] / O.I. Shelukhin, S.D. Kanaev // Hotline - Telecom, 2018.
- Ahmed A., 2020. Alsbakhani, Ahmed Hussein Ali, Farida Ridzuan, A.H. Azni, Mohd Rosmadi Mokhtar//Digital audio steganography: a systematic review, classification and analysis of the modern state of technology. Computer Science Review. Volume 38. November 2020. 100316. <https://doi.org/10.1016/j.cosrev.2020.100316>

NEWS OF THE NATIONAL ACADEMY OF SCIENCES OF THE REPUBLIC OF
KAZAKHSTAN
PHYSICO-MATHEMATICAL SERIES
ISSN 1991-346X
Volume 2. Number 346 (2023). 108–127
<https://doi.org/10.32014/2023.2518-1726.188>

UDC 28.23.25

© **A.S. Omarbekova, A.E. Nazyrova*, N. Tasbolatuly, B.Sh. Razakhova, 2023**
L.N. Gumilyov Eurasian National University, Astana, Kazakhstan.
E-mail: nazyrova_aye_1@enu.kz

ONTOLOGICAL MODEL OF AN INTELLIGENT E-LEARNING SYSTEM AND LEARNING OUTCOMES

Omarbekova Assel Sailaubekovna — Ph. D., associate professor of the Department of artificial intelligence, L.N. Gumilyov Eurasian National University, Astana, Kazakhstan

E-mail: omarbekova78@gmail.com, <https://orcid.org/0000-0002-9272-8829>;

Nazyrova Aizhan Esbolovna — doctoral student of the Department of artificial intelligence, L.N. Gumilyov Eurasian National University, Astana, Kazakhstan

E-mail: nazyrova_aye_1@enu.kz, <https://orcid.org/0000-0002-9162-6791>;

Tasbolatuly Nurbolat — associate professor, Deputy Dean of the Higher School of Information Technologies and Engineering, Astana International University, Astana, Kazakhstan

E-mail: tasbolatuly@gmail.com, <https://orcid.org/0000-0002-0511-7000>;

Razakhova Bibigul Shamshanovna — k.t.k., Head of the Department of Artificial Intelligence Technologies, L.N. Gumilev Eurasian National University, Astana, Kazakhstan

E-mail: razakhova_bsh@enu.kz, <https://orcid.org/0000-0002-8152-8661>.

Abstract. This paper presents issues related to intelligent learning systems, ontology model, as well as the development of evaluation of the achieved learning outcomes. The purpose of this article is to investigate the ontology of intelligent learning systems, which is one of the important technologies in the semantic network. An analysis of sources on intelligent learning systems has been conducted. Issues related to the construction of intelligent systems, ITSS, e-learning and LMS have been considered. Nowadays, intelligent learning systems (ILS) are one of the most relevant technologies in the field of education. They allow you to create personalized learning materials and training programs that take into account the characteristics of each student and his or her needs. However, for ILSs to be as effective as possible, it is necessary to develop an ontology — a set of terms and concepts that define the relationship between them in a particular area of knowledge. The purpose of this article is to explore the ontology of intelligent learning systems, analyze sources on ILSs, and consider issues related to building intelligent systems, information technology in education (ITSS), e-learning, and LMSs. The article presents a study of the ILS ontology, which is one of the important technologies in the semantic network. The authors of the article describe the methods of developing

an ontology for the ILS, which allows to take into account the characteristics of each student and his or her needs. Issues related to the construction of ILS s and assessment of learning outcomes are discussed, including an analysis of various ITSSs and their use in e-learning and LMSs. The article discusses the benefits of using ILSs in education. They allow students to receive individualized instruction and maximize their strengths, which contributes to deeper learning and improved academic performance. ILSs allow teachers to maximize their time and resources by creating personalized learning materials for each student.

Keywords: intelligent systems, ITSSs, e-learning, LMS, intelligent assessment system of achieved learning outcomes

© А.С. Омарбекова, А.Е. Назырова*, Н. Тасболатұлы,
Б.Ш. Разахова, 2023

Еуразия ұлттық университеті. Л.Н. Гумилева, Астана, Қазақстан .
E-mail: nazyrova_aye_1@enu.kz

ИНТЕЛЛЕКТУАЛДЫ ELEARNING ЖҮЙЕСІНІҢ ОНТОЛОГИЯЛЫҚ МОДЕЛІ ЖӘНЕ ОҚЫТУ НӘТИЖЕЛЕРІ

Омарбекова Асель Сайлаубековна — т.ғ.к., жасанды интеллект кафедрасының доценті, Л.Н. Гумилев атындағы Еуразия ұлттық университеті, Астана, Қазақстан
E-mail: omarbekova78@gmail.com, <https://orcid.org/0000-0002-9272-8829>;
Назырова Айжан Есболовна — жасанды интеллект кафедрасының докторанты, Л.Н. Гумилев атындағы Еуразия ұлттық университеті, Астана, Қазақстан
E-mail: nazyrova_aye_1@enu.kz, <https://orcid.org/0000-0002-9162-6791>;
Тасболатұлы Нұрболат — доцент, Ақпараттық технологиялар және инженерия жоғары мектебінің декан орынбасары, Астана халықаралық университеті, Астана, Қазақстан
E-mail: tasbolatuly@gmail.com, <https://orcid.org/0000-0002-0511-7000>;
Разахова Бибиғұл Шамшановна — т.ғ.к., Жасанды интеллект технологиялары кафедрасының меңгерушісі, Л.Н. Гумилев атындағы Еуразия ұлттық университеті,
E-mail: razakhova_bsh@enu.kz, <https://orcid.org/0000-0002-8152-8661>.

Аннотация. Бұл мақалада интеллектуалды оқыту жүйелеріне, онтология моделіне, сондай-ақ қол жеткізілген оқу нәтижелерін бағалауды әзірлеуге қатысты сұрақтар берілген. Бұл мақаланың мақсаты семантикалық желідегі маңызды технологиялардың бірі болып табылатын интеллектуалды оқыту жүйелерінің онтологиясын зерттеу болып табылады. Зияткерлік оқыту жүйелері бойынша дереккөздерге талдау жүргізілді. Интеллектуалды жүйелерді, ITSS, eLearning және LMS құруға қатысты мәселелер қарастырылды. Қазіргі уақытта интеллектуалдық оқыту жүйелері білім беру саласындағы ең өзекті технологиялардың бірі болып табылады. Олар әр студенттердің ерекшеліктері мен қажеттіліктерін ескеретін жеке Оқу материалдары мен оқу бағдарламаларын жасауға мүмкіндік береді. Алайда, интеллектуалдық оқыту жүйелері мүмкіндігінше тиімді болуы үшін онтологияны – белгілі бір білім саласында олардың арасындағы байланысты анықтайтын терминдер мен ұғымдардың жиынтығын әзірлеу қажет. Бұл мақаланың мақсаты-

интеллектуалдық оқыту жүйелерінің онтологиясын зерттеу, интеллектуалдық оқыту жүйелері бойынша дереккөздерге талдау жүргізу, сондай-ақ интеллектуалдық жүйелерді, білім берудегі ақпараттық технологияларды (ITSS) электрондық оқытуды және LMS құруға байланысты мәселелерді қарастыру. Мақалада семантикалық желідегі маңызды технологиялардың бірі болып табылатын интеллектуалдық оқыту жүйелері онтологиясын зерттеу ұсынылған. Мақала авторлары әр студенттердің ерекшеліктері мен қажеттіліктерін ескеруге мүмкіндік беретін интеллектуалдық оқыту жүйелері үшін онтологияны дамыту әдістерін сипаттайды. Интеллектуалдық оқыту жүйелері құруға және оқу нәтижелерін бағалауға байланысты мәселелер қарастырылады, соның ішінде әртүрлі itss-ке талдау және оларды eLearning және LMS-те пайдалану. Мақалада интеллектуалдық оқыту жүйелері білім беруде қолданудың артықшылықтары талқыланады. Олар студенттерге оқуға жеке көзқараспен қарауға және олардың күшті жақтарын барынша тиімді пайдалануға мүмкіндік береді, бұл материалды тереңірек игеруге және оқу үлгерімін арттыруға ықпал етеді. Интеллектуалдық оқыту жүйелері оқытушылардың әр оқушы үшін жеке оқу материалдарын жасау арқылы уақыт пен ресурстарды тиімді пайдалануға мүмкіндік береді.

Түйін сөздер: интеллектуалды жүйелер, ITSSs, eLearning, LMS, қол жеткізілген оқу нәтижелерін бағалаудың интеллектуалды жүйесі

© А.С. Омарбекова, А.Е. Назырова*, Н. Тасболатұлы,
Б.Ш. Разахова, 2023

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

E-mail: nazyrova_aye_1@enu.kz

ОНТОЛОГИЧЕСКАЯ МОДЕЛЬ ИНТЕЛЛЕКТУАЛЬНОЙ СИСТЕМЫ ЭЛЕКТРОННОГО ОБУЧЕНИЯ И РЕЗУЛЬТАТЫ ОБУЧЕНИЯ

Омарбекова Асель Сайлаубековна — к.т.н., доцент кафедры искусственного интеллекта, Евразийский национальный университет им. Л.Н. Гумилева, Астана, Казахстан,
E-mail: omarbekova78@gmail.com, <https://orcid.org/0000-0002-9272-8829>;

Назырова Айжан Есболовна — докторант кафедры искусственного интеллекта, Евразийский национальный университет им. Л.Н. Гумилева, Астана, Казахстан
E-mail: nazyrova_aye_1@enu.kz, <https://orcid.org/0000-0002-9162-6791>;

Тасболатұлы Нұрболат — доцент, заместитель декана Высшей школы информационных технологий и инженерии, Международный университет Астана, Астана, Казахстан.
E-mail: tasbolatuly@gmail.com, <https://orcid.org/0000-0002-0511-7000>;

Разахова Бибигул Шамшановна — к.т.н., заведующий кафедрой технологий искусственного интеллекта Л.Н. Гумилева, Астана, Казахстан
E-mail: razakhova_bsh@enu.kz, <https://orcid.org/0000-0002-8152-8661>

Аннотация. В данной статье представлены вопросы, связанные с интеллектуальными системами обучения, моделью онтологии, а также разработкой оценки достигнутых результатов обучения. Целью данной статьи

является исследование онтологии интеллектуальных обучающих систем, которая является одной из важных технологий в семантической сети. Был проведен анализ источников по интеллектуальным системам обучения. Были рассмотрены вопросы, связанные с построением интеллектуальных систем, ITSS, электронного обучения и LMS. В настоящее время интеллектуальные системы обучения (ИСО) являются одной из самых актуальных технологий в сфере образования. Они позволяют создавать персонализированные учебные материалы и обучающие программы, учитывающие особенности каждого ученика и его потребности. Однако, для того чтобы ИСО были максимально эффективны, необходимо разрабатывать онтологию – совокупность терминов и понятий, которая определяет взаимосвязь между ними в конкретной области знаний. Цель данной статьи – исследовать онтологию интеллектуальных обучающих систем, провести анализ источников по ИСО, а также рассмотреть вопросы, связанные с построением интеллектуальных систем, информационных технологий в образовании (ITSS), электронного обучения и LMS. В статье представлено исследование онтологии ИСО, которая является одной из важных технологий в семантической сети. Авторы статьи описывают методы разработки онтологии для ИСО, которые позволяют учитывать особенности каждого ученика и его потребности. Рассматриваются вопросы, связанные с построением ИСО и оценкой результатов обучения, в том числе проводится анализ различных ITSS и их использование в электронном обучении и LMS. В статье обсуждаются преимущества использования ИСО в образовании. Они позволяют студентам получать индивидуальный подход к обучению и максимально эффективно использовать свои сильные стороны, что способствует более глубокому усвоению материала и повышению успеваемости. ИСО позволяют учителям максимально эффективно использовать свое время и ресурсы, создавая персонализированные учебные материалы для каждого ученика.

Ключевые слова: интеллектуальные системы, ITSS, электронное обучение, LMS, интеллектуальная система оценки достигнутых результатов обучения

Introduction

Digitalization of education involves the use of modern digital educational technologies that differ from the previous e-learning system and LMS content management system in the presence of flexible functions not only for management, but also for analyzing and evaluating student achievements. The development of intelligent assessment system of achieved learning outcomes is urgent. The implementation of the technology of the intelligent assessment system of achieved learning outcomes in general, specifically with the content in the Kazakh language as well, is not sufficiently disclosed both in the world of science and in the Republic of Kazakhstan. In this regard, the disclosure of the essence and method of implementation of such technology is an urgent task.

The demand for intelligent systems is growing at a record pace during the pandemic. In educational system of the Republic of Kazakhstan various educational learning systems, in which the assessment function is partially implemented and is often limited to automatic verification of test and interactive tasks or in the form of manual verification by a teacher.

The crucial aim of the article is to conduct a literature review, analyse methods for assessing and predicting knowledge, skills and competencies of students in the intelligent technologies.

“Intelligent training system is the system with the elements of artificial intelligence. The intelligent training system allows you to adapt the learning process to the each individual student characteristics working with the system. Learning management is defined by the learning system itself based on learning outcomes. Implementation is carried out on the ground of knowledge about the subject area, the learning process and the learner”.

Related works. The literature review of different authors from 2017–2022 has been conducted and analysed.

The analysis of the article Intelligent Tutoring Systems: a systematic review of characteristics, applications, and evaluation methods (Mousavinasab et al., 2018) has been carried out in accordance with the artificial intelligence methods for creating more personalized educational system. Such systems are called Intelligent Tutoring Systems (ITSs). The authors of the article focus on the characteristics of ITSs developed in various fields of education. The authors conducted research from 2007 to 2017, and 53 articles were included in the study on the basis of inclusion criteria. The educational areas in ITSs were mainly computer science (37.73 %). The most frequent artificial intelligence methods used in rule-based reasoning, data mining, and Bayesian networking with frequencies of 33.96 %, 22.64 %, and 20.75 %, respectively. These methods allow ITSs to provide adaptive guidance and learning, evaluate students, determine and update the student model, and classify or group students. The system performance indicators, student performance and experience were used to evaluate ITSs in particular.

Further the Smart Learning (García-Peñalvo et al., 2020) has been considered. In the consequence of pandemic due to COVID-19 artificial intelligence has a huge potential, applied in the educational sphere. Therefore, educational systems claim to use more intelligent learning technologies that do not claim to replace the teaching staff, but facilitate their teaching activities. This special issue is focused on introducing a collection of articles about original advancements in educational applications and services driven by AI, big data, machine and deep learning.

In article An intelligent tutoring system for supporting active learning: A case study on predictive parsing learning (Castro-Schez et al., 2021) authors designed and developed an intelligent mentoring system to allow students to learn through experience, working at the top cognitive level of Bloom's Taxonomy, encouraging them to actively learn and be self-sufficient. Thus, the developed tool called Proletool 3.0 (Figure 1) includes modules that are commonly encountered in

systems of this type (domain module, mentor module, student module, interface module, and evaluation module). In addition, the system provides a web interface that allows you to use it as a learning object on the e-learning platform.

Authors (Grivokostopoulou et al., 2016) have utilised the app for interactive examples and exercises related to search algorithms. The research area is the artificial intelligence curriculum. The issue of using Bsearch algorithms has been considered to offer theoretical descriptions, interactive examples and exercises related to search algorithms in the artificial intelligence training system. The results show that visualized animation and interactivity are two of the most important factors contributing to better learning.

A methodology for assessing the student intelligence level based on the theory of multiple intelligences. The research area is computer engineering. The system is based on a hybrid approach with artificial intelligence generated by an artificial neural network and an optimization algorithm called the Vortex Optimization Algorithm (VOA). It was tried and tested on students and gave positive results for improving self-learning (Kose et al., 2017).

In the article Application intelligent search and recommendation system based on speech recognition technology (Jiang et al., 2020) was considered the speech recognition technology to create smart search and app recommendation. The authors designed a system of extraction levels, speech information analysis and finalized the recommendations. Experimental results show that the proposed method is more efficient and intelligent.

In the article Early segmentation of students according to their academic performance: A predictive modelling approach (Jiang et al., 2020) were considered the the methods Random Forests, decision trees, support vector machines and naive Bayes. The authors propose a two-stage model supported by data mining methods that uses information available at the end of the first year of students ' academic career (path) to predict their overall academic performance. Unlike most educational data mining literature, academic success is measured by both the average score achieved and the time taken to complete the degree. In addition, this study suggests segmenting students based on the dichotomy between evidence of failure or high performance at the beginning of the degree program and student performance levels predicted by the model.

Investigates the relationship of two aspects to objective and subjective academic achievements, as well as their incremental validity and interaction with the two aspects of conscientiousness diligence and orderliness. Cross-sectional data from 424 students showed that a) intellectual openness is a strong positive predictor of academic achievement, while sensory-aesthetic openness is a moderate negative predictor, b) aspects of openness showed significantly greater incremental confidence in aspects of conscientiousness than vice versa, and c) intellectual openness and order interacted in predicting objective academic achievement (Thomas Gatzka, 2021).

Therefore, in the article Improving the expressiveness of black-box models for

predicting student performance (Carlos et al., 2017) has been proposed a black box method that allows us to take advantage of the power and versatility of these methods in making some decisions about the input data and the design of the classifier, providing a rich set of output data. A set of graphical tools has also been offered to use the output information and provide meaningful guidance for teachers and students.

Authors of the article A new methodology for early warning of critical academic performance, based on discrete predictive models (Meca et al., 2019) have presented a methodology based on the hypothesis that the method of sampling the performance variable will significantly affect the generated predictive models. The authors have emphasized the relationship between the factorization criteria and the most relevant attributes selected for instance classification. In addition, the relationship between the obtained predictive model and the criteria for factorization of the target variable has been presented.

Three studies (Mousavinasab et al., 2018), (Garcia-Penalvo et al., 2020), (Castro-Schez et al., 2021) conducted a comparative analysis of the effectiveness of intelligent systems used in the educational sphere.

The analysis showed that there is a large amount of research on assessment and prediction methods in e-learning, students' skills competencies, and academic achievements (Zulfija et al., 2019).

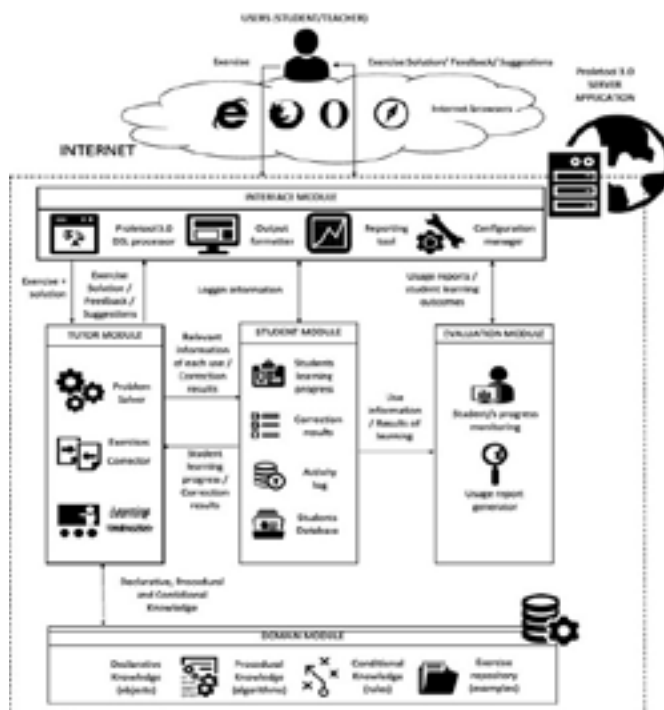


Fig. 1. Architecture of the web interface system software according to Bloom's Taxonomy (Castro-Schez et al.:2021)

In this paper, the authors discussed (Quezada et al., 2021) about using NEON (ontology methodology), Scrum (agile methodology), and web scraping to develop an application of the ontology model of knowledge representation for knowledge arrays as an additional tool to improve the curriculum in the context of software development. The benefits of this paper are as follows: First, it combines and applies ontology methodologies and agile development; second, a Java application using the Jsoup library has been developed to extract and visualize information about a bachelor's degree in software engineering.

One can see the design of a competence ontology for educational program analysis. It is shown as an analytical review of the works on this topic. The methodology of the process of collecting and analyzing professional and personal qualities is proposed. Created data on programmer diploma requirements. The ontology of required competencies for this specialty is created. For example, the ontology of key skills for the profession "Programmer" is given. It is possible to see variants of compositions of all collected qualities. The practical implementation of the transition and the combination of research related to market analysis and research related to the dynamic changes in education is explained. The resulting lists of professional qualities and personality traits will then be classified into generalized groups (Aksenov et al., 2020).

The need to ensure compliance of professional training with both formal and informal requirements of the education system and the labor market of this work is confirmed by the relevance. The reason for the employment of university graduates in the labor market, is the lack of methods and technologies that are able to align the results of the information technology curriculum with the requirements of the labor market. This article (Kurzaeva et al., 2020) describes methods for developing an ontological model that would be capable of shaping the content of IT competencies and assessing them in university students. Firstly, the boundaries of the subject area of awareness must be defined, and secondly, a method for taking into account individual characteristics and needs must be considered and developed in advance to create an individual training and self-study plan, these two theses are our basis. On the example of the subject area of business solutions platform "1C:Enterprise" presented ontological model, developed by the stated method for the formation of the content of IT competencies and their evaluation. Knowledge base, created with the help of ontological model, can be integrated into the intellectual information systems of decision-making, designed for education, corporate training and human resource management. For professionals developing new tools and methods of distance education and distance learning management this article will be useful.

The professor is known as the mentor who guides and guides students throughout their professional learning, this is provided in the competency-based model. Consequently, the professor must determine which courses students have approved and which courses they should take during the next period so that their academic learning does not fall behind. Curricula should be modeled to indicate a complex structure consisting of courses (which are represented as nodes), links (denoted

as edges) between course prerequisites, and semester number (represented by a node weight). The paper (Galvez et al., 2020) provides a bi-directional weighted graph that will represent internal and wire constraints in the flow of courses that students must take and pass, creating an organized context (which is governed by each course weight) in which they can be guided and directed by the instructor. A system is proposed for monitoring student progress and enabling student-teacher interaction that integrates a weighted bi-directional schedule. Developments that are based on an ontology model is a collaborative application, such systems. So, this work integrates graphs with artificial intelligence.

The article (Mandić et al., 2018) also describes a semi-automated software platform developed to harmonize computer science curricula at all levels of education. The algorithms used for ontology matching are described in detail, as well as the principle of matching computer science curricula with ontology models. For computer science teachers from the Republic of Serbia, the model of the selected curriculum was created and compared with the model of the reference curriculum for computer science teachers using the software platform. According to the analyses of the results, which include a comparison with the data obtained for other possible pairs of created input ontology models (ACM K12 high school model and the reference model, high school model and the model of the selected curriculum). The experiments presented in this article indicate that consideration should be given to improving the curriculum for teachers, as well as the use of new matching methods.

In the field of education, the semantic network and the development of ontologies can play a significant role. In this (Katis et al., 2020) discussion, the authors focus on conceptualizing educational knowledge structures in an academic setting. More specifically, we present the methodology and process of developing an educational ontology. The methodology could be reused and applied to any type of course at various institutions and contribute to several objectives of the curriculum and course activities.

In education, ontologies have been used with great success because they allow us to formulate a representation of the subject domain by identifying all the concepts involved, the relations between the concepts, and all the existing properties and conditions. The purpose of this paper (Stancin et al., 2020) is to present the field of ontologies and provide an overview of recent research in the field in the context of education. As this paper presents a literature review, articles from the past five years, have been collected from the IEEE Xplore database, analyzed and classified based on the use of ontologies for: curriculum modeling and management, description of learning domains, learning data and e-learning services. From the resulting work, one can observe a somewhat growing trend in the contribution of ontologies to educational systems. Many analyses used ontologies to describe learning domains, and some of the 95 papers collected could not fit into just one category because the system used more than one ontology. Throughout the work, the following contributions were made: the term "ontology" was defined, the most common types

of ontologies and commonly used methodologies for building ontologies were identified, and an overview of existing systems that use ontologies in education was provided.

With the emergence of data-driven technologies, the importance of education and training in data science is increasing, and of an organizational culture that aims to gain practical value for improving the research process or enterprise business by using a variety of enterprise data and widely available open data and social media data. By today's research and data-driven industries require new types of professionals capable of supporting all phases of the data lifecycle, from production, input, to data processing and delivery of actionable results, visualization and reporting, which can be collectively defined as a family of data science professions. An interdisciplinary approach is required for education and for professional training of data scientists, combining a broad view of the fundamentals of data science and analytics with in-depth practical knowledge of specific subject areas. In current situations with rapidly changing technologies and high demand for professional skills, data science education and training must be customizable and provided in a variety of forms, as well as provide sufficient opportunities for data processing labs for hands-on learning. Article (Demchenko et al., 2020) discusses an approach to creating a customizable data science curriculum for different types of learners based on the EDISON Data Science Framework (EDSF) ontology, developed through the EU-funded EDISON project and widely used by universities and professional learning organizations.

The art and expertise of data management are critical to the widespread adoption of open science and the effective use of data in research, industry, business, and other sectors of the economy. Unbiased data management views (searchable - accessible - compatible - reusable) and data management provide the foundation for effective research data management. The paper (Demchenko et al., 2021) "Turning YARMARKS into Reality" and others recommend that data skills be more widely integrated into university curricula and that a concerted effort be made to coordinate and accelerate the teaching of professional roles in data processing. Reliable Data Skills for Learning is being worked on by many projects and initiatives, across Europe and beyond. Here is the ongoing work of the FAIRsFAIR project to develop a data management competency framework and provide guidance on how to implement this framework into university curricula by defining model data management curricula. In the labor market analysis, the proposed method and identified competencies and knowledge themes are validated. The material presented uses the EDISON Data Science Framework as a framework for defining data management competencies and a methodology for linking competencies, skills, knowledge, and intended learning outcomes in curriculum development.

As a result of the formation of e-learning systems, personalization and adaptability have now become important characteristics of educational technology. By design, the authors describe (Aeiad et al., 2019) the development of an architecture for a

personalized and adaptable e-learning system (APELS) that attempts to contribute to advances in this area. A personalized and adaptable learning environment APELS seeks to provide users from freely available resources on the Internet. Modeling a specific learning subject and extracting relevant learning resources has been used an ontology from the Internet based on the learner model (learners' experiences, needs and learning styles). The APELS processing methods is to use natural language to evaluate content extracted from relevant resources according to a set of learning outcomes defined by standard curricula to ensure appropriate learning of the subject matter. A computer science primer is used to illustrate the working mechanisms of APELS and its assessment based on the ACM / IEEE computing curriculum. Experimental opinions were conducted with subject matter experts to assess whether APELS can create the right instructional content to meet the learning needs of the student. The scores show that the content created by APELS is of good quality and satisfies the learning outcomes for teaching purposes.

Ontologies are one of the pillars of the semantic network. The semantic network tangibly relies on formal ontologies that structure subordinate data and information in a machine-readable way so that this knowledge can be automatically read and processed by machines. In this process, the authors (Piedra et al., 2018) present an ontology approach to the foundations required for the IEEE and ACM computer science curricula (CS2013). The goal of this process is to improve the usability of the curriculum through the development and implementation of a computer science ontology. Through ontology definitions of concepts, data properties, and data objects can be useful to overlay and maintain a logical structure for new types of curriculum concepts that may emerge from autonomous higher education institutions (HEIs). Users gain a resource from obtaining information in open portable data formats. As a result, the developed ontology is a powerful tool that can change the way curriculum knowledge is defined, managed, negotiated, analyzed, and shared. The Semantic Curriculum Viewpoint can be used in a variety of learning modes, facilitate academic mobility, accelerate the recognition of academic credits earned at different institutions, and integrate and interact with different computer science faculty curricula or provide convergence and interoperability of other disciplines.

A course of teaching and learning is a curriculum because it contains an important promise between faculty and students in higher education and in the university. But at the moment, most curriculum management systems provide simple functionality, including the creation, modification and retrieval of unstructured curriculum. The approach of the authors (Chung et al., 2014) consists of defining the ontological structure of the curriculum and the semantic relationships of the curriculum, classifying and integrating the curriculum based on ACM/IEEE computing curriculum, and formalizing the learning goals, learning activities, and learning assessment in the curriculum using Bloom's taxonomy to improve the usability of the curriculum. Among other things, we propose an effective method for enhancing student learning effects through the construction of a subject ontology that is used in discussion, visual presentation, and knowledge sharing between instructor and

students. The authors assure the correctness of the search and classification of our proposed methods according to experiments and performance evaluations.

Although model-based software engineering (MBE) is a widely recognized discipline in the field of software engineering (SE), no agreed basic set of concepts and practices (i.e., body of knowledge) has yet been defined for it (Burgueño et al., 2019). In order to properly the content of the MBE discipline, promote a global, consistent view of it, clarify its scope with respect to other SE disciplines, and define a framework for MBE curriculum development, this paper proposes content for the body of knowledge for MBE. In doing so, the authors outline the methodology they used to compile the proposed list of content, as well as the results of a survey we conducted to solicit community opinion on the importance of the proposed topics and their level of coverage in existing SE curricula.

Theoretical analysis of the literature shows that this problem has been widely considered. At the same time, education implies the use of such forms and methods of learning, which allow to include in the educational process. The most important role in this process can be played by the ontological model, to which additional requirements are imposed in the conditions of higher education.

Materials and methods

In the materials and methods section of the e-learning system, two key aspects are outlined: the definition of the main components of the system and the development of an ontological model to establish the semantic structure and relationships between these components. Here's a breakdown of each aspect:

Definition of Main Components:

The main components of the e-learning system are defined, including:

Courses: These are the overarching units of instruction that cover specific topics or subject areas. Courses typically consist of a series of lessons, activities, and assessments designed to facilitate learning.

Modules: Courses can be divided into smaller units called modules. Modules focus on specific subtopics or learning objectives within a course. They provide a structured approach to organizing the content and activities for effective learning.

Assignments: These are tasks or projects given to students as part of their learning process. Assignments can include essays, research papers, group projects, or any other form of work that assesses the students' understanding and application of the course material.

Tests: Tests are assessments designed to evaluate students' knowledge and understanding of the course material. They can take various forms, such as multiple-choice questions, open-ended questions, or practical exercises.

Other Elements: This refers to any additional components or features that are part of the educational process in the e-learning system. It could include discussion forums, virtual labs, interactive simulations, or any other tools or resources used to enhance the learning experience.

Development of Ontological Model:

An ontological model is created to define the semantic structure and

relationships between the components of the e-learning system. An ontology is a formal representation of knowledge that describes entities, their attributes, and the relationships between them.

In this context, the ontological model would define the relationships between courses, modules, assignments, tests, and other elements of the e-learning system. It establishes how these components are interconnected and how they relate to each other within the overall educational framework.

The ontological model can provide a structured and organized representation of the e-learning system, facilitating effective content management, navigation, and retrieval. It helps ensure coherence, consistency, and interoperability of the system components, enabling seamless integration and interaction between different elements.

Stages of creating an intelligent e-learning system.

1. Learning outcome formulation

The formulation of discipline learning outcomes is carried out taking into account the form of training, as the achievement of learning outcomes may differ in distance, mixed and traditional training. However, the results must be formalized as achievable for any type of training. Each learning result corresponds to (at least does not contradict) the purpose of the educational program.

2. Ontological model of the discipline

The structure and content. The ontological model of the discipline is being built on Protégé. tool (Center:2019). <https://protege.stanford.edu/> It was developed at Stanford University in collaboration with the University of Manchester.

The model includes topics of the discipline and a glossary, each topic of the discipline includes basic concepts from the glossary, control questions on the topic with links to the glossary are being developed, as well as tasks to check the achievement of results on each topic, questions to check knowledge and their answers on ontology in natural language is created, the system can be self-studiable within a given discipline (Fig.2) (Zulfiya et al., 2019).

3. Assessment system development

Assume test question assignments and create a knowledge base with reference answers on the subject of databases:

1 What is a relational data model?

Reference answer: A relational data model is a structured approach to organising data in a database based on representing data as tables consisting of rows (tuples) and columns (attributes) and defining relationships between tables.

2 What is a primary key?

Reference answer: A primary key is a unique identifier that uniquely identifies each record (row) in a database table. It ensures uniqueness and integrity of the data in the table.

3 What is a foreign key?

Reference answer: A foreign key is an attribute (column) in a table that establishes a relationship between two tables in a database. It refers to the primary

key of the other table and is used to determine the relationship between the data in the different tables.

4 What is database normalisation?

Reference answer: Database normalisation is the process of organising data in tables to eliminate redundancy and ensure data integrity. It involves dividing tables into smaller and related tables using certain normal forms.

5 What is SQL?

Reference answer: SQL (Structured Query Language) is a programming language used to manipulate relational databases. It allows you to perform data creation, modification and retrieval operations as well as manage the structure of the database.

By creating a knowledge base with benchmark answers to the test questions, we can compare students' answers with the benchmark answers to assess their understanding and knowledge of databases.

An example of student entries for test questions in a database subject:

Question: What is a relational data model?

Student Answer: A relational data model is a way of organising data in a database.

Question: What is a primary key?

Student's Answer: A primary key is a unique identifier that helps identify records in a table.

Question: What is a foreign key?

Student Answer: A foreign key is an attribute that links two tables in a database.

Question: What is database normalization?

Student's Answer: Database normalization is the process of improving the structure of data in tables.

Question: What is SQL?

Student answer: SQL is a programming language used to work with databases.

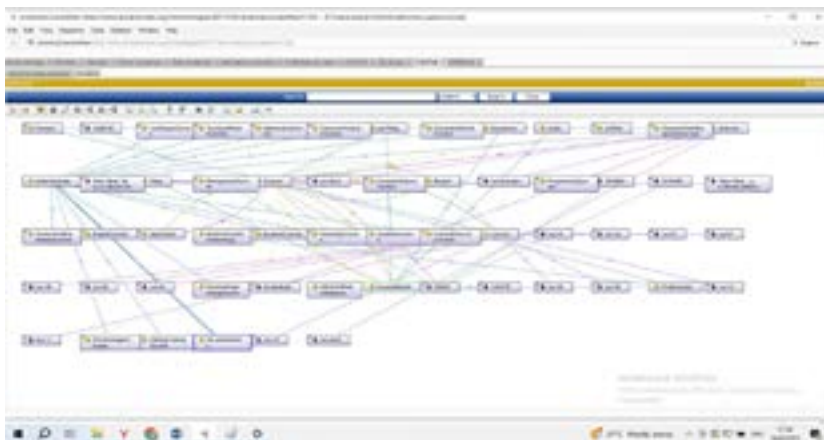


Fig. 2. Ontological model of the discipline

Student's record of answers can be assessed by comparing it to benchmark answers from a knowledge base. For each question, the degree of similarity between the student's answer and the benchmark answer can be determined to calculate the student's score for each question and the overall test score. This will assess the student's understanding and knowledge of the databases.

To calculate the degree of similarity between a student response and a reference response based on a fuzzy binary relationship, a similarity measure, such as cosine similarity, can be used. Cosine similarity measures the angle between the vectors of the student's response and the reference response, and the closer the angle is to 0, the higher the degree of similarity.

An example of calculating the degree of similarity between a student's answer and a reference answer using cosine similarity:

Suppose that the reference answer to the question "What is a relational data model?" is "A relational data model is a way of organizing data in a database."

And the student's answer to this question is "A relational data model is a way of storing information in tables."

We can represent each answer as a vector, where each word is a component of a vector, and then calculate the cosine of the angle between these vectors:

Reference answer: [relational, data, model, is, a, way, of, organising, a, database]

Student answer: [relational, data, model, is, a, way of, storing, information, tables]

Now, let's calculate the Cosine Similarity between these vectors:

$$\text{Cosine Similarity} = (A \cdot B) / (\|A\| * \|B\|)$$

Where A and B are vectors of student and benchmark answers respectively, and $\|A\|$ and $\|B\|$ are their norms (lengths).

After calculating the cosine similarity, we obtain a value that reflects the degree of similarity between the student's answer and the reference answer. The closer the value is to 1, the higher the degree of similarity.

Applying this approach to other questions will calculate the degree of similarity and evaluate student responses based on the fuzzy binary relationship.

To determine scores based on the degree of similarity between the student's answer and the benchmark answer, a similarity threshold value can be used and scores can be assigned depending on how close the answers are.

An example of calculating scores based on a similarity threshold:

Question: What is a relational data model?

Reference answer: "A relational data model is a way of organizing data in a database."

Student response: "A relational data model is a way of storing information in tables."

Let the similarity threshold be 0.8.

We calculate the cosine similarity between the answer vectors and get a value of, for example, 0.85.

Since the similarity value (0.85) is higher than the threshold value (0.8), the student receives a full grade for this question.

Question: What is a primary key?

Reference answer: "A primary key is a unique identifier that helps identify records in a table."

Student answer: "A primary key is an identifier for records in a table."

Let the similarity threshold be 0.7.

We calculate the cosine similarity between the response vectors and obtain a value of, for example, 0.65.

Since the similarity value (0.65) is below the threshold value (0.7), the student receives a partial score for this question, such as 0.5.

Scores for the remaining questions can be calculated using the similarity threshold value and setting appropriate scores based on the proximity of the student's answers to the benchmark answers.

To calculate the competence score based on the discipline scores for each module, a weighted sum of scores can be used, where each discipline contributes to the overall competence score.

An example of calculating competence scores based on discipline scores:

Suppose we have two competencies, Competency 1 and Competency 2.

Module 1:

- Discipline 1: Databases
- Discipline 2: Database Management Systems
- Discipline 3: SQL Language

The student received the following marks for the disciplines in Module 1:

- Databases: 85 points
- Database Management Systems: 90 points
- SQL Language: 80 points

Now, to calculate competency scores, assign weights to each discipline according to their importance for each competency.

Let us assume that the weights for the competencies are as follows:

- Competency 1: Databases - 0.4, Database Management Systems - 0.3, SQL Language - 0.3

- Competency 2: Databases - 0.2, Database Management Systems - 0.5, SQL Language - 0.3

Now let's calculate competency scores:

Competency 1 scores = (Database scores * Database weight) + (DBMS scores * DBMS weight) + (SQL Language scores * SQL Language weight) = $(85 * 0.4) + (90 * 0.3) + (80 * 0.3) = 34 + 27 + 24 = 85$ points

Competency 2 Scores = (Database Scores * Database Weight) + (RDBMS Scores * RDBMS Weight) + (SQL Language Scores * SQL Language Weight) = $(85 * 0.2) + (90 * 0.5) + (80 * 0.3) = 17 + 45 + 24 = 86$ points

Thus, the student received 85 points on competency 1 and 86 points on competency 2 based on the points received on the disciplines included in module 1.

It is possible to calculate competence scores for other modules by applying weights of disciplines depending on their importance for each competence.

To calculate the average achievement of the compulsory learning outcomes of an educational programme, the sum of the points for each discipline can be used and divided by the total number of disciplines. This will provide an overall assessment of the achievement of the learning outcomes throughout the programme.

An example of calculating the average achievement of compulsory learning outcomes:

Suppose we have the following subjects in an educational programme:

1. databases - 85 points
2. Database management systems - 90 points
3. SQL Language - 80 points

To calculate the average of the required learning outcomes, we add up the scores for each discipline and divide them by the total number of disciplines:

$$\text{Sum of points} = 85 + 90 + 80 = 255$$

$$\text{Total number of disciplines} = 3$$

$$\text{Average Achievement} = \text{Sum of Scores} / \text{Total Number of Disciplines} = 255 / 3 = 85$$

Thus, the average achievement of the compulsory learning outcomes of the disciplines in the educational programme is 85. This makes it possible to assess the overall level of achievement of the learning outcomes throughout the programme.

It is important to note that thresholds and scores can be adjusted according to the requirements and preferences of the assessor or the educational institution.

It is proposed to use an assessment system on the basis of open tests in a specific discipline. The student writes his answer, which is compared with the reference answer based on a fuzzy binary relation set between the student's answers and the reference answers synthesized from the knowledge base of the given discipline in accordance with the test questions. Competence scores can then be calculated for all modules based on the scores obtained in the disciplines included in those modules. As a result, it is possible to assess the achievement of learning outcomes throughout the educational program, which is calculated as the average value of the achievement of compulsory learning outcomes in the disciplines of the educational program.

Results and discussion

The proposed assessment system based on open tests in a specific discipline appears to be designed to evaluate student performance and measure the achievement of learning outcomes. Here's a breakdown of the process and its potential implications:

Open Tests: The assessment system utilizes open tests, where students provide their answers in written form. This format allows for more flexibility and diverse responses compared to multiple-choice questions.

Fuzzy Binary Relation: The student's answers are compared with reference answers using a fuzzy binary relation set. This approach suggests that the evaluation is not limited to strict correctness or incorrectness but considers the degree of similarity or dissimilarity between the student's response and the reference answer.

Knowledge Base: The reference answers are synthesized from the knowledge base of the specific discipline. This knowledge base likely contains a comprehensive collection of information and expertise related to the discipline, which serves as a basis for determining the correctness or quality of the answers.

Competence Scores: Competence scores can be calculated for all modules based on the scores obtained in the disciplines included in those modules. This allows for assessing the overall performance and competence level of students across different areas of knowledge covered in the educational program.

Achievement of Learning Outcomes: The proposed system aims to assess the achievement of learning outcomes throughout the educational program. This involves calculating the average value of the achievement of compulsory learning outcomes in the disciplines of the program. It provides a comprehensive overview of students' progress and proficiency in meeting the defined learning goals.

The strengths of this approach lie in its potential objectivity, scalability, and adaptability. By employing a fuzzy binary relation set and referencing a knowledge base, the system can provide a more nuanced assessment that goes beyond simple right or wrong answers. Additionally, the calculation of competence scores and the assessment of learning outcomes can offer valuable insights into students' overall performance and program effectiveness.

Conclusion

The proposed approach you mentioned seems to focus on creating a teaching system that doesn't require a human teacher. Instead, it relies on some other mechanism or process to facilitate learning. Additionally, it suggests using a large knowledge base to verify the attainment of learning outcomes.

This type of approach aligns with the concept of machine learning and artificial intelligence, where systems can learn and improve from data without explicit programming. By leveraging a large knowledge base, the system can access a vast amount of information to enhance the learning process and verify the achievement of learning goals.

Using a teaching system without a human teacher can have several advantages. It can provide scalability, as the system can potentially handle a large number of learners simultaneously. It can also offer personalized learning experiences, tailoring the content and pace of instruction to individual learners' needs. Furthermore, such a system can adapt and improve over time based on feedback and data analysis.

However, it's important to note that while a teaching system without a human teacher can be valuable in certain contexts, it may not completely replace the role of human educators. Human interaction, guidance, and support are often crucial for fostering deeper understanding, critical thinking, and socio-emotional development in learners.

REFERENCES

- Aksenov Alexander & Borisov Vasilii & Shadrin Denis & Porubov Andrey & Kotegova Anna & Sozykin Andrey, 2020. Competencies Ontology for the Analysis of Educational Programs. Pp. 368–371. 10.1109/USBREIT48449.2020.9117793. (in Eng.).
- Aciad E., Meziane F., 2019. An adaptable and personalised E-learning system applied to computer science Programmes design. *Educ Inf Technol* 24. Pp.1485–1509 <https://doi.org/10.1007/s10639-018-9836-x>. (in Eng.).
- Burgueño L., Ciccozzi F., Famelis M. et al., 2019. Contents for a Model-Based Software Engineering Body of Knowledge. *Softw Syst Model* 18. Pp. 3193–3205. <https://doi.org/10.1007/s10270-019-00746-9>. (in Eng.).
- Gross N., Frickel S., 2005. A General Theory of Scientific. Intellectual Movements // *American Sociological Review*. - Vol. 70. - Pp. 204–232. (in Eng.).
- García-Peñalvo, Francisco & Casado-Lumbreras, Cristina & Colomo-Palacios, Ricardo & Yadav, Aman. (2020). Smart Learning. *Applied Sciences*. 10. 6964. 10.3390/app10196964. (in Eng.).
- Grivokostopoulou I. Perikos and I. Hatzilygeroudis, 2016. An innovative educational environment based on virtual reality and gamification for learning search algorithms. *IEEE Eighth International Conference on Technology for Education (T4E)*, Mumbai. Pp.110–115, doi:10.1109/T4E.2016.029. (in Eng.).
- Galvez Luz & Anzures-García Mario & Gregorio Álvaro, 2020. Weighted Bidirectional Graph-based Academic Curricula Model to Support the Tutorial Competence. *Computación y Sistemas*. 24. 10.13053/cys-24-2-3397. (in Eng.).
- Demchenko Yuri & Communiello Luca & Reali Gianluca, 2019. Designing Customisable Data Science Curriculum Using Ontology for Data Science Competences and Body of Knowledge. *ICBDE'19: Proceedings of the 2019 International Conference on Big Data and Education*. Pp. 124–128. 10.1145/3322134.3322143. (in Eng.).
- Mousavinasab E., Zarifsanaiy N., R. Niakan Kalhori S., Rakhshan M., Keikha L. & Ghazi Saeedi M, 2018. Intelligent Learning Systems: A systematic review of characteristics, applications and assessment methods. *Interactive Learning Environment*. <https://doi.org/10.1080/10494820.2018.1558257>. (in Eng.).
- Meca Inmaculada & Mollá-Campello Nuria & Rabasa A., 2019. A new methodology for early warning of critical academic performance, based on discrete predictive models. *TEEM'19: Proceedings of the Seventh International Conference on Technological Ecosystems for Enhancing Multiculturality*. Pp. 680–685. 10.1145/3362789.3362937. (in Eng.).
- Mandić Milinko, 2018. Semantic Web based software platform for curriculum harmonization. *WIMS '18: Proceedings of the 8th International Conference on Web Intelligence, Mining and Semantics*. Pp. 1–9. 10.1145/3227609.3227654. (in Eng.).
- N. Piedra and E.T. Caro, 2018. "LOD-CS2013: Multilearning through a semantic representation of IEEE computer science curricula," 2018 IEEE Global Engineering Education Conference (EDUCON). Pp. 1939–1948, doi: 10.1109/EDUCON.2018.8363473. (in Eng.).
- J.J. Castro-Schez, C. Glez-Morcillo, J. Albusac, D. Vallejo, 2021. An intelligent tutoring system for supporting active learning: A case study on predictive parsing learning, *Information Sciences*. Volume 544. Pp. 446–468. ISSN 0020–0255. <https://doi.org/10.1016/j.ins.2020.08.079>. (in Eng.).
- Jiang J., Wang H.H., 2020. application of an intelligent search and recommendation system based on speech recognition technology. *Int J Speech Technol* <https://doi.org/10.1007/s10772-020-09703-0>. (in Eng.).
- Stancin K., Poscic P. & Jaksic D., 2020. Ontologies in education – state of the art. *Educ Inf Technol* 25. Pp. 5301–5320. <https://doi.org/10.1007/s10639-020-10226-z>. (in Eng.).
- Carlos J. Villagrà-Arnedo, Francisco J. Gallego-Durán, Faraón Llorens-Largo, Patricia Compañ-Rosique, Rosana Satorre-Cuerda, Rafael Molina-Carmona. (2017). Improving the expressiveness of black-box models for predicting student performance, *Computers in Human Behavior*. Volume 72. Pp. 621–631. ISSN 0747-5632, <https://doi.org/10.1016/j.chb.2016.09.001>. (in Eng.).
- Center S., 2019. protégé. [Stanford.edu](https://protege.stanford.edu/). <https://protege.stanford.edu/>. (in Eng.).

Chung Hyun-Sook & Kim J., 2014. Semantic Model of Syllabus and Learning Ontology for Intelligent Learning System. Pp. 175–183. 10.1007/978-3-319-11289-3_18. (in Eng.).

Kose U. and Arslan A., 2017. Optimization of selflearning in Computer Engineering courses: An intelligent software system supported by Artificial Neural Network and Vortex Optimization Algorithm. *Comput Appl Eng Educ*, 25: Pp.142–156. <https://doi.org/10.1002/cae.21787>. (in Eng.).

Kurzaeva, L & Povitukhin, S & Usataya, T. & Usatiy, D. (2020). The development of ontological model for increasing the competitiveness of university graduates in information technologies. *Journal of Physics: Conference Series*. 1691. 012003. 10.1088/1742-6596/1691/1/012003. (in Eng.).

Katis, E., Kondylakis, H., Agathangelos, G., Vassilakis, K. (2018). Developing an Ontology for Curriculum and Syllabus. In: , et al. *The Semantic Web: ESWC 2018 Satellite Events. ESWC 2018. Lecture Notes in Computer Science()*, vol 11155. Springer, Cham. https://doi.org/10.1007/978-3-319-98192-5_11. (in Eng.).

V.L. Miguéis, Ana Freitas, Paulo J.V. Garcia, André Silva. (2018). Early segmentation of students according to their academic performance: A predictive modelling approach, *Decision Support Systems*, Volume 115, Pages 36-51, ISSN 0167-9236, <https://doi.org/10.1016/j.dss.2018.09.001>. (in Eng.).

Thomas Gatzka. (2021). Aspects of openness as predictors of academic achievement, *Personality and Individual Differences*, Volume 170, 110422, ISSN 0191-8869, <https://doi.org/10.1016/j.paid.2020.110422>. (in Eng.).

Y. Demchenko and L. Stoy. (2021). "Research Data Management and Data Stewardship Competences in University Curriculum," 2021 IEEE Global Engineering Education Conference (EDUCON), pp. 1717-1726, doi: 10.1109/EDUCON46332.2021.9453956. (in Eng.).

Quezada Sarmiento, Pablo & Elorriaga, Jon & Arruarte, Ana & Jumbo-Flores, Luis. (2021). Used of Web Scraping on Knowledge Representation Model for Bodies of Knowledge as a Tool to Development Curriculum. 10.1007/978-3-030-72651-5_58. (in Eng.).

Zulfiya, K., Gulmira, B., & Altynbek, S. (2019). Ontological model for student's knowledge assessment. In *ACM International Conference Proceeding Series. Association for Computing Machinery*. <https://doi.org/10.1145/3330431.3330449>. (in Eng.).

NEWS OF THE NATIONAL ACADEMY OF SCIENCES OF THE REPUBLIC OF
KAZAKHSTAN
PHYSICO-MATHEMATICAL SERIES
ISSN 1991-346X
Volume 2. Number 346 (2023). 128–140
<https://doi.org/10.32014/2023.2518-1726.189>

UDC 004.931

© **M. Bolsynbek¹, G. Abdikerimova¹, S. Serikbayeva^{1*}, A. Tanirbergenov¹,
Zh. Tashhurekova², 2023**

¹Eurasian National University named after L.N. Gumilyov, Department
of Information Systems, Astana, Kazakhstan;

²Taraz Regional University named after M.KH. Dulaty, Taraz, Kazakhstan.
E-mail: inf_8585@mail.ru

RESEARCH OF INFORMATION SYSTEMS AND METHODS OF FORECASTING SOIL AND SOIL EROSION

Bolsynbek Mukhammed Kurmanbekuly — doctoral student of the Department of Information Systems of the L.N. Gumilyov Eurasian National University, Astana, Kazakhstan
E-mail: mbolsynbek@bk.ru;

Abdikerimova Gulzira Bakhytbekovna — PhD, acting associate professor of the Department of Information Systems of L.N. Gumilyov Eurasian National University, Astana, Kazakhstan
E-mail: gulzira1981@mail.ru;

Serikbayeva Sandugash Kurmanbekovna — PhD. Senior Lecturer of the Department of Information Systems of L.N. Gumilyov Eurasian National University, Astana, Kazakhstan
E-mail: inf_8585@mail.ru;

Tanirbergenov Adilbek Zhumatayevich — associate professor of the Department of Information Systems, algebra and geometry of the L.N. Gumilyov Eurasian National University, Astana, Kazakhstan
E-mail: t.adilbek@mail.ru;

Tashhurekova Zhazira Kudaibergenovna — acting associate professor of the Department "applied informatics and programming" Taraz regional university named after M.KH. Dulaty, Taraz, Kazakhstan
E-mail: tashhurekova@mail.ru.

Abstract. Information systems and methods of forecasting soil and soil erosion are an important tool for the sustainable use of land resources. Research in this area helps to understand the mechanisms of soil erosion and develop measures to prevent it. The article describes several information systems in which it is possible to use information systems for predicting soil preparation to determine the optimal time and methods of tillage in order to increase productivity and efficiency of agricultural production, describes the main work of information systems. The methods of soil erosion control and management used in the development of an information system for predicting soil preparation are considered. In this article, we have described some methods for determining mathematical models and soil

erosion that can be used in an information system. The main methods of developing an information system for predicting soil preparation for erosion are proposed.

Keywords: information systems, agriculture, soil, soil erosion

Financing: There is no source of funding for this study.

Conflict of interest: The authors declare that there is no conflict of interest.

© **М.Қ. Болсынбек¹, Г.Б. Абдикеримова¹, С.К. Серикбаева^{1*},
А.Ж. Танирбергенов¹, Ж.К. Тасжурекова², 2023**

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

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

E-mail: inf_8585@mail.ru

ТОПЫРАҚ ЖӘНЕ ТОПЫРАҚ ЭРОЗИСЫН БОЛЖАУЖЫҢ АҚПАРАТТЫҚ ЖҮЙЕЛЕРІ МЕН ӘДІСТЕРІН ЗЕРТТЕУ

Болсынбек Мұхаммед Құрманбекұлы — Л.Н. Гумилев атындағы Еуразия ұлттық университетінің Ақпараттық жүйелер кафедрасының докторанты, Астана, Қазақстан
E-mail: mbolsynbek@bk.ru;

Абдикеримова Гульзира Бахытбековна — Л.Н. Гумилев атындағы Еуразия ұлттық университетінің Ақпараттық жүйелер кафедрасының м.а.доценты, PhD, Астана, Қазақстан
E-mail: gulzira1981@mail.ru;

Серикбаева Сандугаш Құрманбековна — Л.Н. Гумилев атындағы Еуразия ұлттық университетінің Ақпараттық жүйелер кафедрасының аға оқытушысы, PhD, Астана, Қазақстан
E-mail: inf_8585@mail.ru;

Танирбергенов Адилбек Жуматаевич — Л.Н. Гумилев атындағы Еуразия ұлттық университетінің Ақпараттық жүйелер кафедрасы, алгебра және геометрия кафедрасының доценті, Астана, Қазақстан
E-mail: t.adilbek@mail.ru;

Тасжурекова Жазира Қудайбергеновна — М.Х. Дулати атындағы Тараз өңірлік университеті, «Қолданбалы информатика және бағдарламалау» кафедрасының доцент м.а., т.ғ.к., Тараз, Қазақстан
E-mail: tashjurekova@mail.ru.

Аннотация. Топырақ пен топырақ эрозиясын болжаудың ақпараттық жүйелері мен әдістері Жер ресурстарын тұрақты пайдаланудың маңызды құралы болып табылады. Осы саладағы зерттеулер топырақ эрозиясының даму механизмдерін түсінуге және оның алдын алу шараларын жасауға көмектеседі. Мақалада ауылшаруашылық өндірісінің өнімділігі мен тиімділігін арттыру мақсатында топырақты өңдеудің оңтайлы уақыты мен әдістерін анықтау үшін топырақты дайындауды болжаудың ақпараттық жүйелерін пайдалануға болатын бірнеше ақпараттық жүйелерге сипаттама келтіріліп, ақпараттық жүйелердің негізгі жұмыстарына сипаттама келтірілді. Топырақты дайындауды болжаудың ақпараттық жүйесін әзірлеуде қолданылатын, топырақ эрозиясын бақылау және басқару әдістері қарастырылды. Бұл мақалада біз ақпараттық жүйеде қолдануға болатын, математикалық модельдерді және топырақ эрозиясын анықтаудың кейбір әдістері сипатталды. Топырақты

эрозияны дайындауды болжаудың ақпараттық жүйесін әзірлеу негізгі әдістері ұсынылды.

Түйінді сөздер: ақпараттық жүйелер, ауылшаруашылық, топырақ, топырақ эрозиясы

Қаржыландыру: Бұл зерттеу жұмысын қаржыландыру көзі жоқ.

Мүдделер қақтығысы: Авторлар осы мақалада мүдделер қақтығысы жоқ деп мәлімдемейді.

© **М.Қ. Болсынбек¹, Г.Б. Абдикеримова¹, С.К. Серикбаева^{1*},
А.Ж. Танирбергенов¹, Ж.К. Тасжурекова², 2023**

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

²Таразский региональный университет им. М.Х. Дулати, Тараз, Казахстан.
E-mail: inf_8585@mail.ru

ИССЛЕДОВАНИЕ ИНФОРМАЦИОННЫХ СИСТЕМ И МЕТОДОВ ПРОГНОЗИРОВАНИЯ ПОЧВЕННОЙ И ПОЧВЕННОЙ ЭРОЗИИ

Болсынбек Мұхаммед Құрманбекұлы — докторант кафедры информационных систем Евразийского национального университета имени Л.Н. Гумилева, Астана, Казахстан
E-mail: mbolsynbek@bk.ru;

Абдикеримова Гульзира Бахытбековна — PhD, и.о. доцент кафедры информационных систем Евразийского национального университета им. Л.Н. Гумилева, Астана, Казахстан
E-mail: gulzira1981@mail.ru;

Серикбаева Сандугаш Қурманбековна — PhD, старший преподаватель кафедры информационных систем Евразийского национального университета им. Л.Н. Гумилева, Астана, Казахстан
E-mail: inf_8585@mail.ru;

Танирбергенов Адилбек Жуматаевич — доцент кафедры алгебры и геометрии, кафедра информационных систем Евразийского национального университета им. Л.Н. Гумилева, Астана, Казахстан
E-mail: t.adilbek@mail.ru;

Тасжурекова Жазира Қудайбергеновна — к.т.н., и.о. доцента кафедры «Прикладная информатика и программирование» Таразского регионального университета им. М.Х. Дулати, Тараз, Казахстан
E-mail: tashjurekova@mail.ru.

Аннотация. Информационные системы и методы прогнозирования почвенной и почвенной эрозии являются важными инструментами для устойчивого использования земельных ресурсов. Исследования в этой области помогают понять механизмы развития почвенной эрозии и разработать меры по ее предотвращению. В статье приводятся описания нескольких информационных систем, в которых можно использовать информационные системы прогнозирования подготовки почвы для определения оптимального времени и методов обработки почвы с целью повышения производительности и эффективности сельскохозяйственного производства, приводятся описания

основных работ информационных систем. Рассмотрены методы контроля и управления эрозией почв, применяемые при разработке информационной системы прогнозирования подготовки почв. В этой статье мы описали некоторые методы определения математических моделей и эрозии почвы, которые можно использовать в информационной системе. Предложены основные методы разработки информационной системы прогнозирования подготовки почв к эрозии.

Ключевые слова: информационные системы, сельское хозяйство, почва, эрозия почвы

Финансирование: в данной исследовательской работе нет источника финансирования.

Конфликт интересов: авторы заявляют об отсутствии конфликта интересов.

Кіріспе

Топырақ экожүйенің маңызды компоненттерінің бірі болып табылады, олар өсімдіктерді қоректендіреді, су мен көміртекті сақтайды және басқа да бірқатар функцияларды орындайды. Климаттың өзгеруіне және жерді пайдалану және қоршаған ортаның ластануы сияқты басқа факторларға байланысты Қазақстанда топырақтың тозу қаупі жоғары.

Топырақтың қасиеттерін анықтайтын негізгі аспектілердің бірі-климат. Қазақстандағы климаттық жағдайлар өте алуан түрлі және әрбір климаттық аймақтың өзіндік топырақ сипаттамалары бар. Осыған байланысты Қазақстанның әртүрлі климаттық аймақтарындағы топырақты зерттеу олардың қасиеттері мен пайдалану мүмкіндіктерін түсіну үшін өте маңызды болып табылады.

Топырақ эрозиясы Қазақстанда елеулі проблема болып табылады, өйткені ел аумағының 80 %-ға жуығы эрозияның әртүрлі түрлеріне, ең алдымен су және жел эрозиясына ұшырайды. Елдің құрғақ және жартылай құрғақ климаты жерді пайдаланудың тұрақсыз әдістерімен бірге елдің топырақ ресурстарының деградациясына әкелді, Бұл ауыл шаруашылығы өнімділігінің төмендеуіне, биоәртүрліліктің жоғалуына және су тасқыны қаупінің жоғарылауына әкелді. Жаңбыр мен қардың еруінен туындаған су эрозиясы Қазақстанда эрозияның ең көп таралған түрі болып табылады, ол барлық эрозияның шамамен 70 %-приходится құрайды. Эрозияның бұл түрі жауын-шашын мөлшері топырақтың суды сіңіру қабілетінен асып, каналдардың, сайлардың және сайлардың пайда болуына әкелгенде пайда болады. Бұл каналдар мен сайлардың көлемі ұлғаюы мүмкін, сайып келгенде, жерге терең траншеяларды кесіп, дақылдарды өсіруді қиындатады. Екінші жағынан, жел эрозиясы елдегі барлық эрозияның шамамен 30 % құрайды және Қазақстанның оңтүстік және оңтүстік-шығыс аудандарында күрделі проблема болып табылады. Эрозияның бұл түрі топырақ бөлшектері қатты

желмен ауаға көтерілгенде пайда болады, нәтижесінде құм төбелері пайда болады және дақылдардың өсуіне әсер етеді. Елдегі кең көлемді тау-кен өндірісі және ормандарды кесу топырақтың тұрақсыздығына байланысты жел эрозиясының таралуына ықпал етті. Қазақстандағы топырақ эрозиясының негізгі факторлары жер пайдаланудың тұрақсыз әдістері, оның ішінде малды шамадан тыс жаю, тік беткейлерде жерді өңдеу және ормандарды кесу болып табылады. Атап айтқанда, шамадан тыс мал жаю елдің топырақ ресурстарына айтарлықтай әсер етті, өйткені ол топырақтың тығыздалуына және су мен жел эрозиясының пайда болуына ықпал ететін өсімдік жамылғысының бұзылуына әкеледі. Қазақстанда топырақ эрозиясының салдары өте ауыр: ауыл шаруашылығы өнімділігінің төмендеуі азық-түлік тапшылығына және ауыл тұрғындары арасында кедейлік деңгейінің жоғарылауына әкеледі. Сонымен қатар, топырақтың үстіңгі қабатының жоғалуы және топырақ ресурстарының деградациясы елдің экожүйесіне қатты әсер етеді, бұл топырақтың бедеулігіне және биоәртүрліліктің төмендеуіне әкеледі. Топырақ эрозиясы сонымен қатар үйлер мен инфрақұрылымды зақымдауы мүмкін су тасқыны қаупін арттырады. Қазақстандағы топырақ эрозиясына қарсы күрес Үкімет тарапынан да, жергілікті қоғамдастықтар тарапынан да келісілген күш-жігерді талап етеді. Жерді тұрақты пайдалану, орман өсіру және террасалар мен тіреу қабырғаларын салу сияқты стратегиялар эрозияның елдің топырақ ресурстарына әсерін азайтуға көмектеседі. Сонымен қатар, топырақты сақтаудың маңыздылығы және эрозияның алдын алу үшін адамдар қабылдауы мүмкін шаралар туралы хабардар болу елдің тұрақты болашағына ықпал етуі мүмкін (Матыченков, 2012).

Топырақ пен топырақ эрозиясын болжаудың ақпараттық жүйелері мен әдістерін зерттеу қазіргі экология мен ауыл шаруашылығындағы өзекті тақырып болып табылады. Топырақ эрозиясы-су, жел және басқа факторлардың әсерінен топырақ жамылғысының бұзылуы және топырақтың тасымалдану процесі.

Топырақ эрозиясын болжау үшін компьютерлік бағдарламаларды модельдеу, ландшафт пен жер жамылғысын талдау және геоақпараттық жүйелерді қолдану сияқты әртүрлі әдістер қолданылады. Бұл әдістер топырақ эрозиясының пайда болу қаупін бағалауға және оның алдын алу шараларын жасауға мүмкіндік береді.

Сонымен қатар, топырақ эрозиясын бақылау үшін қолданылатын ақпараттық жүйелер бар. Олар жер жамылғысының жағдайы туралы ақпарат береді және уақыт өте келе ондағы өзгерістерді бақылауға мүмкіндік береді. Бұл жүйелер ғылыми зерттеулер үшін де, ауыл шаруашылығында шешім қабылдау үшін де қолданылады.

Осылайша, топырақ пен топырақ эрозиясын болжаудың ақпараттық жүйелері мен әдістерін зерттеу жер жамылғысын сақтау және жер ресурстарын үнемдеу үшін маңызды міндет болып табылады.

Топырақ пен топырақты эрозияға дайындауды болжаудың қолданыстағы ақпараттық жүйелеріне шолу

Қазіргі ақпараттық жүйелер (АЖ) топырақ ресурстарын талдау мен басқарудың қуатты құралы болып табылады. Олар топырақ жағдайы мен оның қасиеттері туралы мәліметтер алуға, климаттық жағдайларды, географиялық деректерді және топырақ жағдайына әсер ететін басқа факторларды талдауға мүмкіндік береді. АЖ сонымен қатар деректерді талдау және әртүрлі модельдер негізінде топырақ ресурстарындағы өзгерістерді болжауға мүмкіндік береді.

Топырақ ресурстарын басқару үшін қолданылатын заманауи ақпараттық жүйелердің ішінде SoilGrids, globalsoilmap және басқалары бар. Олар геоақпараттық технологияларды қолдануға негізделген және топырақ ресурстарын талдау және басқару үшін қажетті мәліметтердің кең спектрін ұсынады (Акшалов, 2010).

Ауылшаруашылық өндірісінің өнімділігі мен тиімділігін арттыру мақсатында Топырақты өңдеудің оңтайлы уақыты мен әдістерін анықтау үшін топырақты дайындауды болжаудың ақпараттық жүйелерін пайдалануға болады. Мұндай жүйелердің бірнеше мысалдары төменде келтірілген:

AgroClimate-Флорида Университеті әзірлеген жүйе, ол АҚШ-тағы әрбір аймақ үшін ауа райы жағдайлары, топырақ жағдайы және оңтайлы өңдеу тәжірибелері бойынша нұсқаулар береді.

SoilGrids-жер бетіндегі әрбір нүкте үшін органикалық заттар, рН және құрылым сияқты топырақ сипаттамалары туралы ақпарат беретін жаһандық топырақ деректер базасы.

Agro-MetShell-бұл гидрология және метеорология институты жасаған жүйе В.Н. Сукачева Ресейдегі ауылшаруашылық кәсіпорындары үшін ауа-райы, топырақ жағдайы және топырақты өңдеу бойынша ұсыныстар беретін жүйе әзірленген.

CropProphet-Gro Intelligence компаниясы әзірлеген жүйе, ол өнімділікті болжау және АҚШ-тағы ауылшаруашылық кәсіпорындары үшін топырақты өңдеудің оңтайлы әдістерін анықтау үшін машиналық оқыту мен деректерді талдауды пайдаланады.

ADAPT — Дүниежүзілік азық-түлік және ауыл шаруашылығы ұйымы (ФАО) әзірлеген жүйе, ол бүкіл әлем бойынша ауыл шаруашылығы кәсіпорындары үшін ауа райы жағдайлары, топырақ жағдайлары және топырақты өңдеу бойынша ұсыныстар туралы ақпарат береді.

AgroAPI – АӨК мен фермерлерге тыңайтқыштарды пайдалануды оңтайландыруға және топырақ өңдеуге қажетті техника түрлерін таңдауға көмектесетін ақпараттық жүйе.

Storio – ауыл шаруашылығы кәсіпорындарына тыңайтқыштар, өсімдіктерді қорғау, сондай-ақ қолайлы топырақ өңдеу жабдықтарын таңдау саласында шешім қабылдауға көмектесетін дақылдарды бақылаудың біріктірілген жүйесі.

AgroNavigator – фермерлерге топырақ өңдеудің оңтайлы шешімдерін

таңдауға, оңтайлы ауыспалы егіс үлгілерін таңдауға, сондай-ақ егіс пен егін жинаудың оңтайлы сәттерін анықтауға көмектесетін сервис.

Climate FieldView - бұл ауа райы жағдайлары, топырақ жағдайлары, сондай-ақ өсімдіктер мен дақылдар туралы деректер беретін ақпараттық жүйе. Сондай-ақ тыңайтқыштарды пайдалануды оңтайландыруға және топырақ өңдейтін техниканың қолайлы түрлерін таңдауға көмектеседі.

FieldNET – фермерлерге суды пайдалануды азайтуға және топырақ өңдеудің ең жақсы шешімдерін таңдауға көмектесетін суаруды автоматтандыру және топырақтың ылғалдылығын басқару жүйесі.

SoilGrids: Бұл жер бетіндегі топырақ қасиеттері туралы жоғары сапалы деректерді қамтамасыз ететін топырақ туралы ақпараттың дүниежүзілік дерекқоры. Онда топырақтың физикалық, химиялық және биологиялық қасиеттері туралы мәліметтер, сондай-ақ топырақ эрозиясының болжамдары бар.

APEX (Agricultural Policy Environmental eXtender): бұл ауылшаруашылық тәжірибелерінің топырақ, су және ауа сапасына әсерін бағалауға арналған модель. Бұл топырақ эрозиясын және оның қоршаған ортаға әсерін болжауға мүмкіндік береді.

WEPP (water Erosion Prediction Project): бұл ауылшаруашылық жерлеріндегі топырақ эрозиясын болжауға арналған модель. Ол ауылшаруашылық тәжірибелерінің топырақ эрозиясына әсерін бағалау және оны азайту стратегияларын әзірлеу үшін қолданылады.

RUSLE2 (Revised Universal Soil Loss Equation 2): бұл әртүрлі егіншілік жағдайында топырақ эрозиясын болжауға арналған модель. Ол әртүрлі ауылшаруашылық тәжірибелерінің топырақ эрозиясына әсерін бағалау және оны азайту стратегияларын әзірлеу үшін қолданылады.

SWAT (Soil and Water Assessment Tool): бұл ауылшаруашылық тәжірибелерінің су мен топырақ сапасына әсерін бағалауға арналған модель. Бұл топырақ эрозиясын және оның қоршаған ортаға әсерін болжауға, сондай-ақ топырақ эрозиясын азайту үшін ең жақсы егіншілік тәжірибесін анықтауға мүмкіндік береді.

Бұл ақпараттық жүйелер топырақтың дайындалуы мен топырақ эрозиясын болжау үшін, сондай-ақ топырақ эрозиясын азайту және ауылшаруашылық жерлеріндегі топырақ пен судың сапасын жақсарту стратегияларын әзірлеу үшін қолданылады (Скрябина, 2013).

Бұл топырақты дайындауды болжаудың қолданыстағы ақпараттық жүйелеріне шағын шолу ғана. Әрбір ауылшаруашылық кәсіпорнының нақты қажеттіліктері мен жағдайларына сәйкес келетін басқа жүйелер бар.

Топырақты дайындауды болжаудың ақпараттық жүйесін әзірлеуде қолданылатын зерттеу әдістерінің сипаттамасы

Топырақ эрозиясының ақпараттық жүйелері мен әдістері топырақ эрозиясын бақылауға және басқаруға көмектеседі және тиісті бақылау

шараларының орындалуын қамтамасыз етеді (Engelen, 1995). Бұл Ақпараттық жүйелер топырақ эрозиясының жылдамдығы туралы нақты және өзекті ақпаратты ұсынады, сонымен қатар эрозия орын алатын жерлерді анықтайды, бұл тиісті бақылау шараларын жоспарлауға және жүзеге асыруға көмектеседі. Топырақ эрозиясын бақылау мен басқарудың кейбір әдістері мен ақпараттық жүйелері:

1. ГАЖ (географиялық ақпараттық жүйелер) негізіндегі топырақ эрозиясын бағалау модельдері: бұл модельдер ландшафтқа топырақ эрозиясының әсерін картаға түсіру және бағалау үшін географиялық ақпараттық жүйелерді (ГАЖ) пайдаланады. Олар топырақ эрозиясының көздерін, жолдарын және салдарын анықтауға, сондай-ақ топырақты сақтау стратегияларының тиімділігін анықтауға көмектеседі. ГАЖ негізіндегі топырақ эрозиясын бағалау үлгілерінің мысалдарына RUSLE (қайта қаралған әмбебап топырақ жоғалту теңдеуі) және WEPP (су эрозиясын болжау жобасы) жатады.

2. Қашықтықтан зондтау. Қашықтықтан зондтау топырақ эрозиясының ауқымы мен ауырлығын анықтау үшін спутниктік суреттер, аэрофотосуреттер және басқа көздер деректерін пайдалануды қамтиды. Бұл әдіс әсіресе жету қиын немесе шалғай жерлерде топырақ эрозиясының жылдамдығын бағалау үшін пайдалы.

3. Далалық бақылау (мониторинг): далалық бақылау топырақ эрозиясының дәрежесін тікелей өлшеуді қамтиды. Мұны шөгінділерді ағызу үшін эрозиялық түйреуіштерді, тұндырғыштарды және бассейндерді пайдалану арқылы жасауға болады. Далалық бақылау топырақ эрозиясын бағалау үлгілерін калибрлеуге және тексеруге көмектесетін маңызды ақпаратты береді.

4. Топырақ туралы мәліметтер базасы. Топырақ туралы мәліметтер базасы топырақ эрозиясын модельдеуге пайдалы топырақтың сипаттамалары, заңдылықтары мен тенденциялары туралы ақпарат береді. Мәліметтер базасын топырақтың гидрологиялық қасиеттерін, мысалы, инфильтрация, ағын және ылғал сыйымдылығы, сондай-ақ топырақтың физикалық және химиялық қасиеттерін бағалау үшін пайдалануға болады.

5. Эрозия мен шөгінділерді бақылау жоспарлары: бұл жоспарлар топырақ эрозиясының алдын алу немесе бақылау үшін қажетті тәжірибелер мен шараларды анықтау үшін қолданылады.

Топырақ эрозиясын анықтау Қазақстанның топырақ-климаттық аймақтарында топырақтың дайындалуын болжау үшін әзірленген ақпараттық жүйенің маңызды құрамдас бөлігі болып табылады.

Бұл әр аймақтағы топырақ эрозиясының деңгейін болжауға және оны азайту үшін ұсыныстар жасауға мүмкіндік береді. Бұл мақалада біз ақпараттық жүйеде қолдануға болатын топырақ эрозиясын анықтаудың кейбір әдістерін қарастырамыз.

1. USLE (Universal Soil Loss Equation) әдісі-эртүрлі климаттық жағдайларда жел мен судың әсерінен топырақтың жоғалуын бағалау үшін қолданылады.

Ол бес факторға негізделген: желдің күші немесе жаңбырдың қарқындылығы, көлбеу көлбеу ұзындығы, көлбеу көлбеу, топырақ түрі және жабын дәрежесі. Әрбір фактор аймақтың ерекшеліктеріне сәйкес бағаланады және нәтижелер топырақ шығындарының жалпы бағасын алу үшін біріктіріледі.

2. RUSLE (Revised Universal Soil Loss Equation) әдісі-жаңа деректер мен технологияларды ескеретін USLE әдісінің жаңартылған нұсқасы. Ол сондай-ақ бес факторды пайдаланады, бірақ қашықтан зондтау және ГАЗ арқылы алынған деректерді пайдалана отырып, оларды дәлірек бағалайды. Ол сондай-ақ егіншілік және ағаш кесу сияқты адам әсерінен туындаған жер жамылғысының өзгеруін ескереді.

3. Wepp (water Erosion Prediction Project) әдісі-жаңбыр эрозиясынан туындаған топырақ шығынын бағалау үшін қолданылады. Ол топырақтың ерекшеліктерін, өсімдік түрін және климаттық жағдайларды ескереді. WEPP дәлірек нәтиже алу үшін гидрологиялық процестер мен ГАЗ модельдеуін қолданады.

4. GLEAMS (Groundwater Loading Effects of Agricultural Management Systems) әдісі-топырақ эрозиясын және жаңбыр эрозиясы мен жер үсті ағынының әсерінен агрохимикаттардың жоғалуын анықтау үшін қолданылады. Ол топырақтың сипаттамалары, топографиясы, климаттық жағдайлары, өсімдік жамылғысының түрі және егіншілік әдістері сияқты көптеген факторларды ескереді

Олар эрозияны бақылау, жайылымдардағы қысымды төмендету немесе ағынды суларды азайту үшін буферлік жолақтарды салу сияқты топырақтың жоғалуын азайтудың нақты стратегияларын сипаттайды. Осылайша, топырақ эрозиясының ақпараттық жүйелері мен әдістері топырақ эрозиясын басқаруда маңызды рөл атқарады. Олар топырақ эрозиясының дәрежесін бағалауға, оның себептерін анықтауға, топырақты сақтау стратегияларының тиімділігін бағалауға және топырақты басқарудың болашақ тәжірибесін жоспарлауға көмектеседі (Tempel, 2002). Осы әдістер мен жүйелерді біріктіру арқылы топырақ эрозиясын бақылау және топырақ ресурстарының болашақ ұрпақ үшін өнімді және сау болып қалуын қамтамасыз ету үшін тиімді шараларды әзірлеуге болады.

Топырақ эрозиясы ауыл шаруашылығы мен қоршаған орта үшін үлкен проблема болып табылады. Математикалық модельдер топырақ эрозиясының механизмдерін түсінуге және оның әсерін болдырмаудың немесе азайтудың ең тиімді әдістерін анықтауға көмектеседі.

Топырақ эрозиясын сипаттау үшін қолданылатын бірнеше математикалық модельдер бар. Олардың кейбіреулері эрозия кезінде болатын физикалық процестерді сипаттаса, басқа модельдер статистикалық деректерге негізделген (Finke, 2001).

Ең танымал математикалық модельдердің бірі 1960 жылдары АҚШ-та жасалған әмбебап топырақ жоғалту теңдеуі (USLE) болып табылады. Бұл

модель топырақ эрозиясының жылдамдығын анықтау үшін еңіс, топырақ түрі, мәдени тәжірибе және жауын-шашын сияқты факторларды ескереді.

Басқа модель — Куслик моделі, ол уақыт бойынша топырақ құрамы мен құрылымының өзгеруін, сондай-ақ өзгермелі климаттық жағдайлардың эрозия процестеріне әсерін ескереді.

Моррисон моделі белгілі бір аймақтағы топырақ эрозиясының жылдамдығын анықтау үшін жер бедерін, өсімдіктерді, топырақ түрін, жауын-шашынды және басқа факторларды пайдаланады.

Жалпы, математикалық модельдер топырақ эрозиясын зерттеу мен бақылауда маңызды рөл атқарады. Олар тәуекелдерді бағалауға және эрозия процестерінің алдын алу және басқарудың ең тиімді жолдарын анықтауға көмектеседі.

Жоғарыда аталған әдістерге сүйене отырып, топырақты дайындаудың топырақ эрозиясын болжаудың ақпараттық жүйесін әзірлеу келесі әдістерді қамтуы мүмкін:

Математикалық модельдеу әдістері: бұл жағдайда топырақ эрозиясына әсер ететін әртүрлі параметрлер мен факторларды ескеретін математикалық модель жасалады. Бұл модель физикалық заңдарға, статистикалық әдістерге немесе машиналық оқытуға негізделуі мүмкін.

Географиялық ақпараттық жүйелер (ГАЗ): ГАЗ топография, көлбеу, топырақ түрі және климаттық жағдайлар сияқты топырақ параметрлері туралы Кеңістіктік деректерді жинау, сақтау және талдау үшін қолданылады. Бұл деректерді әртүрлі аймақтардағы топырақ эрозиясының әлеуетін болжау үшін пайдалануға болады (Клебанович, 2013).

Растрлық талдау әдістері: растрлық талдау кеңістіктік деректерді растрлық кескін түрінде талдауға мүмкіндік береді. Топырақ эрозиясын болжау контекстінде эрозия қаупі жоғары аймақтарды анықтау және эрозияның алдын алудың оңтайлы шараларын анықтау үшін растрлық талдауды қолдануға болады.

Қашықтықтан зондтау жүйелері: спутниктік немесе авиациялық зондтау арқылы алынған деректерді топырақтың күйін бағалау және оның эрозияға бейімділігін болжау үшін пайдалануға болады. Мұндай деректер өсімдік жамылғысы, өсімдік жамылғысының тығыздығы, топырақтың ылғалдылығы және эрозияға байланысты басқа параметрлер туралы ақпаратты қамтуы мүмкін.

Сараптамалық жүйелер: сараптамалық жүйелерді әзірлеу топырақ эрозиясын болжау үшін сараптамалық білім мен ережелерді біріктіруге мүмкіндік береді. Сараптамалық жүйелерді кірістерді талдау, сарапшылардың білімін қолдану және эрозияның алдын алу бойынша ұсыныстар беру үшін пайдалануға болады.

Статистикалық талдау әдістері: статистикалық әдістерді топырақ параметрлері туралы деректерді және олардың топырақ эрозиясымен байла-

нысын талдау және модельдеу үшін қолдануға болады. Бұған корреляциялық талдау, регрессиялық талдау және басқа статистикалық талдау кіруі мүмкін.

Машиналық оқыту әдістері: Машиналық оқыту алгоритмдерін топырақ параметрлері мен топырақ эрозиясы туралы үлкен көлемдегі деректерді талдау және өңдеу үшін пайдалануға болады. Бұл әдістер әртүрлі факторларға байланысты үлгілерді анықтауға және эрозия ықтималдығын болжауға мүмкіндік беретін жіктеуді, регрессияны, кластерлеуді және басқа әдістерді қамтуы мүмкін.

Гидрологиялық модельдеу әдістері: гидрологиялық модельдерді сайттағы су ағынын және оның топырақ эрозиясына әсерін талдау үшін пайдалануға болады. Бұл модельдер эрозия процестерін болжау үшін топографияны, көлбеуді, топырақ түрлерін, климаттық жағдайларды және басқа параметрлерді қарастыруы мүмкін.

Инженерлік геология әдістері: топырақтың геологиялық құрамын, топырақтың күйін және учаскенің геологиялық құрылымын зерттеуді қамтиды. Бұл деректерді топырақтың тұрақтылығын анықтау және оны эрозиядан қорғау үшін шаралар қабылдау үшін пайдалануға болады.

Желді модельдеу әдістері: егер топырақ эрозиясы жел процестерінен туындаса, желдің бағыты мен жылдамдығын анықтау және эрозиялық материалдардың қозғалысын болжау үшін желді модельдеуді қолдануға болады (Клебанович, 2015).

Қашықтықтан бақылау әдістері: сайтқа орналастырылған сенсорлар мен сенсорлардың көмегімен топырақтың ылғалдылығы, желдің жылдамдығы, жауын-шашын мөлшері және эрозияға байланысты басқа факторлар сияқты әртүрлі параметрлер туралы мәліметтер алуға болады. Бұл деректерді топырақ жағдайын бақылау және талдау және эрозия процестерін болжау үшін пайдалануға болады.

Экологиялық модельдеу әдістері: экожүйедегі экологиялық процестер мен өзара әрекеттесулерді модельдеу топырақ эрозиясын болжауға көмектеседі. Бұл модельдер топырақтың тұрақтылығына және эрозияның алдын алуға әсер етуі мүмкін өсімдіктер, микроорганизмдер мен жануарларды қоса алғанда, биологиялық факторларды ескереді.

Геоморфологиялық талдау әдістері: Жердің пішіні, топырақ түрлері және гидрологиялық сипаттамалары сияқты геоморфологиялық деректерді топырақтың эрозияға осалдығын бағалау және ықтимал эрозия аймақтарын болжау үшін пайдалануға болады.

Өсімдік жамылғысын модельдеу әдістері: өсімдік жамылғысын және оның топыраққа әсерін модельдеу эрозияны болжауға көмектеседі. Бұл модельдер өсімдік түрлерін, өсімдік жамылғысының тығыздығы мен таралуын, оның топырақпен және басқа факторлармен өзара әрекеттесуін ескереді.

Жерді пайдаланудағы өзгерістерді талдау әдістері: ормандарды кесу, қарқынды ауыл шаруашылығы немесе құрылыс сияқты жерді пайдаланудағы

өзгерістер топырақтың эрозияға осалдығының жоғарылауына әкелуі мүмкін. Бұл өзгерістерді талдау эрозия қаупі жоғары аймақтарды анықтауға және оларды қорғау үшін тиісті шараларды әзірлеуге көмектеседі.

Деректерді жинау және біріктіру: топырақ эрозиясын сәтті болжау үшін географиялық, климаттық, топырақтық және басқа да тиісті деректерді қоса алғанда, әртүрлі деректер түрлерін жинау және біріктіру қажет. Бұл мәліметтер базасын, сенсорлық желілерді, тарихи деректерді, зерттеу нәтижелерін және басқа ақпарат көздерін пайдалануды қамтуы мүмкін.

Модельдерді тексеру және тексеру: әзірленген модельдер мен әдістер сенімді деректер мен белгілі зерттеу нәтижелерін пайдалана отырып тексеріліп, тексерілуі керек. Бұл топырақ эрозиясын болжау кезінде олардың дәлдігі мен сенімділігін қамтамасыз етеді.

Пайдаланушы интерфейсінің интеграциясы: ақпараттық жүйе пайдаланушыларға қол жетімді және ыңғайлы болуы керек. Пайдаланушыға ыңғайлы интерфейсті біріктіру пайдаланушыларға жүйемен өзара әрекеттесуге, болжау нәтижелерін көруге, ұсыныстар алуға және топырақ эрозиясының алдын алу үшін тиісті шараларды қабылдауға мүмкіндік береді.

Мониторинг және ескерту жүйесін әзірлеу: мониторинг және ескерту жүйесін әзірлеу үшін ақпараттық жүйені кеңейтуге болады. Бұл топырақ эрозиясының қаупін уақытылы анықтауға және пайдаланушыларға зақымданудың алдын алу үшін нұсқаулар мен нұсқаулар беруге мүмкіндік береді.

Жүйені орнату және жаңарту: әзірленген ақпараттық жүйе тиісті инфрақұрылымға орнатылып, жаңа деректерді, ғылыми зерттеулер мен технологиялық әзірлемелерді ескере отырып, үнемі жаңартылып отыруы керек.

Осы әдістер мен қадамдардың барлығы топырақ эрозиясын болжау мен топырақты дайындаудың нақты жағдайлары мен талаптарына бейімделуі керек. Топырақ эрозиясын болжаудың ақпараттық жүйесін әзірлеуде қолда бар деректерге, Жобаның мақсаттары мен талаптарына байланысты жоғарыда аталған әдістердің комбинациясы қолданылуы мүмкін.

Қорытынды

Топырақ пен топырақ эрозиясын болжаудың ақпараттық жүйелері мен әдістерін зерттеу ауыл шаруашылығы мен экология үшін үлкен маңызға ие маңызды міндет болып табылады. Топырақ эрозиясы-құнарлы топырақ қабатының жоғалуына, топырақ сапасының төмендеуіне және өнімділіктің төмендеуіне әкелетін ауыл шаруашылығының негізгі мәселелерінің бірі. Ақпараттық жүйелер мен болжау әдістері топырақ эрозиясына байланысты процестерді зерттеуге және болжауға, сондай-ақ оның алдын алу және онымен күресудің тиімді стратегияларын жасауға көмектеседі. Бұл жұмыста ақпараттық жүйелерді зерттеудің және топырақ пен топырақ эрозиясын болжаудың негізгі әдістері қарастырылды. Келешекте зерттелінген әдістерге сүйене отырып, топырақ және топырақ эрозиясын зерттеуге арналған ақпараттық жүйені құру барысында деректер қоры жинақталады, машиналық

оқыту әдістерін қолдана отырып, автоматтандырылған қосымша әзірленетін болады.

ӘДЕБИЕТТЕР

Матыченков Д.В., 2012. Информационно-логические схемы банка данных программно-информационного комплекса по оптимизации использования почвенных ресурсов Республики Беларусь / Д.В. Матыченков, Г.С. Цытрон, В.В. Северцов // Почвоведение и агрохимия. – 2012. – № 2(49). – С. 49–57.

Акшалов К.А., 2010. Методические рекомендации по изучению и контролю водной эрозии почв в современных системах земледелия. - Астана: Научно-производственный центр зернового хозяйства им. А.И. Бараева, 2010. - 20 с.

«Эрозия и охрана почв»: Методические указания / О.А. Скрыбина, Н.В. Флягина. Пермь: Изд-во ПГСХА, 2013. – 43 с.

A. Hartemink, S. Nortcliff, D. Dent, 2008. Soil - The living skin of planet earth. – Soil Flyer IYPE 2008. - IUSS, 2008. - 2p.

Global and National Soils and Terrain Digital Databases (SOTER). / Ed. V.W.P. van Engelen and T.T. Wen. –International Soil Reference and Information Centre, 1995, - 138 p.

P. Tempel, 2002. SOTER - Global and National Soils and Terrain Digital Databases. Database Structure v3. -Working paper No. 02/01, September 2002. – 93 p.

P. Finke, R. Hartwich, R. Dudal, J. Ibáñez, M. Jamagne, D. King, L. Montanarella, N. Yassoglou, 2001. Georeferenced Soil Database for Europe. Manual of procedures. Version 1.1. - Edited by European Soil Bureau, 2001. - 178 p.

Клебанович Н.В., 2013. Почвоведение и земельные ресурсы : учеб. пособие / Н. В. Клебанович. - Минск : БГУ. 2013. - 343 с.

Клебанович Н.В., 2015. Интерактивная генерализация в среде ArcGIS как основной способ создания цифровых разномасштабных почвенных карт / Н. В. Клебанович, С. Н. Прокопович, А. И. Чаюк // Земля Беларуси. - 2015. - № 2. - С. 42–47.

REFERENCES

Matychenkov D.V., 2012. Information and logic schemes of the data bank of the software and information complex for optimizing the use of soil resources of the Republic of Belarus / D.V. Matychenkov, G.S. Tsytron, V.V. Severtsov // Soil science and agrochemistry. – 2012. – № 2(49). – Pp. 49–57.

Akshalov K.A., 2010. Methodological recommendations for the study and control of water erosion of soils in modern farming systems.- Astana: Scientific and Production Center of grain economy named after A.I. Barayev, 2010.- 20 p.

"Erosion and soil protection": Methodological guidelines / O.A. Scriabina, N.V. Flyagina. Perm: Publishing house of the PGSHA, 2013. – 43 p.

A. Hartemink, S. Nortcliff, D. Dent Soil, 2008. The living skin of planet earth. – Soil Flyer IYPE 2008. - IUSS, 2008. - 2p.

Global and National Soils and Terrain Digital Databases (SOTER). / Ed. V.W.P. van Engelen and T.T.Wen. –International Soil Reference and Information Centre, 1995, - 138 p.

P. Tempel. SOTER, 2002. Global and National Soils and Terrain Digital Databases. Database Structure v3. -Working paper No. 02/01, September 2002. – 93 p.

P. Finke, R. Hartwich, R. Dudal, J. Ibáñez, M. Jamagne, D. King, L. Montanarella, N. Yassoglou, 2001. Georeferenced Soil Database for Europe. Manual of procedures. Version 1.1. - Edited by European Soil Bureau, 2001. - 178 p.

Klebanovich N.V., 2013. Soil science and land resources : textbook. manual / N.V. Klebanovich. - Minsk : BSU. 2013. - 343p.

Klebanovich N.V., 2015. Interactive generalization in the ArcGIS environment as the main way to create digital multi-scale soil maps / N.V. Klebanovich, S.N. Prokopovich, A.I. Chayuk // Land of Belarus. - 2015. - № 2. -Pp. 42–47.

NEWS OF THE NATIONAL ACADEMY OF SCIENCES OF THE REPUBLIC OF KAZAKHSTAN
PHYSICO-MATHEMATICAL SERIES

ISSN 1991-346X

Volume 2. Number 346 (2023). 141–153

<https://doi.org/10.32014/2023.2518-1726.190>

UDC 004-93

© **L. Zholshiyeva^{1*}, T. Zhukabayeva¹, Sh. Turaev², M. Berdieva³,
B. Khu Ven-Tsen⁴, 2023**

¹Astana International University, Eurasian National University named
after L.N. Gumilyov, Astana, Kazakhstan;

²College of Information Technology, United Arab Emirates University,
Al Ain, United Arab Emirates;

³South Kazakhstan Medical Academy, Shymkent, Kazakhstan;

⁴M. Auezov South Kazakhstan University, Shymkent, Kazakhstan.

E-mail: lazzat.zhol.81@gmail.com

DEVELOPMENT OF AN INTELLECTUAL SYSTEM FOR RECOGNIZING KAZAKH DACTYL GESTURES BASED ON LSTM AND GRU MODELS

Zholshiyeva Lazzat Zulpuharkyzy — PhD student. Astana International University. Astana, Kazakhstan

E-mail: lazzat.zhol.81@gmail.com. ORCID: 0000-0002-2526-8471;

Zhukabayeva Tamara Kokenovna — PhD, assoc. Professor. Astana International University, Eurasian National University named after L.N. Gumilyov. Astana, Kazakhstan

E-mail: tamara_kokenovna@mail.ru. ORCID: 0000-0001-6345-5211;

Turaev Sherzod — PhD, assoc. Professor. College of Information Technology, United Arab Emirates University Al Ain, United Arab Emirates

E-mail: sherzod@uaeu.ac.ae. ORCID: 0000-0001-6661-8469;

Berdieva Meruert Aimambetovna — PhD. South Kazakhstan Medical Academy, Shymkent, Kazakhstan

E-mail: meruert_berdieva@mail.ru. ORCID: 0009-0008-5259-7618;

Khu Ven-Tsen Boris Aleksandrovich — Doctor of technical science, professor. M. Auezov South Kazakhstan University. Shymkent, Kazakhstan;

E-mail: gbcba@bb.mail.ru. ORCID: 0009-0001-0824-8858.

Abstract. The paper deals with the problem of automation of Kazakh Sign Language by converting gestures into text, and proposes two possible approaches to the solution. There is given the method for implementing the recognition of dynamic Kazakh dactyl signs in a video stream based on visual analysis of the position of hands with the positioning of key points. MediaPipe has been applied to read and position key points from images. Advanced versions of recurrent neural networks have been described to detect dactyl gestures from images, which can be used as sign classifiers. For gesture recognition, own datasets are prepared, which allow achieving high efficiency and accuracy of recognition by accepted models in

real time. The specified dataset consists of dynamic gestures of the Kazakh Sign Language. The advantage of the considered approaches is the relative simplicity of the proposed methods in a computational sense, so that their implementation does not require the use of high-performance computer technology and special equipment. The recognition accuracy of dynamic dactyl gestures of the Kazakh Sign Language was 85 % and 100 %, respectively, in the conducted experimental studies using the proposed LSTM and GRU models. In this case, recognition was carried out in real time using a conventional computer camera. Moreover, the recognition of images of same gestures obtained using a webcam has been tested. The results of the used LSTM and GRU models was developed with recognition evaluation of dynamic gestures of the Kazakh Sign Language.

Keywords: Kazakh Sign Language, Hand Gesture Recognition, Dynamic Gestures, Media Pipe, LSTM, GRU

© Л.З. Жолшиева^{1*}, Т.К. Жукабаева¹, Ш. Тураев², М.А. Бердиева³,
Б.А. Ху Вен-Цен⁴, 2023

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

²Ақпараттық технологиялар колледжі, Біріккен Араб Әмірліктері университеті, Әл-Айн, Біріккен Араб Әмірліктері;

³Оңтүстік Қазақстан медициналық академиясы, Шымкент, Қазақстан;

⁴М. Әуезов атындағы оңтүстік Қазақстан университеті, Шымкент, Қазақстан.
E-mail: lazzat.zhol.81@gmail.com

LSTM ЖӘНЕ GRU ҮЛГІЛЕРІ НЕГІЗІНДЕ ҚАЗАҚ ДАКТИЛЬДЕРІН ТАНУДЫҢ ИНТЕЛЛЕКТУАЛДЫ ЖҮЙЕСІН ҚҰРУ

Жолшиева Лаззат Зулпухаркызы — PhD студент. Астана халықаралық университеті.
Астана, Қазақстан

E-mail: lazzat.zhol.81@gmail.com. ORCID: 0000-0002-2526-8471;

Жукабаева Тамара Кокеновна — PhD, қауымдастырылған профессор. Астана халықаралық университеті. Л.Н. Гумилев атындағы Евразия ұлттық университеті. Астана, Қазақстан

E-mail: tamara_kokenovna@mail.ru. ORCID: 0000-0001-6345-5211;

Тураев Шерзод — PhD, қауымдастырылған профессор. Ақпараттық технологиялар колледжі, Біріккен Араб Әмірліктері университеті, Әл-Айн, Біріккен Араб Әмірліктері

E-mail: sherzod@uaeu.ac.ae. ORCID: 0000-0001-6661-8469;

Бердиева Мерuert Аймамбетовна — PhD. Оңтүстік Қазақстан медициналық академиясы, Шымкент, Қазақстан

E-mail: meruert_berdieva@mail.ru. ORCID: 0009-0008-5259-7618;

Ху Вен-Цен Борис Александрович — техника ғылымдарының докторы, профессор. М. Әуезов атындағы оңтүстік Қазақстан университеті, Шымкент, Қазақстан

E-mail: gbcba@bb.mail.ru. ORCID: 0009-0001-0824-8858.

Аннотация. Мақалада қазақ ым тілін мәтінге айналдыру арқылы қазақша дактильді қимылдарды автоматтандыру мәселесі қарастырылып, оны шешудің екі ықтимал тәсілі ұсынылады. Негізгі нүктелердің орналасуы мен қолдардың

орналасуын визуалды талдауға негізделген бейне ағынындағы динамикалық қазақ дактилдерін тану әдісі ұсынылған. Кескіндерден негізгі нүктелерді оқып, орналастыру үшін MediaPipe құралдар жүйесі пайдаланылды. Қимыл кескіндерін анықтау үшін белгі классификаторлары ретінде қайталанатын нейрондық желілердің жетілдірілген нұсқалары пайдаланылады. Қимылдарды тану үшін, нақты уақыт режимінде қабылданған үлгілерді пайдалана отырып жоғары тиімділік пен тану дәлдігіне қол жеткізуге мүмкіндік беретін жеке деректер жинақтары дайындалды. Аталған деректер жиынтығы қазақ дактилдінің динамикалық қимылдарынан тұрады. Қарастырылып отырған тәсілдердің артықшылығы — ұсынылған әдістердің есептеу тұрғысынан алғанда салыстырмалы түрде қарапайымдылығы болып табылады, сондықтан да оларды іске асыру өнімділігі жоғары компьютерлер мен арнайы жабдықты қолдануды қажет етпейді. Ұсынылған LSTM және GRU үлгілері арқылы жүргізілген эксперименттік зерттеулер нәтижесінде қазақ динамикалық дактилдік қимылдарын тану дәлдіктері сәйкесінше 85 % және 100 % құрады. Бұл жағдайда тану кәдімгі компьютерлік бейнекамера көмегімен нақты уақыт режимінде жүзеге асырылды. Сондай-ақ, веб-камера арқылы алынған бірдей қимылдардың кескіндерін тану сыналды. Зерттеу нәтижелері бойынша қазақ дактилінің динамикалық ым-ишарасын тануға баға бере отырып, LSTM және GRU үлгілерін қолдану нәтижелері көрсетілді.

Түйін сөздер: қазақ ым тілі, қол қимылын тану, динамикалық қимылдар, Media Pipe, LSTM, GRU

© Л.З. Жолшиева^{1*}, Т.К. Жукабаева¹, Ш. Тураев², М.А. Бердиева³,
Б.А. Ху Вен-Цен⁴, 2023

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

²Колледж информационных технологий, Университет Объединенных Арабских Эмиратов, Эль-Айн, Объединенные Арабские Эмираты;

³Южно-Казахстанская медицинская академия, Шымкент, Казахстан; ⁴Южно-Казахстанский университет М. Ауезова, Шымкент, Казахстан.
E-mail: lazzat.zhol.81@gmail.com

РАЗРАБОТКА ИНТЕЛЛЕКТУАЛЬНОЙ СИСТЕМЫ РАСПОЗНАВАНИЯ КАЗАХСКИХ ДАКТИЛЬНЫХ ЖЕСТОВ НА ОСНОВЕ МОДЕЛЕЙ LSTM И GRU

Жолшиева Лаззат Зулпухаркызы — PhD студент. Международный университет Астана. Астана, Казахстан

E-mail: lazzat.zhol.81@gmail.com. ORCID: 0000-0002-2526-8471;

Жукабаева Тамара Кокеновна — PhD, ассоциированный профессор. Международный университет Астана. Евразийский национальный университет им. Л.Н. Гумилева. Астана, Казахстан

E-mail: tamara_kokenovna@mail.ru. ORCID: 0000-0001-6345-5211;

Тураев Шерзод — PhD, ассоциированный профессор. Колледж информационных технологий. Университет Объединенных Арабских Эмиратов, Аль-Айн, Объединенные Арабские Эмираты E-mail: sherzod@uaeu.ac.ae. ORCID: 0000-0001-6661-8469;

Бердиева Меруерт Аймамбетовна — PhD. Южно-Казахстанская медицинская академия, Шымкент, Казахстан

E-mail: meruert_berdieva@mail.ru. ORCID: 0009-0008-5259-7618;

Ху Вен-Цен Борис Александрович — доктор технических наук, профессор. Южно-Казахстанский университет имени М. Ауэзова, Шымкент, Казахстан

E-mail: gbcbba@bb.mail.ru. ORCID: 0009-0001-0824-8858.

Аннотация. Рассматривается проблема автоматизированного перевода дактильного казахского языка с преобразованием жестов в текст и предлагаются два возможных подхода к решению проблемы. Представлен способ распознавания динамических казахских дактильных знаков в видеопотоке, основанный на визуальном анализе положения рук с позиционированием ключевых точек. Для считывания и позиционирования ключевых точек из изображений использована инструментальная система MediaPipe. В качестве классификаторов знаков применяются усовершенствованные варианты рекуррентных нейронных сетей для обнаружения жестовых изображений. Для распознавания жестов подготовлены собственные наборы данных, которые позволяют достичь высокой эффективности и точности распознавания по принятым моделям в режиме реального времени. Указанный набор данных состоит из динамических жестов казахского дактильного языка. Достоинством рассматриваемых подходов является относительная простота предлагаемых методов в вычислительном отношении, вследствие чего их реализация не требует использования высокопроизводительной вычислительной техники и специального оборудования. В проведенных экспериментальных исследованиях с использованием предложенных моделей LSTM и GRU достигнута точность распознавания динамических дактильных жестов казахского языка 85 % и 100 % соответственно. При этом распознавание осуществлялось в режиме реального времени с использованием обычной компьютерной видео камеры. Проведено также тестирование распознавания изображений этих же жестов, полученных с помощью веб-камеры. Получены результаты использованных моделей LSTM и GRU с оценкой распознавания динамических жестов казахского дактильного языка.

Ключевые слова: казахский жестовый язык, распознавание жестов рук, динамические жесты, Media Pipe, LSTM, GRU

Introduction

Sign language recognition has been an active area of research for nearly two decades. In sign language, the extraction of letters, words and sentences is expressed through hand gestures, body movements, facial expressions and facial emotions.

Analysis of the literature can allow us to reveal vision-based and sensor-based hand gesture recognition methods. Sensor-based methods are expensive, which

contribute to implement limited movement, and require specialized equipment (Chen et al., 2020: 25). While vision-based computer technologies use bare hands without any sensors. Thus, vision-based method using cameras become applicable later because it is convenient and easy to use (Farid et al., 2022: 19). One of the benefits of visual methods is their minimal cost (Jacob et al., 2021: 6). This study proves that there is currently no need for sophisticated and expensive equipment for sign language recognition. All you need is a modern cell phone or computer camera. In addition, the issue of hand tracking remains when using all vision-based methods.

We proposed a methodology, which uses an approach including Media Pipe open source and a machine learning algorithm on top of this framework to get a faster and simpler pipeline that can be used as a sign language recognition system. The camera captures real-time gesture images and builds a gesture dataset to recognize each Kazakh dactyl (Zholshiyeva et al., 2021: 4).

There are two types of sign systems: sign language and dactyl alphabet. The Kazakh dactyl alphabet is a system of hand gestures corresponding to the letters of the Kazakh alphabet, which includes 42 characters, reproduced by hand. For the study, dynamic Kazakh dactyl gestures were applied (Figure 1).

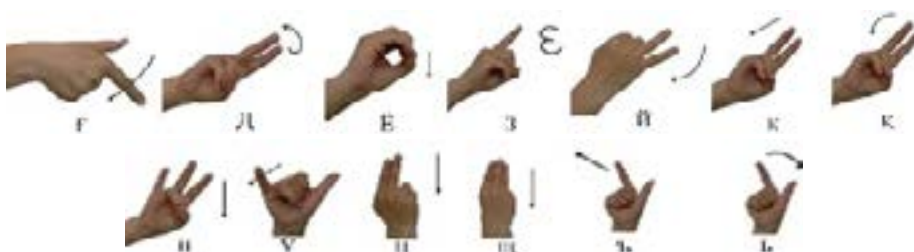


Fig. 1. Dynamic Kazakh Dactyl Gestures

Listening to manual communication is an expression of respect for the deaf people.

The purpose of that study is to improve the culture of interpersonal communication among the deaf in our country and help them recognize the Kazakh gestures. The proposed system uses the translation of the Kazakh dactyl gestures into text. The LSTM and GRU models have proven to be very effective for modeling data sequences and are used for gesture recognition.

Materials and Methods

There is a lot of research on image processing and machine learning for gesture recognition. Many good results have been obtained with high recognition accuracy of static as well as dynamic gestures. This proves that it is possible to recognize various gestures in real time, using a mobile phone or a computer camera. This study examines vision-based methods using MediaPipe (Google. MediaPipe) and machine learning methods.

There are few studies on Kazakh hand gestures recognition. For example, the review (Zholshiyeva et al., 2020: 8) proposes a CNN-based approach using advanced methods and algorithms for hand gesture detection and Kazakh sign language recognition. The same authors (Zholshiyeva et al., 2021: 4) proposed their ideas for tracking hand postures and for creating their own datasets of the Kazakh dactyl alphabet. In addition, the same authors recognized 42 Kazakh gestures with a high percentage of recognition using the machine learning method in the study (Zholshiyeva et al., 2023: 12). (Sarfaraz et al., 2018: 10) show that CNN along with RNN can be successfully used for video gesture classification and gives an accuracy of 95.217 %. There is a method for implementing recognition of gesture commands in a video stream based on Media Pipe and recognition of a dynamic sign language through GRU, LSTM and LSTM recurrent neural network (Samaan et al., 2022: 15). Experiments have shown that the named model has an accuracy of more than 99 %. The following paper (Grif et al., 2022: 7) presents a system for recognizing isolated static and dynamic gestures of the dactyl alphabet of Russian Sign Language. The system is based on a machine learning method such as LSTM with Media Pipe Hands. The network showed an F-measure value of 91 % on the test data. In the paper by (Haldera et al. 2021: 9) demonstrates a methodology that facilitates sign language recognition using the Media Pipe framework and a machine learning algorithm with an average accuracy of 99%. Research experiments (Junyoung Chung et.al) showed that GRU is indeed better than more traditional repeating units such as tanx units and GRU is comparable to LSTM.

Long Short Term Memory (LSTM) and Gated Recurrent Unit (GRU)

Problems with sequence prediction have been around for a long time. For sequence prediction, long short-term memory networks, also known as LSTMs, are considered the most efficient solution. They can relate previous information to the current task, for example, using previous video frames could help understand the current frame, which other neural networks do not have. LSTMs have an advantage over conventional feedforward neural networks. This is due to their ability to selectively remember patterns for long periods. They work extremely well on a wide variety of problems, and they are now widely used.

LSTM is designed to avoid long-term addiction problems. Thanks to a specially designed memory cell, by remembering information for long periods, LTSM can handle long-term dependencies.

GRU is very similar to LSTM and is presented as an improvement over LSTM. Just like the LSTM, the GRU uses gates to control the flow of information. In addition, it has a simpler architecture and better performance compared to LSTM, but it does it with fewer parameters and operations. Another interesting feature of GRU is that it does not have a separate cell state, there is only a hidden state, and thanks to the simpler architecture, GRUs learn faster.

Figure 2 shows that LSTM has three gates, while GRU has only two gates. In the LSTM, they are input elements, forget elements, and output elements. Whereas

GRU has reset and update gates. There are two states in the LSTM: the cell state, or long-term memory, and the latent state, also known as short-term memory. In the case of GRU, there is only one state, i.e. the hidden state (Vasilev et al. 2019: 379).



Fig. 2. LSTM and GRU architectures

The peculiarities of these architectures are that their element is not individual neurons, but a whole set of neurons, which is called a cell or a module. This cell contains a memory cell to which data can be stored for a long period and several gates that control what will be stored in this memory cell, what will be output, and when the data from this cell will be deleted.

The advantages of LSTM and GRU networks are that they allow you to store information for a long period. In addition, the vanishing gradient problem is solved in these networks.

Despite the fact that LSTM and GRU networks have a complex structure, traditional approaches are used to train them:

- Training with a teacher
- Error backpropagation.

At the same time, the networks simultaneously learn how to give the correct output signal and select the weights for the gates in order to remember the correct values in the cell and produce the correct signal.

Results and Discussion

Proposed Model Description

This study presents two models for recognizing dynamic Kazakh dactyl gestures and uses machine-learning methods, namely LSTM and GRU. The algorithm was implemented in the Python 3.9.13, with code execution in the Jupyter Notebook development environment.

Proposed Methodology

Dataset

Own data set of dynamic dactyl gestures of the Kazakh sign language is used in this study, such as “F”, “Д”, “E”, “З”, “Й”, “К”, “Қ”, “Н”, “Ү”, “Ц”, “Щ”, “Ъ”, “ь”. It consists of 390 videos recorded by the computer camera.

Feature Extraction Using MediaPipe Framework

MediaPipe Hands was used for feature extraction. MediaPipe Hands is a list of

key points for hands and pose evaluation. For each hand, it extracts 21 key points, and the key points are calculated in 3D space: X, Y, and Z ([Zholshiyeva et al., 2021: 4]), representing real 3D coordinates with origin at the geometric center of the hand. MediaPipe Hand Landmarker allows you to determine the landmarks of the hands in the image (Figure 3, 4).

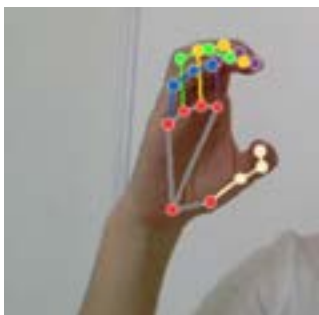


Fig. 3. Hand landmark of dactyl “Ə”

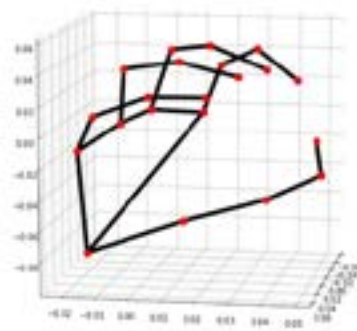


Fig. 4. Plot3d hand landmark of dactyl “Ə”

MediaPipe Pose Landmarker allows determining the landmarks of human bodies in the image. For pose estimation, MediaPipe extracts 33 key points; for a face, it extracts 468 key points (Samaan et al., 2022: 15).

The Models Results

In our paper, we attempted to show a solution from two models GRU, LSTM with parameters: the number of nodes between 32 and 128, “relu” or “softmax” activation function, “Adam” optimizer. GRU is similar to LSTM with forget gate, but it tends to have fewer parameters. The models are a neural network for multi-class classification. Each class corresponds to one sign of the dactyl gesture. The network contains six layers - three LSTM (or GRU) layers and three fully connected layers. The first, second and third layers include 64, 128 and 64 neurons, respectively, while the fully connected layers consist of 64 and 32 neurons and have the “relu” activation function. The results of each model can be seen in Figures 5, 6.

```

Model: "sequential"
-----
Layer (type)                Output Shape                Param #
-----
lstm (LSTM)                  (None, 30, 64)             48896
lstm_1 (LSTM)                (None, 30, 128)           98816
lstm_2 (LSTM)                (None, 64)                 49488
dense (Dense)                (None, 64)                 4160
dense_1 (Dense)              (None, 32)                 2080
dense_2 (Dense)              (None, 13)                 429
-----
Total params: 203,789
Trainable params: 203,789
Non-trainable params: 0
    
```

Fig. 5. Result of LSTM model

```

Model: "sequential_3"
-----
Layer (type)                Output Shape                Param #
-----
gru_6 (GRU)                  (None, 30, 64)             36864
gru_7 (GRU)                  (None, 30, 128)           74496
gru_8 (GRU)                  (None, 64)                 37248
dense_9 (Dense)              (None, 64)                 4160
dense_10 (Dense)             (None, 32)                 2080
dense_11 (Dense)             (None, 13)                 429
-----
Total params: 155,277
Trainable params: 155,277
Non-trainable params: 0
    
```

Fig. 6. Result of GRU model

In this way, it is possible to analyze the movement of the character and predict the sign of the hand gesture. Therefore, we can say that 13 dynamic Kazakh gestures such as “F”, “Д”, “E”, “З”, “Й”, “K”, “К”, “H”, “Y”, “Ц”, “Щ”, “Ъ”, “Ь” are recognized effectively (Figure 7).



Fig.7. Screenshots of the Results

Models Evaluation

Accuracies of 85 %, 100 % were obtained for the LSTM and GRU models, respectively, and models were evaluated. Tables 1, 2 present the results of recognition accuracy for 13 dactyls of the Kazakh Sign Language.

Table 1 – Metrics for evaluating the quality of LSTM model

	precision	recall	f1-score	support
GG	0.88	1.00	0.93	7
D	0.75	0.60	0.67	10
IO	0.86	0.86	0.86	7
Z	1.00	0.77	0.87	13
II	0.75	0.50	0.60	6
K	0.82	0.90	0.86	10
KK	0.82	1.00	0.89	8
NN	0.82	0.69	0.75	13
UU	0.75	1.00	0.86	9
CC	0.67	0.86	0.75	7
SHI	0.92	0.92	0.92	13
JU	1.00	1.00	1.00	9
JI	1.00	1.00	1.00	5
accuracy			0.85	117
macro avg	0.85	0.85	0.84	117
weighted avg	0.85	0.85	0.84	117

Table 2 – Metrics for evaluating the quality of GRU model

	precision	recall	f1-score	support
GG	1.00	1.00	1.00	21
D	1.00	1.00	1.00	23
IO	1.00	1.00	1.00	21
Z	1.00	1.00	1.00	23
II	1.00	1.00	1.00	21
K	1.00	1.00	1.00	20
KK	1.00	1.00	1.00	18
NN	1.00	1.00	1.00	22
UU	1.00	1.00	1.00	20
CC	1.00	1.00	1.00	19
SHI	1.00	1.00	1.00	22
JU	1.00	1.00	1.00	22
JI	1.00	1.00	1.00	21
accuracy			1.00	273
macro avg	1.00	1.00	1.00	273
weighted avg	1.00	1.00	1.00	273

Fig.8, 9 show the graphs of accuracy and loss of the above models.

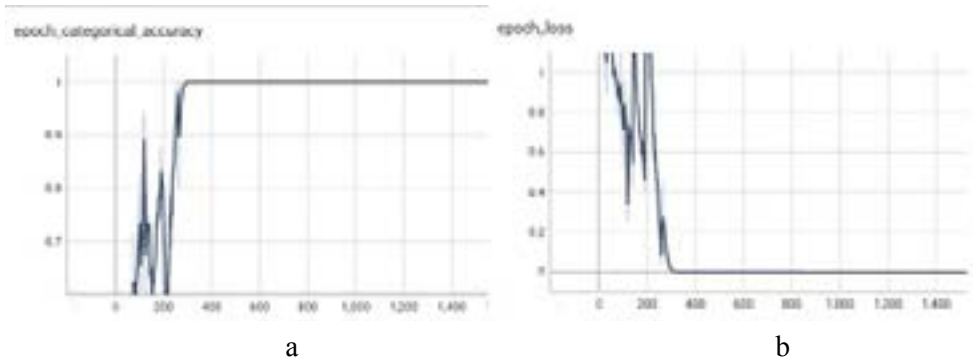


Fig. 8. Values of accuracy (a) and loss (b) of LSTM model from Tensorboard

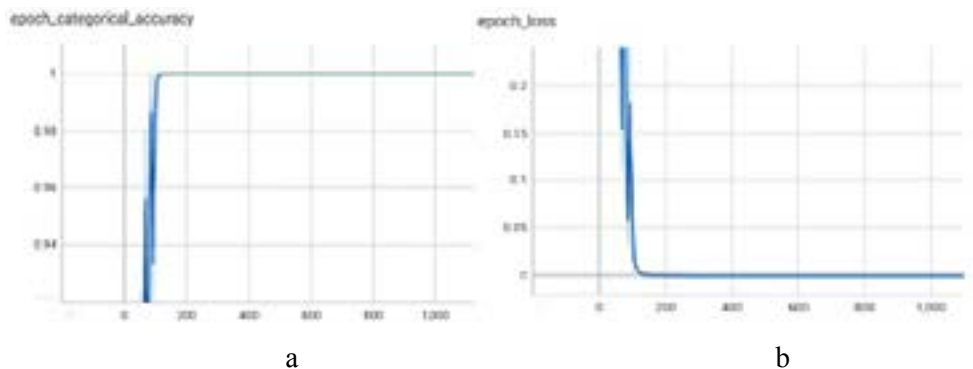


Fig. 9. Values of accuracy (a) and loss (b) of GRU model from Tensorboard

The results show that for dynamic Kazakh gestures the GRU model works perfectly, and the LSTM model showed a less good rate.

Conclusion

Thus, we presented the dynamic Kazakh dactyl recognition with two GRU, LSTM models on own dataset of Kazakh Sign Language. The MediaPipe was used to determine the location, shape and orientation, extraction of key points of the hands, body and face.

First, the dataset is prepared from videos with the same number of frames. The input was passed to the MediaPipe. Finally, the extracted key points are inserted into the GRU, LSTM models for training.

The results of the study show that MediaPipe in combination with GRU, LSTM models are effective in solving the problems of dynamic sign language recognition. Advanced RNN models like GRU, LSTM solve the problem of frame dependency in character movement. GRU trained less than LSTM and showed a better 100% recognition result, LSTM requires more parameters and nodes than GRU, and the recognition percentage is much lower than the GRU model. In the future, we plan a real-time recognition of phrases and sentences for deaf (Ergaliyeva et al., 2005: 405) in Kazakh language to enhance the culture of interpersonal communication among the deaf in our country.

REFERENCES

- A. Jacob, M. Koshy and K.K. Nisha, 2021 — *A. Jacob, M. Koshy and K.K. Nisha*. Real Time Static and Dynamic Hand Gestures Cognizance for Human Computer Interaction, 2021 International Conference on Advances in Computing and Communications (ICACC, Kochi, Kakkannad, India. - Pp. 1–6, doi: 10.1109/ICACC-202152719.2021.9708249. (in Eng.).
- Al Farid F., Hashim N., Abdullah J., Bhuiyan M.R., Shahida Mohd Isa W.N., Uddin J., Haque M.A., Husen M.N., 2022 — *Al Farid F., Hashim N., Abdullah J., Bhuiyan M.R., Shahida Mohd Isa W.N., Uddin J., Haque M.A., Husen M.N.* A Structured and Methodological Review on Vision-Based Hand Gesture Recognition System. *J. Imaging*, 8 (6), 153. <https://doi.org/10.3390/jimaging8060153> (in Eng.).
- Arpita Haldera, Akshit Tayadeb, 2021 — *Arpita Haldera, Akshit Tayadeb*. Real-time Vernacular Sign Language Recognition using MediaPipe and Machine Learning, *International Journal of Research Publication and Reviews*. Vol (2) Issue (5). 7: Pp. 9–17. (in Eng.).
- Ergaliyeva A., Kanagatova A., 2005 — *Ergaliyeva A., Kanagatova A.* The world of gestures. Dictionary for hearing-impaired citizens in Kazakhstan. Almaty, "Umit", 2005 - 401 p. ISBN 9965-9619-0-5. (in Eng.).
- Google Media Pipe, 2023 — url: [https://Gesture recognition task guide/MediaPipe/Google Developers](https://Gesture%20recognition%20task%20guide/MediaPipe/Google%20Developers). (in Eng.).
- Grif M.G., Kondratenko Y.K., 2021 — *Grif M.G., Kondratenko Y.K.* Development of a software module for recognizing the fingerspelling of the Russian Sign Language based on LSTM // *International Conference on IT in Business and Industry*. 2032 (1): Pp. 1–8. DOI: 10.1088/1742-6596/2032/1/012024. (in Eng.).
- Ivan Vasilev, Daniel Slater, Gianmario Spacagna, Peter Roelants, Valentino Zocca, 2019 — *Ivan Vasilev, Daniel Slater, Gianmario Spacagna, Peter Roelants, Valentino Zocca*. Python Deep Learning Second Edition, Packt Publishing Ltd., UK. ISBN 978-1-78934-846-0, p. 379. (in Eng.).
- Junyoung Chung, Caglar Gulcehre, KyungHyun Cho, Yoshua Bengio, 2014 — *Junyoung Chung, Caglar Gulcehre, KyungHyun Cho, Yoshua Bengio*. Empirical Evaluation of Gated Recurrent Neural Networks on Sequence Modeling, arXiv:1412.3555v1 [cs.NE]. (in Eng.).

Lugaresi C., Tang J., Nash H., McClanahan C., Uboweja E., Hays M., Zhang F., Chang C.L., Yong M.G., Lee J. et al., 2019 — *Lugaresi C., Tang J., Nash H., McClanahan C., Uboweja E., Hays M., Zhang F., Chang C.L., Yong M.G., Lee J. et al.* Mediapipe: A framework for building perception pipelines. arXiv, arXiv:1906.08172. (in Eng.)

Samaan G.H., Wadie A.R., Attia A.K., Asaad A.M., Kamel A.E., Slim S.O., Abdallah M.S., Cho Y.-I., 2022 — *Samaan G.H., Wadie A.R., Attia A.K., Asaad A.M., Kamel A.E., Slim S.O., Abdallah M.S., Cho Y.-I.* MediaPipe's Landmarks with RNN for Dynamic Sign Language Recognition. Electronics, 11 (19), 3228. <https://doi.org/10.3390/electronics11193228>. (in Eng.)

Sarfaraz Masood, Adhyan Srivastava, Harish Chandra Thuwal, and Musheer Ahmad, 2018 — *Sarfaraz Masood, Adhyan Srivastava, Harish Chandra Thuwal, and Musheer Ahmad.* Real-Time Sign Language Gesture (Word Recognition from Video Sequences Using CNN and RNN". In: Intelligent Engineering Informatics. Ed. by Vikrant Bhateja, Carlos A. Coello Coello, Suresh Chandra Satapathy, and Prasant Kumar Pattnaik. Singapore: Springer Singapore. Pp. 623–632. isbn: 978-981-10-7566-7. (in Eng.)

Weiya Chen, Chenchen Yu, Chenyu Tu, Zehua Lyu, Jing Tang, Shiqi Ou, Yan Fu, and Zhidong, 2020 — *Weiya Chen, Chenchen Yu, Chenyu Tu, Zehua Lyu, Jing Tang, Shiqi Ou, Yan Fu, and Zhidong.* A Survey on Hand Pose Estimation with Wearable Sensors and Computer-Vision-Based Methods, Sensors, 20 (4), 1074, doi:10.3390/s20041074. (in Eng.)

Zholshiyeva L., Zhukabayeva T., Turaev Sh., Berdiyeva M., Jambulova D., 2021 — *Zholshiyeva L., Zhukabayeva T., Turaev Sh., Berdiyeva M., Jambulova D.* Hand Gesture Recognition Methods and Applications: A Literature Survey," ICEMIS'21: The 7th International Conference on Engineering & MIS 2021, 7: Pp.1-8. <https://doi.org/10.1145/3492547.3492578>. (in Eng.)

Zholshiyeva L., Zhukabayeva T., Turaev Sh., Berdiyeva M., Sengirbaeva R., 2022 — *Zholshiyeva L., Zhukabayeva T., Turaev Sh., Berdiyeva M., Sengirbaeva R.* A Real-Time Approach to Recognition of Kazakh Sign Language. 2022 IEEE International Conference on Smart Information Systems and Technologies (SIST), Nur-Sultan, Kazakhstan, 28–30 April. (in Eng.)

Zholshiyeva L., Zhukabayeva T., Turaev Sh., Berdiyeva M., Sengirbaeva R., 2023 — *Zholshiyeva L., Zhukabayeva T., Turaev Sh., Berdiyeva M., Sengirbaeva R.* Real-time Kazakh Sign Language Recognition Using Media Pipe and SVM. News of the National Academy of Sciences of the Republic of Kazakhstan Physico-mathematical Series. ISSN 1991–346X. 345 (1): PP.82–93. <https://doi.org/10.32014/2022.2518-1726.170>. (in Eng.)

NEWS OF THE NATIONAL ACADEMY OF SCIENCES OF THE REPUBLIC OF
KAZAKHSTAN
PHYSICO-MATHEMATICAL SERIES
ISSN 1991-346X
Volume 2. Number 346 (2023). 154–171
<https://doi.org/10.32014/2023.2518-1726.191>

UDC 665.733.3: 519.816

© **M. Kabibullin¹, B. Orazbayev¹, K. Orazbayeva^{2*}, S. Iskakova³,
Zh. Amanbayeva³, 2023**

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

²Esil University, Astana, Kazakhstan;

³Atyrau Oil and Gas University, Atyrau, Kazakhstan.

E-mail: madyar_kabibullin@mail.ru

DEVELOPMENT OF MODELS OF UNITS OF COMPLEX CHEMICAL-TECHNOLOGICAL SYSTEMS UNDER CONDITIONS OF DEFICIENCY AND FUZZY OF INITIAL INFORMATION

Kabibullin M.D. — doctoral student of the Department of System Analysis and Control, L.N. Gumilyov Eurasian National University, Satpayev str.2A, Astana, Kazakhstan

E-mail: madyar_kabibullin@mail.ru. ORCID: 0000-0002-3972-7894;

Orazbayev B.B. — doctor of technical sciences, academician of the Engineering academy of the republic of Kazakhstan, professor of the department of System analysis and Control, L.N. Gumilyov Eurasian National University, Satpayev str.2A, Astana, Kazakhstan

E-mail: batyr_o@mail.ru. ORCID: 0000-0003-2109-6999;

Orazbayeva K.N. — doctor of technical sciences, professor of the department Management, Esil University, Zhubanov str.7, Astana, Kazakhstan

E-mail: kulman_o@mail.ru. ORCID: 0000-0002-1741-7553;

Iskakova S.Sh. — Candidate of Technical Sciences, Associate Professor of Information Technologies, Atyrau Oil and Gas University, Baymukhanov str.45A, Atyrau, Kazakhstan

E-mail: iskakova_sh@mail.ru. ORCID: 0000-0002-6589-854X;

Ananbayeva Zh.Sh. — Senior Lecturer of Information Technologies, Atyrau Oil and Gas University, Baymukhanov str.45A, Atyrau, Kazakhstan

E-mail: zhak-7777@mail.ru. ORCID: 0009-0005-2200-4453.

Abstract. Based on the methodology of system analysis, an integrated approach is developed in the work, which allows creating a system of models of interconnected aggregates of complex chemical-technological systems (CTS), characterized by a deficit and fuzziness of the initial information. The developed method is based on the evaluation of models that can be created for individual CTS units, using expert evaluation methods based on selected criteria, determining the most effective type of model and building it based on the available information. Based on the available information of a different nature, a block diagram of the proposed system method for creating models of CTS units is constructed and described. In the proposed

method, an effective model is created depending on the type of information of a different nature about the state and modes of operation of each unit and the value of the integral criteria for evaluating models. Then the created models are integrated into a single system of models, taking into account the interconnections of the CTS units and the flows of technological processes occurring in them. The proposed system method allows you to create deterministic models based on analytical methods, statistical models using experimental statistical methods and fuzzy or linguistic models using the methods proposed in this paper. Fuzzy models will be created in the form of fuzzy regression equations (when the input parameters of the object are clear, and linguistic models (when both the input and output parameters of the object are fuzzy will be developed based on the rules of logical conditional inference. The novelty of the proposed system method, which is being developed to create models of interconnected aggregates of CTS aggregates in conditions of scarcity and fuzziness of the initial information, lies in its synergistic effect and the emergence property, which make it possible to carry out system modeling of CTS.

Keywords: chemical-technological system, model system, fuzzy model, linguistic model, system method, membership function

© М.Д. Кабибуллин¹, Б.Б. Оразбаев¹, К.Н. Оразбаева^{2*}, С.Ш. Искакова³,
Ж.Ш. Аманбаева³, 2023

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

²Esil University, Астана, Қазақстан;

³С. Өтебаев атындағы Атырау мұнай және газ университеті,
Атырау, Қазақстан.

E-mail: madyar_kabibullin@mail.ru

КҮРДЕЛІ ХИМИЯЛЫҚ-ТЕХНОЛОГИЯЛЫҚ ЖҮЙЕЛЕР АГРЕГАТТАРЫНЫҢ МОДЕЛЬДЕРІН БАСТАПҚЫ АҚПАРАТТЫҢ ЖЕТІСПЕУШІЛІГІ МЕН АЙҚЫНСЫЗДЫҒЫ ЖАҒДАЙЫНДА ҚҰРУ

Кабибуллин М.Д. — Жүйелік талдау және басқару кафедрасының докторанты, Л.Н.Гумилев атындағы Еуразия ұлттық университеті, Сәтпаев көш. 2А, Астана қ., Қазақстан

E-mail: madyar_kabibullin@mail.ru. ORCID: 0000-0002-3972-7894;

Оразбаев Б.Б. — техника ғылымдарының докторы, Қазақстан Республикасы Инженерлік академиясының академигі, Жүйелік талдау және басқару кафедрасының профессоры, Л.Н.Гумилев атындағы Еуразия ұлттық университеті, Сәтпаев көш. 2А, Астана қ., Қазақстан
E-mail: batyr_o@mail.ru. ORCID: 0000-0003-2109-6999;

Оразбаева К.Н. — техника ғылымдарының докторы, Менеджмент кафедрасының профессоры, Esil University, Жұбанов көш. 7, Астана қ., Қазақстан
E-mail: kulman_o@mail.ru. ORCID: 0000-0002-1741-7553;

Искакова С.Ш. — техника ғылымдарының кандидаты, Ақпараттық технологиялар факультетінің қауымдасқан профессоры, С. Өтебаев атындағы Атырау мұнай және газ университеті, Баймуханов көш. 45а, Атырау қ., Қазақстан
E-mail: iskakova_sh@mail.ru. ORCID: 0000-0002-6589-854X;

Аманбаева Ж.Ш. — Ақпараттық технологиялар факультетінің аға оқытушысы, С. Өтебаев атындағы Атырау мұнай және газ университеті, Баймуханов көш. 45а, Атырау қ., Қазақстан
E-mail: zhak-7777@mail.ru. ORCID: 0009-0005-2200-4453.

Аннотация. Жұмыста жүйелік талдау методологиясы негізінде бастапқы ақпараттарының тапшылығы және айқын еместігімен сипатталатын күрделі химиялық-технологиялық жүйелердің (ХТЖ) өз ара байланысқан агрегаттарының модельдер жүйесін құруға мүмкіндік беретін кешенді тәсіл жасақталады. Жасақталатын тәсіл ХТЖ-дің жеке агрегаттарына құрылуы мүмкін модельдерін таңдалған критерийлер негізінде эксперттік бағалау тәсілі арқылы бағалап, ең тиімді модель түрін анықтап, оны қол жетімді ақпарат негізінде құруға негізделеді. Қолжетімді әр түрлі сипаттағы ақпарат негізінде ХТЖ агрегаттарының модельдерін құру үшін ұсынылатын жүйелік тәсілдің блок-сызбасы тұрғызылып, сипатталады. Ұсынылатын тәсілде әр агрегаттың күйі мен жұмыс режимдері жайлы түрлі сипаттағы ақпараттың түріне және модельдерді бағалаудың интегрленген критерийлерінің мәніне байланысты тиімді модель құрылады. Содан кейін құрылған модельдер ХТЖ-дің агрегаттарының өз ара байланыстың мен оларда жүретін технологиялық процесстердің ағындарын ескере отырып бір модельдер жүйесіне біріктірілді. Ұсынылатын жүйелік тәсіл аналитикалық тәсілдер негізінде детерминдік модель, эксперименталдық-статистикалық тәсілдері көмегімен статистикалық модельдер және осы жұмыста ұсынылатын тәсілдер көмегімен айқын емес немесе лингвистикалық модельдер құруға мүмкіндік береді. Айқын емес модель (нысанның кіріс параметрлері айқын, ал шығыс параметрлері айқын емес болғанда) айқын емес регрессиялық теңдеулер түрінде, ал лингвистикалық модельдер (нысанның кіріс және шығыс параметрлері айқын емес болғанда) логикалық шартты қорытындалау ережелері негізінде құрылатын болады. Ұсынылатын ХТЖ-дің өз ара байланысқан агрегаттарының модельдерін бастапқы ақпарат тапшылығы мен айқынсыздығы жағдайында құруға әзірленетін жүйелік тәсілдің жаңашылдығы оның ХТЖ-ді жүйелік модельдеуге мүмкіндік беретін синергизм эффектісі мен эмердженттік қасиетінің болуында.

Түйін сөздер: химиялық-технологиялық жүйе, модельдер жүйесі, айқын емес модель, лингвистикалық модель, жүйелік тәсіл, тиістілік функция

© М.Д. Кабибуллин¹, Б.Б. Оразбаев¹, К.Н. Оразбаева^{2*}, С.Ш. Искакова³,
Ж.Ш. Аманбаева³, 2023

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

² Esil University, г. Астана, Казахстан;

³ Атырауский университет нефти и газа им. С. Утебаева, Атырау, Казахстан.
E-mail: madyar_kabibullin@mail.ru

РАЗРАБОТКА МОДЕЛЕЙ АГРЕГАТОВ СЛОЖНЫХ ХИМИКО-ТЕХНОЛОГИЧЕСКИХ СИСТЕМ В УСЛОВИЯХ ДЕФИЦИТА И НЕЧЕТКОСТИ ИСХОДНОЙ ИНФОРМАЦИИ

Кабибуллин М.Д. – докторант кафедры Системного анализа и управления Евразийский национальный университет им. Л.Н. Гумилева, ул. Сатпаева 2А, Астана, Казахстан
E-mail: madyar_kabibullin@mail.ru. ORCID: 0000-0002-3972-7894;

Оразбаев Б.Б. – академик Инженерной академии Республики Казахстан, профессор кафедры Системного анализа и управления, Евразийский национальный университет им. Л.Н. Гумилева, ул. Сатпаева 2А, Астана, Казахстан
E-mail: batyr_o@mail.ru. ORCID: 0000-0003-2109-6999;

Оразбаева К.Н. – доктор технических наук, профессор кафедры Менеджмента, Esil University, ул. Жубанова 7, Астана., Казахстан
E-mail: kulman_o@mail.ru. ORCID: 0000-0002-1741-7553;

Искакова С.Ш. – кандидат технических наук, ассоциированный профессор факультета Информационных технологий, Атырауский университет нефти и газа имени С. Утебаева, ул. Баймухановаш. 45а, Атырау, Казахстан
E-mail: iskakova_sh@mail.ru. ORCID: 0000-0002-6589-854X;

Аманбаева Ж.Ш. – старший преподаватель факультета Информационных технологий, Атырауский университет нефти и газа имени С. Утебаева, ул. Баймуханова, 45а, Атырау, Казахстан
E-mail: zhak-7777@mail.ru. ORCID: 0009-0005-2200-4453.

Аннотация. В работе на основе методологии системного анализа разрабатывается комплексный подход, позволяющий создать систему моделей взаимосвязанных агрегатов сложных химико-технологических систем (ХТС), характеризующиеся дефицитом и нечеткостью исходной информации. Разрабатываемый метод основан на оценке моделей, которые могут быть созданы для отдельных агрегатов ХТС с использованием методов экспертной оценки на основе выбранных критериев, определении наиболее эффективного типа модели и построении ее на основе имеющейся информации. На основе доступной информации различного характера строится и описывается блок-схема предлагаемого системного метода создания моделей агрегатов ХТС. В предлагаемом методе эффективная модель создается в зависимости от вида информации разного характера о состоянии и режимах работы каждого агрегата и значения интегральных критериев оценки моделей. Затем созданные модели интегрируются в единую систему моделей с учетом взаимосвязей агрегатов ХТС и потоков протекающих в

них технологических процессов. Предлагаемый системный метод позволяет создать детерминированные модели на основе аналитических методов, статистических моделей с использованием экспериментально-статистических методов и нечетких или лингвистических моделей с использованием методов, предложенных в данной работе. Нечеткие модели будут создаваться в виде уравнений нечетких регрессий (когда входные параметры объекта четкие, а лингвистические модели (когда и входные, и выходные параметры объекта нечеткие) будут разработаны на основе правил логического условного вывода. Новизна предлагаемого системного метода, разрабатываемого для создания моделей взаимосвязанных агрегатов агрегатов ХТС в условиях дефицита и нечеткости исходной информации, заключается в его эффекте синергизма и свойстве эмерджентности, позволяющие проводить системное моделирование ХТС.

Ключевые слова: химико-технологическая система, система моделей, нечеткая модель, лингвистическая модель, системный метод, функция принадлежности

Кіріспе

Мұнай өңдеу, химия, металлургия тағы басқа өндіріс салаларының негізгі технологиялық процесстері күрделі химиялық-технологиялық жүйеде (ХТЖ) жүреді. Мұндай ХТЖ шикізатты дайындауға, химиялық түрлендіруге, аралық және мақсаттық өнімдерді алуға арналған технолологиялық процесстердің тізбегі жүретін өз-ара байланысқан, біртұтас жүйе ретінде жұмыс жасайтын агрегаттардың жиынтығы болып табылады. Әдетте өндірістік жағдайда ХТЖ-дің математикалық модельдерін құрып, оптимизациялауға және басқаруға қажетті ақпараттардың тапшылығы және айқынсыздығымен сипатталады (Липин, 2018; Плановский, 2020; Ушева, 2019). Математикалық модельдеу ХТЖ-ді зерттеу, олардың жұмыс режимдерін оптимизациялау және басқару үшін қолданатын негізгі тәсілдердің бірі болып табылады. Қазіргі компьютерлік техниканың дамыған кезінде математикалық модельдер негізінде компьютерлік модельдеу ХТЖ-дің оптималды жұмыс режимдерін табудың, оларды басқарудың ең тиімді, көрнекті және көрнекі құралы ретінде саналады (Ovechkin, 2018; Orazbayev, 2018). ХТЖ-дің оптималды жұмыс режимдерін анықтау және тиімді басқару үшін, олардың өз ара байланысқан агрегаттарының модельдері пакетін қолдану қажет.

Ғылыми әдебиеттерде осы кезге дейін ХТЖ-дің өз ара байланысқан агрегаттарының модельдерінің пакетін анықсыздық жағдайда құру сұрақтары әлі де толықтай зерттеліп шешілмеген мәселелердің бірі болып табылады. Сондықтан бұл жұмыстың негізгі мақсаты бастапқы ақпараттың тапшылығымен және айқынсыздығымен сипатталатын күрделі ХТЖ-дің өз ара байланысқан агрегаттарының модельдерін құрып, бір пакетке біріктіру сұрақтарын зерттеп, оларды шешудің жүйелі тәсілін ұсыну болап табылады.

ХТЖ-дің түрлі агрегаттарының құру сұрақтары зерттелген жұмыстар талдау жасап, оларда шешілмей қалған сұрақтарды қарастырайық. Мұнай өңдеу және басқа өндіріс салаларының түрлі ХТЖ-нін математикалық модельдеу сұрақтары Кафаров тағы басқалары (Kafarov, 2018), Жоров (Zhorov, 2015) және Mohaddec (Mohaddecu, 2006) еңбектерінде қарастырылған. Бұл және тағы басқа зерттеулерде каталикалық риформинг және басқа ХТЖ агрегаттарында өтетін типтік процесстердің теориялық зерттеулер мәліметтері мен процесстердің кинетикасын негізінде детерминді модельдері ұсынылған. Бұл жұмыстарда зерттелген ХТЖ-дің жеке агрегаттары мен элементарлық процесстердің детерминді модельдері масса және энергия сақталу заңдарына негізделгендіктер әмбебап болып табылады және олардың артықшылықтарына жатады. Алайда мәселе мынада, аталған модельдер тек типтік процесстерді сипаттайды және олар агрегаттардың байланыстары мен өз ара әсерлерін, сондай-ақ бастапқы ақпараттың тапшылығы мен айқынсыздығын ескермейді. Сондықтан мұндай модельдерді оларды құруға қажетті ақпарат жетіспеушілігі мен айқынсыздығы жағдайында құру мүмкін емес және мұндай модельдер ХТЖ жұмысын жүйеле модельдеуге жарамсыз.

Pinheiro және басқалары каталикалық риформинг процесі жүретін ХТЖ-нің риформингтеу реакторының жұмысын модельдеу мен басқару тәсілдерін өздерінің зерттеулерінде ұсынған (Pinheiro, 2018). Бірақ бұл жұмыста риформинг реакторының моделі каталикалық риформинг процесі жүретін ХТЖ-нің пеші, колонналары сияқты басқа агрегаттарымен байланысын ескермейді. Сондықтан аталған жұмыста ұсынылған риформинг реакторының моделі ол реактор жұмысын модельдеу мен оптимизациялауға жарамды болғанымен, тұтастай каталикалық риформинг қондырғысының жұмысын жүйелі модельдеу мен оптимизациялауға жарамсыз. Стохастикалық тәсілдеме негізінде ХТЖ технолоргиялық агрегаттарын басқару тәсілдері Coleman және Babu (Coleman, 2018) жұмыстарында зерттелген. Ал Bequette еңбегінде (Mircea, 2019) статистикалық модельдер негізінде ХТЖ-дің түрлі агрегаттарында өтетін технологиялық процесстерді басқару мәселелері зерттелген. Мұндай статискалық модельдер күрделі нысандар жұмысын оптимизациялау мен басқаруға жарамды. Мұндай модельдерді құру үшін нысандардың жұмысы жайында сенімді статистикалық деректердің массивы, мұндай деректерді жинау үшін көптеген эксперименттер жүргізу және оларды дұрыс өңдеу қажет. Алайда айқынсыздықпен сипатталатын Атырау мұнай өңдеу зауыты каталикалық риформинг қондырғысы сияқты көптеген күрделі ХТЖ үшін мұндай эксперименттер, модель құруға қажетті кейбір маңызды параметрлер өлшенбейтіндіктен, мүмкін емес.

ХТЖ жеке агрегаттарының жұмыс режимдерін оптимизациялауға мүмкіндік беретін және ол агрегаттардың модельдеріне негізделген түрлі басқару жүйелері белгілі. Мысалы (Keller, 2020) жұмыстың авторлары модельдер негізінде крекинг реакторлары шығысындағы температураның оптималды

таралуын анықтауға мүмкіндік беретін каталитикалық процесстерді басқару жүйесін жасақтаған, ал Gumen (Gumen, 2019) каталитикалық риформинг қондырғысының тұрақтандыру блогында бензин сапасын басқару жүйесін ұсынған. Модельдер негізінде ХТЖ-де жүретін технологиялық процесстерді жағдаяттық басқаруды қодану сұрақтарын Поспелов (Pospelov, 2018) жұмысында қарастырылған. Қарастырылған басқару жүйелері ХТЖ-дің агрегаттарының жұмыс режимдері мен оларда жүретін процесстерді детерминдіә немесе стохастикалық жағдайларда оптимизациялауға мүмкіндік береді, бірақ ол модельдер мен басқару жүйелері бастапқы ақпараттардың айқынсыздығы жағдайларында тиімсіз немесе жарамсыз.

Dzhambekov тағы басқалары (Dzhambekov, 2018), Keller мен Gorak (Keller, 2019) және Алиев тағы басқалары (Алиев, 2018) жұмыстарында бастапқы ақпараттың кей бөлігінің айқынсыздығы жағдайында ХТЖ агрегаттарының модельдерін құру және олардың негізінде агрегаттардың жұмыс режимдерін басқару тәсілдемелері зерттеліп, ұсынылған. Бұл ұсынылған тәсілдемелері эксперттік бағалау және айқын емес жиындар тәсілдерін қолдануға негізделген. Бұл тәсілдер нысанның шығыс параметрлері айқын емес, ал кіріс параметрлері айқын болған жағдайларда α деңгейлі жиындар негізінде нысандардың айқын емес модельдерін құруға мүмкіндік береді. Бірақ бұл және омы сияқты тәсілдерде нысанның барлық кіріс және шығыс параметрлері айқынсыз болғанда ол нысандардың лингвистикалық модельдерін құру сұрақтары ашшылмаған. Ал шынайы практикада көптеген ХТЖ-дің лингвистикалық модельдерін құруға тура келеді. Соныме, қатар аталған (Dzhambekov, 2018; Keller, 2019; Алиев, 2018) жұмыстарында ХТЖ-дің өз ара байланысқан агрегаттарының модельдері пакеттерін құру сұрақтары қарастырылмаған.

Сонымен зертеу жұмытарын талдау нәтижелері ХТЖ-дің жеке агрегаттарының математикалық модельдерін құрудың белгілі тәсілдері дәстүрлі аналитикалық және эксперименталдық-статистикалық тәсілдеріне негізделген, ХТЖ агрегаттарының өз ара байланысқан модельдер пакетін құру және оларды жүйелі модельдеу мәселелерін шешпейді. Бұл мәселелерді шешу үшін жүйелік талдау, эксперттік бағалау және айқын емес жиындар тәсілдеріне негізделген, күрделі айқынсыздықпен сипатталатын ХТЖ-дің жұмыс режимдерін модельдеу мен оптимизациялаудың жүйелі тәсілін құрып, қолдану қажет. Бұл жұмыс ХТЖ өз ара байланысқан агрегаттарының модельдері пакетін құруға мүмкіндік беретін жүйелі тәсілді құруға бағытталған. Жеке технологиялық нысандардың модельдерін құрудың белгілі тәсілдерінен ұсынылатын жүйелік тәсілдің айырмашылығы, ол қол жетімді түрлі сипаттағы ақпараттар негізінде ХТЖ-дің өз ара байланысқан агрегаттарының модельдер пакетін құруға мүмкіндік береді. Сонымен қатар ұсыныллатын жүйелік тәсіл ХТЖ агрегаттарының белгілі математикалық модельдермен қатар айқын емес, лингвистикалық, гибридтік модельдерін құрып, бір пакетке біріктіруге мүмкіндік береді.

Зерттеу мақсаты мен міндеттері және тәсілдері

Зерттеу мақсаты ХТЖ жұмысын жүйелі модельдеу үшін, олардың өз ара байланысқан агрегаттарының модельдері пакетін бастапқы ақпараттың жетіспеушілі және айқынсыздығы жағдайында құруға мүмкіндік беретін жүйелі тәсіл жасақтау. Қойылған мақсатқа жету үшін келесі зерттеу міндеттері қойылып шешіледі:

ХТЖ жеке агрегаттарының құрылуы мүмкін модельдері түрін анықтау және оларды тиімді құру тәсілдерін таңдау не жасақтау;

ХТЖ агрегаттары жағдайы, жұмыс жайындағы бастапқы ақпараттың айқынсыздығы деңгейіне байланысты, олардың модельдерін құру әдістемесін анықтау;

Агрегаттардың нақты бір моделін құруға қажетті ақпараттың тапшылығы жағдайында, қол жетімді түрлі ақпарат негізінде олардың гибридті, құрымдалған моделін құру әдістемесін ұсыну;

Түрлі қол жетімді ақпарат негізінде құрылатын жеке агрегаттардың модельдерін бір модельдер пакетіне біріктіру принципін әзірлеу.

Зерттеу барысында ХТЖ-дің жеке агрегаттарының модельдерін құруда олардың жұмысы мен жағдайын сипаттайтын қол жетімді ақпараттың түріне байланысты модель құрудың түрлі белгілі тәсілдері, және ақпараттың тапшылығы мен айқынсыздығы жағдайында осы жұмыста жүйелендіріліп, ұсынылған тәсілдер қолданылады.

Жұмыста ХТЖ агрегаттары жұмыс жайында теориялық мәліметтер не статистикалық деректер жеткілікті болса, сәйкесінше детерминді модель құруға арналған аналитикалық тәсілдер (Кафаров, 2018; Жоров, 2015; Sharikov, 2018) мен эксперименталдық-статистикалық тәсілдер (Zhi-Wen, 2018; Карманов, 2019) пайдаланылады. Ал айқын емес эксперттік ақпаратты алу және өңдеу үшін эксперттік бағалау тәсілдері (Гуцыкова, 2017; Jorgensen, 2019; Sabzi, 2017) мен айқын емес жиындар теориясының тәсілдері (Рыжов, 2017; Zimmermann, 2018) қолданылады. Сонымен қатар бұл жұмыста түрлі сипаттағы ақпараттар негізінде модельдер пакетін құру және жүйелік модельдеуді қамтамасыз ету үшін жүйелік талдау тәсілдері де қолданылған (Pavlov, 2016; Reverberi, 2019).

Зерттеу нәтижелері

Зерттеу барысында каталитикалық риформинг блогы сияқты күрделі және кейбір параметрлері айқынсыздықпен сипатталатын ХТЖ-дің өз ара байланысқан агрегаттарының модельдері пакетін құру тәсілін қарастырып, сипаттамасын иелендірілді.

Заманауи компьютерлер көмегімен ХТЖ жұмыс режимін модельдеу арқылы олардың түрлі жұмыс режимдерінің нәтижелерін жылдам және қажетті дәлдіпен анықтауға және оларды басқаруға болады. Мұндай компьютерлік жүйелер көмегімен ХТЖ-ден алынатын өнімдердің көлемі мен сапаларын есептеуге, сондай-ақ оларды жақсартудың резервтерін анықтауға

болады. Ол үшін әрине ХТЖ-ді жүйелі модельдеу қажет, ал оны жүзеге асыру үшін зерттелетін ХТЖ-дің жеке агрегаттарының бір пакетке біріктірілген модельдер жүйесін тұрғызу керек.

Бірінші кезекте зерттеу нәтижесін алуға қажетті шарттар мен негіздемелерді келтірейік. Әдетте математикалық модельдер, құрылатын модельді қанлай мақсатқа қолданалытына, яғни модельдеу мақсатына байланысты жсақталады. Сондықтан ХТЖ-дің модельдер пакетін құру процесінде ол модельдердің жүйелі модельдеу арқылы ХТЖ-де алынатын мақсаттық өнімдердің шығысын барынша арттыру және олардың сапа көрсеткіштерін жақсартуға жарамдығын ескерген жөн. Сонымен қатар басқару критерийлері ретінде өндірістің экологиялық көрсеткіштері және басқалары да анықталуы мүмкін. Сол себептен құрылатын модель, қойылған шектеулерді ескере отырып, өндірісті басқару критерийлер векторы бойынша тиімді шешім қабылдауды қамтамасыз етуі шарт..

Модельдеу мақсаты, басқару критерийлері және шектеулер анықталғаннан анықталған кейін ХТЖ-дің әр агрегаттарының модельдерін құру мәселесі туындайды. Бұл жерде нысанның нақты моделі сол модельді құруға қажетті ақпараттыр жеткілікті болғанда құрылуы мүмкін екен атап өткен жөн. Ол ақпараттар негізінде зерттеу нысанының модельдерін құрылымдық және параметрлік идентификациялау есептері шешіледі.

Бастапқы ақпараттардың тапшылығы мен айқынсыздығы жағдайында күрделі ХТЖ-дің өз ара байланысқан агрегаттарының модельдері пакетін құруға ұсынылған жүйелік тәсілінің негізгі кезеңдерінің блок-сызбасы сурет 1-де келтірілген.

Ұсынылған жүйелі тәсілді негізгі блоктары бойынша жете сипаттамасын келтірейік.

Тәсілдің 2-ші блогында зерттелетін ХТЖ-нің күйі, оның жұмысы жайлы ақпарат, оның жұмыс режиміне, шығысына әсер ететін кіріс, режимдік параметрлері векторы $\mathbf{x} = (x_1, \dots, x_n)$ анықталып, ендіріледі. Сонымен қатар бұл блокта ХТЖ-нің және оның агрегаттарының бақыланатын, басқарылатын шығыс параметрлері, оның жұмыс сапасын сипаттайтын критерийлер векторы $\mathbf{y} = (y_1, \dots, y_m)$ анықталады. Шынайы өндірістік жағдайда аталған векторлардың бір бөлігі не толықтай айқын емес, яғни $\tilde{\mathbf{x}} = (\tilde{x}_1, \dots, \tilde{x}_n)$ және $\tilde{\mathbf{y}} = (\tilde{y}_1, \dots, \tilde{y}_m)$ болуы мүмкін.

3 және 4 блоктарда ХТЖ мен оның агрегаттары (элементтері) сипаттамалы мен байланыстарын жүйелік талдау, модельдеу мақсатын, сондай-ақ әр агрегатқа құрылуы мүмкін және тиімді болатын модельдерді бағалдау критерийлерін анықтау жүзеге асырылады.

5-ші блокта алдыңғы блокта анықталған критерийлер негізінде ХТЖ агрегаттарының құрылуы мүмкін модельдер типін эксперттік бағалау арқылы әр агрегатқа құрылуы тиімді модель типін таңдау орындалады.

6-шы блоктан ХТЖ агрегаттарының қол жетімді ақпарат сипатына

байланысты таңдалған тиімді модельдерін құру процедурасы басталады. Қай агрегатқа қандай модель типін құру тиімді екенін анықтау үшін эсперттік бағалау нәтижелері баллдық жүйеде анықталатындықтан, әр агрегатқа әр модельдің бағалары қосылып, интегрленген критерийдің мәні анықталады. Содан кейін келесі шарттар тексеріледі.

Егер теориялық мәліметтер модель құруға жеткілікті және интегрленген критерий мәні бойынша детерминді модель максималды мәнге ие болса (блок 7), онда аналитикалық тәсілдер көмегімен ХТЖ-нің қарастырылған агрегатының детерминді моделі құрылады (блок 8).

Детерминді модель құру мүмкін емес немесе тиімсіз болса 9-шы блокта келесі шарт тексеріледі.

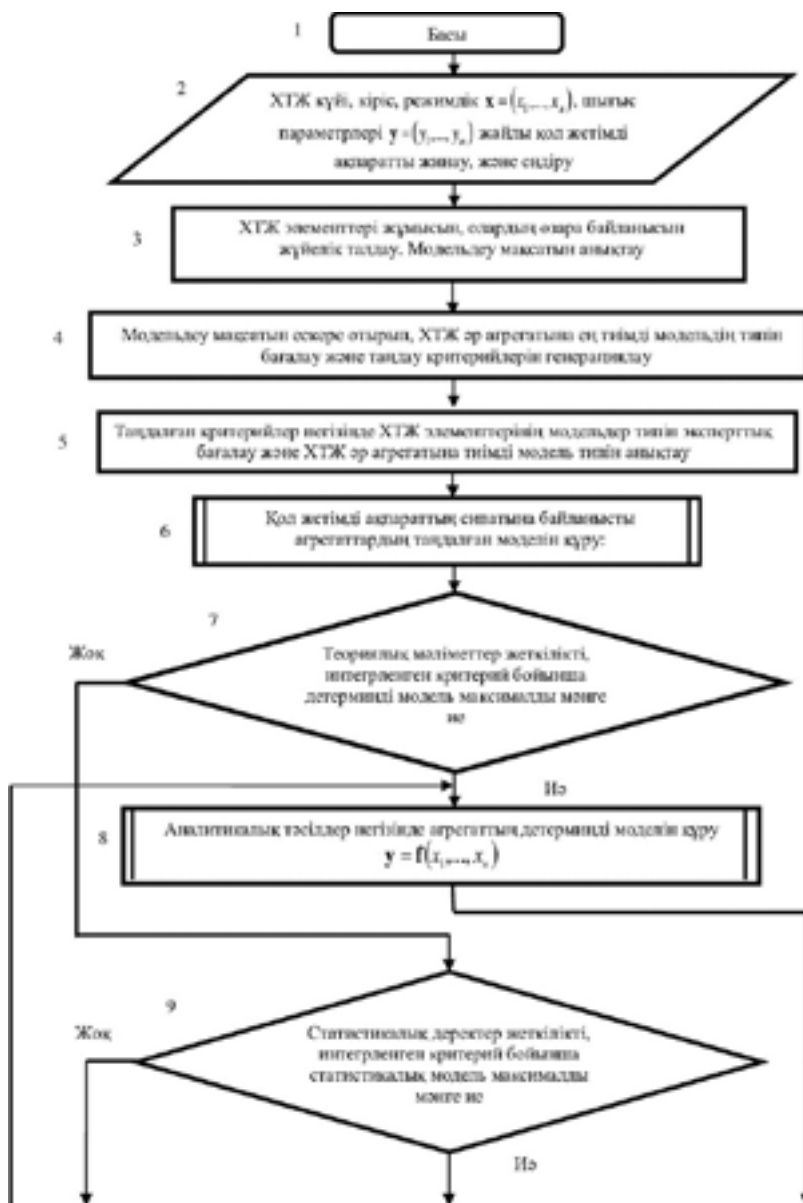
Егер ХТЖ-нің агрегаты жұмысы жайлы дұрыс статистикалық мәліметтер қажетті көлемі жеткілікті не қол жетімді және интегрленген критерий мәні бойынша статистикалық модель ең үлкен мәнге ие болса, онда онда эксперименталдық-статистикалық негізінде агрегаттың статистикалық моделі құрылады (блок 10). Бұл блокта модельдің құрылымы регрессорларды тізбектей қосу тәсілі негізінде идентификациялауға болады. Ал алынған регрессорлық модельдердің параметрлерін идентификациялау үшін белгілі ең кіші квадраттар тәсілі арқылы жүзеге асыруға болады.

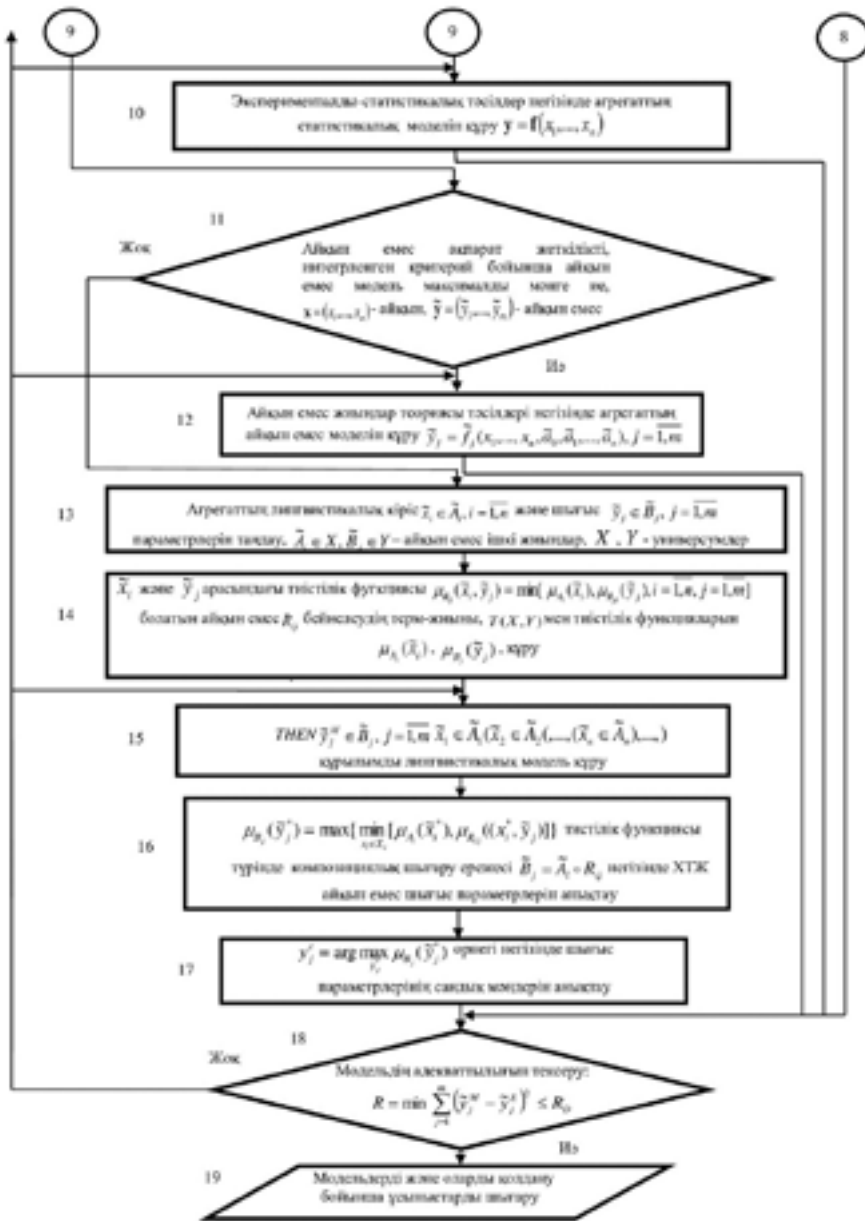
9-шы блоктағы шарт орындалмаса, онда 11-ші шартты блогында келесі шарт тексеріледі.

Егер ХТЖ-нің агрегаты параметрлері жайлы айқын емес ақпарат жеткілікті не қол жетімді, интегрленген критерий мәні бойынша айқынсыздықты модель құру максималды мәнге, ал векторлар $\mathbf{x} = (x_1, \dots, x_n)$ - айқын, ал $\tilde{\mathbf{y}} = (\tilde{y}_1, \dots, \tilde{y}_m)$ - айқын емес болса, онда 12-блокта айқын емес модель құрылады. Бұл блокта агрегаттың кірісі мен шығысын сипаттайтын айқын кіріс, режимдік параметрлері және айқын емес параметрлері таңдалады. Айқын емес модельдердің құрылымы $\tilde{y}_j = \tilde{f}_j(x_1, \dots, x_n, \tilde{a}_0, \tilde{a}_1, \dots, \tilde{a}_n)$, $j = \overline{1, m}$ құрылымы айқын емес көпкерегрессиялық тендеулер түрінде регрессорларды тізбектей қосу әдісі көмегімен анықтауға болады:

$$\tilde{y}_j = \tilde{a}_{0j} + \sum_{i=1}^n a_{ij} x_{ij} + \sum_{i=1}^n \sum_{k=i}^n a_{ikj} x_{ij} x_{kj}, j = \overline{1, m}.$$

Айқын емес модельдің құрылымы анықталғаннан кейін оның айқын емес параметрлерін идентификациялау есебі шешіледі, яғни, $\tilde{a}_0, \tilde{a}_1, \dots, \tilde{a}_n$ айқын емес регрессиялық коэффициенттер бағаланады. Бұл параметрлік идентификациялау есебін α деңгейлі жиын негізінде модификацияланған ең кіші квадраттар тәсілі көмегімен шешуге болады.





Сурет 1 – Қолжетімді әр түрлі сипаттағы ақпарат негізінде ХТЖ модельдерін құрудың жүйелік тәсілінің блок-сызбасы

Егер 9-шы блокта шарты тексерілгенде егрегаттың кіріс, режимдік және шығыс параметрлері векторлары айқын емес, яғни $\tilde{x} = (\tilde{x}_1, \dots, \tilde{x}_n)$ және $\tilde{y} = (\tilde{y}_1, \dots, \tilde{y}_m)$ болса, онда 13-ші блокта агрегаттың лингвистикалық моделі құрылады.

13-ші блокта лингвистикалық модельді құру үшін ХТЖ агрегатының айқын емес шығысына $\tilde{y}_j \in \tilde{B}_j, j = \overline{1, m}$ әсер ететін кіріс, режимдік $\tilde{x}_i \in \tilde{A}_i, i = \overline{1, n}$ параметрлері анықталады. Бұл параметрлерлер лингвистикалық айнымалылар арқылы сипатталады мұнда $\tilde{A}_i \in X, \tilde{B}_j \in Y$ – айқын емес ішкі жиындар, X, Y – әмбебап жиындар. Содан соң эксперттік бағалау тәсілдері арқылы агрегаттың айқын емес параметрлерін сипаттайтын терм-жинақтар $T(X, Y)$ анықталады және оның айқынсыз параметрлерінің тиістілік функциялары $\mu_{A_i}(\tilde{x}_i), i = \overline{1, n}, \mu_{B_j}(\tilde{y}_j), j = \overline{1, m}$ құрылады.

Айқынсыздықта мұнай өңдеу өндірісінің ХТЖ-ін модельдеу тәжірибесі негізінде айқын емес параметрлерді формалды сипаттайтын тиістілік функцияның келесі бейімделетін коэффициенттері бар құрылымы ұсынуға болады:

$$\mu_{B_j}^p(\tilde{y}_j) = \exp(Q_{B_j}^p \left| (y_j - y_{mdj})^{N_{B_j}^p} \right|), \quad (1)$$

Айқын емес параметрлердің тиістілік функциясының ұсынылған өрнегінде (1) енгізілген белгілемер: $\mu_{B_j}^p(\tilde{y}_j)$ - агрегаттың айқын емес шығыс параметрлерінің \tilde{B}_j -айқын емес ішкіжиынына тиістілік функциясы; p терм нөмері; $Q_{B_j}^p$ - тиістілік функцияның 0.5 деңгейінде идентификациялау барысында анықталатырн коэффицент; $N_{B_j}^p$ -айқынсыз параметрлердің термдерінің тиістілік функциясының анықталу облысын сипаттайтын коэффицент; y_{mdj}^p - анықталған p термге ең сәйкес келетін айқын емес айнымалы. Бұл айнымалы келесі шарттан анықталады $\mu_{B_j}(y_{mdj}^p) = \max_j \mu_{B_j}(y_j)$.

Бұл блокта содан соң кіріс және шығыс лингвистикалық айнымалылар арасындағы байланыстар анықталады, яғни \tilde{x}_i және \tilde{y}_j арасында айқын емес бейнелеулер \tilde{R}_{ij} құралады. Бұл бейнелеулерді есептеулерді қолдану ыңғайлылығы болу мақсатында олар келесі тиістілік функциялары арқылы байланысатар матрицасы түрінде келтіріледі (блок 14):

$$\tilde{\mu}_{\tilde{R}_{ij}}(\tilde{x}_i, \tilde{y}_j) = \min[\mu_{A_i}(\tilde{x}_i), \mu_{B_j}(\tilde{y}_j), i = \overline{1, n}, j = \overline{1, m}].$$

Лингвистикалық модельді келесідей логикалық қорытындылау ережелері түрінде құруға болады (блок 15):

$$IF \tilde{x}_1 \in \tilde{A}_1 (\tilde{x}_2 \in \tilde{A}_2, (\dots, (\tilde{x}_n \in \tilde{A}_n), \dots)) THEN \tilde{y}_j^M \in \tilde{B}_j, j = \overline{1, m}. \quad (2)$$

ХТЖ-нің зерттелетін агрегатының айқынсыз шығыс параметрлері мәндерін композициялық қорытынды $\tilde{B}_j = \tilde{A}_i \circ R_{ij}$ ережесінің арқылы анықтаға болады. Ол үшін шығыс айнымалыларының айқын емес мәндері максиминді

көбейтуді өрнектейтін тиістілік функциялар ретінде келесідей анықталады (блок 16):

$$\mu_{\tilde{B}_j}(\tilde{y}_j^*) = \max_{x_i \in X_i} \{ \min[\mu_{\tilde{A}_i}(\tilde{x}_i^*), \mu_{\tilde{R}_{ij}}((x_i^*, \tilde{y}_j^*))] \}. \quad (3)$$

Шығыс параметрлерінің сандық мәндерін $y_j^c, j = \overline{1, m}$ тиістілік функцияларды максималды мәндерінің аргументі ретінде келесі формула көмегімен анықтауға болады: $y_j^c = \arg \max \mu_{\tilde{B}_j}(\tilde{y}_j^*)$ (блок 17).

Егер де ХТЖ-нің агрегаттары күйін, жұмысын сипаттайтын теориялық мәліметтер, статистикалық деректер және айқын емес ақпарат жеткіліксіз немесе оларды жинау экономикалық тұрғыдан тиімсіз болса, онда түрлі тәсілге негізделген гибридік тәсіл негізінде құралымдалған (гибриді) модель құрылады. Бұл жағдайда құралымдалған модель түрлі сипаттағы қолжетімді ақпарат негізінде (теориялық, статистикалық, айқын емес) құрылады. Ол үшін нысанның белгілі бір параметрін сипаттау үшін әдістеменің жоғарыда сипатталған блоктарының әртүрлі комбинациялары қолданылады.

18-ші блокта құрылған модельдердің адекваттығы тексеріледі. Модель адекваттығын бағалау үшін келесі критерийді қолдануға болады:

$$R = \min \sum_{j=1}^m (y_j^M - y_j^E)^2 \leq R_D, \quad (4)$$

мұнда y_j^M и y_j^E – шығыс параметрлерінің модельдік және шынайы (эксперименталдық) мәндері; R_D – рұқсат етілген ауытқу.

Келтірілген (4) шарт орындалса, құрылған модель адекватты деп қабылданып, нысанның жұмыс режимдерін модельдеуге және оның тиімді жұмыс режимдерін анықтау үшін ұсынылады. Ал адекваттық шарты (4) орындалмаса, яғни модель адекватты болмаса, онда адекватсыздықтың себебі анықталып, осы жойып, адекваттықты қамтамасыз ету үшін сипатталған жүйелі тәсілдің сәйкес блоктарына өту қажет. Модельдің адекватсыздығының негізгі себептері кеесілер болуы мүмкін:

модельде агрегат жұмысына әсер ететін кейбір параметрлердің ескерілмей қалуы;

модельдің құрылымдық немесе параметрлік дұрыс идентификацияланбауы, т.б.

Нәтижелерді талқылау

Жұмыста жасақталып, өз ара байланысқан технологиялық агрегаттардың кешенінен тұратын ХТЖ-ді жүйелік модельдеу үшін ұсынылған тәсіл жүйелік талдау методологиясы негізінде түрлі формалды және формалсыз тәсілдерді жүйелі қолдануға негізделген. Бастапқы ақпараттың айқын емес жағдайында, нысан параметрлерінің айқынсыздық деңгейіне байланысты айқын емес модель және лингвистикалық модель құру тәсілдері ұсынылған. Ұсынылған тәсілде түрлі белгілі және әзірленген (айқын емес, лингвистикалық

модельдерді құру) тәсілдер біртұтас жүйе ретінде қолданатындықтан, жасақталған жүйелік тәсіл синергизм эффектісі мен әмерджендік қасиетке ие болады. Ал бұлар ұсынылған жүйелік тәсілдің артықшылығын білдіреді.

Технологиялық нысандардың айқын емес параметрлері мен көрсеткіштерін сипаттауға ұсынылған тиістілік функцияның құрвлымының (1) артықшылығы, функцияны эксперттер көмегімен тұрғызылған оның графикалық байланысына тез және дәл бейімдеуге мүмкіндік беретін Q_B^p және $N_{B_j}^p$ коэффициенттерінің болуында.

ХТЖ-дің өз ара байланысқан агрегаттарының модельдер жүйесін құрып, жүйелік модельдеуге ұсынылған жүйелік тәсілде нысанның айқын емес шығыс параметрлері максимінді көбейтуді өрнегі көмегімен олардың тиістілік функциялары арқылы (3) формула көмегімен анықталады, Ал ол айқынсыз шығыс параметрлерінің сандық мәндері $y_j^c, j = \overline{1, m}$, бұл (3) формула бойынша анықталған тиістілік функциясының максималды мәндерін қамтамасыз ететін аргументі ретінде анықталады. келесі формула көмегімен анықтауға болады. Сонымен ХТЖ-дің өз ара байланысқан агрегаттарының тиімді модельдерін құрып, бір модельдер жүйесіне біріктіретін ұсынылған тәсіл бастапқы ақпараттың тапшылығы мен айқынсыздығы жағдайларында түрлі қол жетімді мәліметтер мен ақпараттар негізінде агрегаттардың түрлі моделдерін құруға мүмкіндік беретін жүйелі тәсіл болып табылады. Мұнда қол жетімді бастапқы ақпараттың типіне және интеграленген критерий мәнене байланысты агрегаттардың детерминдік, стохастикалық, айқын емес немесе лингвистикалық модельдерін құрылып, олар ХТЖ-де жүретін процесс пен агрегаттар арасындағы ағындарды ескере отырып, бір жүйеге бірітіріледі.

Қорытынды

Бастапқы ақпарат тапшылығы және айқынсыздығы жағдайында күрделі ХТЖ агрегаттарының бір жүйеге біріктірілген модельдерін құруға арналған жүйелік тәсіл ұсынылып, зерттелген. Ұсынылған жүйелік тәсіл жүйелік талдау методологиясына, модельер құрудың дәстурлы тәсілдері мен айқынсыздықта модель құрудың әзірленген тәсілдеріне, оларды кешенді қолдануға негізделген. Зерттеу нәтижесінде жұмыстың мақсатына қол жеткізілгені, яғни ХТЖ жұмысын жүйелі модельдеуге, олардың өз ара байланысқан агрегаттарының модельдері пакетін бастапқы ақпараттың жетіспеушілі және айқынсыздығы жағдайында құруға мүмкіндік беретін жүйелі тәсіл жасақталған.

Зерттеу барысында келесі негізгі нәтижелер алынған:

ХТЖ-нің жеке агрегаттарының құрылуы мүмкін модельдерінің түрін жүйелік талдау және эксперттік бағалау тәсілдер негізінде олардың тиімді модельдерін таңдау және оларды құру тәсілдерін анықтап, жасақтау процедурасы әзірленген;

ХТЖ-дің агрегаттарының жағдайы, жұмыс жайындағы бастапқы ақпараттың айқынсыздығы деңгейіне байланысты, олардың айқын емес және лингвистикалық модельдерін құру әдістемесі ұсынылған жүйелік тәсіл құрамында жасақталған;

ХТЖ агрегтарының нақты бір моделін құруға қажетті ақпараттың тапшылығы жағдайында, қол жетімді түрлі ақпарат негізінде олардың құралымдалған, гибридті моделін құру ұсынылған;

Түрлі қол жетімді ақпарат негізінде құрылатын жеке агрегаттардың модельдерін бір модельдер пакетіне ХТЖ агрегаттарының өз ара байланыстары мен оларда өтетін технологиялық процесс тізбектерін ескере отырып біріктіру принципі ұсынылған. Бұл принцип бойынша технологиялық процесстің жүруіне байланысты бір агрегат моделінің шығысы басқа агрегаттардың кірісі ретінде қолданылады.

Сонымен, ХТЖ-дің өз ара байланысқан агрегаттарының модельдерін бастапқы ақпарат тапшылығы мен айқынсыздығы жағдайында құру үшін әзірленген жүйелік тәсіл дәстүрлі және ұсынылған айқын емес, лингвистикалық модельдер құру тәсілдерін бір жүйеге бірітіруге негізделген. Әзірленген жүйелік тәсілдің ерекшелігі мен артықшылығы, ол синергизм эффектісі мен әмердендік қасиеті негізінде аталған жағдайда ХТЖ-ді жүйелік модельдеуге мүмкіндік беретін тиімді модельдер жүйесін құруға мүмкіндік беретіндігінде.

ӘДЕБИЕТТЕР

Алиев Р.А., Церковный А.Э., Мамедова Г.А., 2018. Управление производством при нечеткой исходной информации. -М.: Энергоатомиздат. 2-изд. 2018. -253 с.

Гуцыкова С.В., 2017. Метод экспертных оценок. Теория и практика, -М., ИП РАН. 2017. -278 с. ISBN: 978-5-9270-0209-2.

Жоров М., 2015. Расчеты и моделирование термokatалитических процессов в нефтепереработке. -М: Энергоатомиздат, 2-изд. 2015, с. 307.

Карманов Ф.И., Острейковский В.А., 2019. Статистические методы обработки экспериментальных данных с использованием пакета MathCad, -М.: Радио и связь. 2019. -187 с. ISBN: 978-5-905554-96-4.

Кафаров В.В., Дорохов И.Н., 2018. Системный анализ химико-технологических процессов. Основы стратегии. -М.: Наука, 3-изд. 2018, с. 517.

Липин А.Г., 2018. Математическое моделирование химико-технологических систем: -М.: Химия. 2-изд. 2018. ISBN 978-5-9616-0243-2 (in Russ.).

Овечкин Г.В., 2018. Компьютерное моделирование. -М. Академия, 2018. 432 с.

Плановский А.Н., 2020. Процессы и аппараты химической и нефтехимической технологии. – 3-е изд., испр. и доп. – М.: Химия, 2020. – 498 с. (in Russ.).

Поспелов Д.А., 2018. Ситуационное управление: теория и практика. -М.: Наука, 3-е изд. 2018, стр. 305.

Рыжов А.П., 2017. Теория нечетких множеств и ее приложения. -М.: 2017. -278. ISBN: 978-3-540-70777-6

Ушева Н.В., Мойзес О.Е., Митянина О.Е., Кузьменко Е.А., 2019. Математическое моделирование химико-технологических процессов. -Томск: Изд-во Томского политехнического университета, 2019. – 135 с.

Coleman B., Babu J., 2018. Techniques of Model Based Control. UpperSaddle River: Prentice Hall PTR, 2018, p.576.

Dzhambekov A.M., Fyodorova O.V., 2018. Optimal control of the process of catalytic reforming of petrol fractions”, Bulletin of the Astrakhan State Tech. University Ser. control, computing tech. inform., 2018. V. 329, №. 2. -Pp.34–42.

Gumen M.I., 2019. Increasing of the Efficiency of the Reforming LG-35-11/300. Petroleum Processing and Petrochemistry. 2019. №11. -Pp. 54–57.

Jorgensen M., 2019. A Review of Studies on Expert Estimation of Software Development Effort. *Journal of Systems and Software*. 2019. V.70. -Pp. 37–60.

Keller T., Gorak A., 2020. Modelling of homogeneously catalysed reactive distillation processes in packed columns: Experimental model validation". *Comput. Chem. Eng.*, 2020. V. 48. -Pp. 74–93.

Mircea C., Agachi S., Marimoiu V., 2019. Simulation and Model Predictive Control of a UOP Fluid Catalytic Cracking. *Chemical Engineering and Processing*. 2019. V. 42. -Pp.67–75.

Mohaddecy S.R., Zahedi S., Sadighi S., Bonyad H., 2006. Reactor modeling and simulation of catalytic reforming process. // *Petroleum & Coal*. 2006. V. 48, № 3. -Pp. 28–35.

Orazbayev B.B., Ospanov E.A., Orazbayeva K.N., Kurmangazieva L.T., 2018. A Hybrid Method for the Development of Mathematical Models of a Chemical Engineering System in Ambiguous Condition // *Mathematical Models and Computer Simulations*, ISSN 2070-0482, 2018. V. 10. № 6, -Pp.748–758.

Pavlov S.Yu., Kulov N.N., Kerimov R.M., 2016. Improvement of Chemical Engineering Processes Using Systems Analysis.//*Theoretical Foundations of Chemical Engineering*. 2016. V. 53. № 2. -Pp.117–133.

Pinheiro C.C., Fernandes J.L., Domingues L., 2018. Catalytic Cracking (FCC) Process Modeling, Simulation, and Control // *Industrial Engineering Chemistry Research*, 2018. V. 51. № 1. -Pp. 1–29.

Reverberi A.P., Kuznetsov N.T., Meshalkin V.P., Salerno M., Fabiano B., 2019. Systematical Analysis of Chemical Methods in Metal Nanoparticles Synthesis//*Theoretical Foundations of Chemical Engineering*. 2019. V. 50. №1. -Pp. 63–75.

Sabzi H.Z., 2017. Developing an intelligent expert system for streamflow prediction, integrated in a dynamic decision support system for managing multiple reservoirs: a case study// *Expert systems with applications*. 2017. V. 82. №3. -Pp. 145–163.

Sharikov Yu.V., Petrov P.A., 2018. Universal model for catalytic reforming, *Chemical and Petroleum Engineering*, 2018. V. 25, №4. -Pp. 33–47.

Zhi-Wen Zhao, De-Hui Wang. Statistical inference for generalized random coefficient autoregressive model // *Mathematical and Computer Modelling*, 2018. V.56. № 3. -Pp.152–166.

Zimmermann H.-J., 2018. *Fuzzy Set Theory – and Its Applications*. Springer Science+Business Media, LLC. Fifth Edition 2018. p.525. ISBN: 978-94-010-3870-6.

REFERENCES

Aliiev R.A. Tserkovnyj A.E., Mamedova G.A., 2018. Production management with fuzzy source information. -M.: Energoatomizdat. 2nd ed. 2018. -253 p.

Coleman B., Babu J., 2018. *Techniques of Model Based Control*. UpperSaddle River: Prentice Hall PTR, 2018, p.576.

Dzhambekov A.M., Fyodorova O.V., 2018. Optimal control of the process of catalytic reforming of petrol fractions", *Bulletin of the Astrakhan State Tech. University Ser. control, computing tech. inform.*, 2018. V. 329. №. 2. -Pp.34–42.

Gumen M.I., 2019. Increasing of the Efficiency of the Reforming LG-35-11/300. *Petroleum Processing and Petrochemistry*. 2019. №11. -Pp. 54–57.

Gutsykova S.V., 2017. The method of expert assessments. Theory and practice, -M., IP RAN. 2017. -278 p. ISBN: 978-5-9270-0209-2.

Jorgensen M., 2019. A Review of Studies on Expert Estimation of Software Development Effort. *Journal of Systems and Software*. 2019. V.70. -Pp.37–60.

Kafarov V.V., Dorohov I.N., 2018. System analysis of chemical and technological processes. Basics of strategy. -M.: Nauka, 3rd ed. 2018, p. 517.

Karmanov F.I., 2019. Ostrejkovskij V.A Statistical methods for processing experimental data using the MathCad package, -M.: Radio i svyaz. 2019. -187 p. ISBN: 978-5-905554-96-4.

Keller T., Gorak A., 2020. Modelling of homogeneously catalysed reactive distillation processes in packed columns: Experimental model validation". *Comput. Chem. Eng.*, 2020. V. 48. -Pp. 74–93.

Keller T., Gorak A., 2019. Modelling of homogeneously catalysed reactive distillation processes in packed columns: Experimental model validation. *Comput. Chem. Eng.* 2019. V. 48. -Pp. 74–93.

- Lipin A.G., 2018. Mathematical modeling of chemical and technological systems: -M.: Chemistry. 2nd ed. 2018. ISBN 978-5-9616-0243-2 (in Russ.).
- Mircea C., Agachi S., Marimoiu V., 2019. Simulation and Model Predictive Control of a UOP Fluid Catalytic Cracking. *Chemical Engineering and Processing*. 2019.V. 42. -Pp. 67–75.
- Mohaddecy S.R, Zahedi S., Sadighi S., Bonyad H., 2006. Reactor modeling and simulation of catalytic reforming process. // *Petroleum & Coal*. 2006. V. 48. № 3, -Pp. 28–35.
- Orazbayev B.B., Ospanov E.A., Orazbayeva K.N., Kurmangazieva L.T., 2018. A Hybrid Method for the Development of Mathematical Models of a Chemical Engineering System in Ambiguous Condition // *Mathematical Models and Computer Simulations*, ISSN 2070-0482, 2018. V. 10. № 6. -Pp. 748–758.
- Ovechkin G.V., 2018. Computer simulation. -M. Akademiya, 2018. 432 p.
- Pavlov S.Yu., Kulov N.N., Kerimov R.M., 2016. Improvement of Chemical Engineering Processes Using Systems Analysis.//*Theoretical Foundations of Chemical Engineering*. 2016. V. 53. № 2. -Pp.117–133.
- Pinheiro C.C., Fernandes J.L., Domingues L., 2018. Catalytic Cracking (FCC) Process Modeling, Simulation, and Control // *Industrial Engineering Chemistry Research*, 2018. V 51. № 1. -Pp. 1–29.
- Planovskij A.N., 2020. Processes and devices of chemical and petrochemical technology. – 3rd ed. corrected and expanded. – M.: Chemistry, 2020. – 498 p. (in Russ.).
- Pospelov D.A., 2018. Situational management: theory and practice. -M.: Nauka, 3rd ed. 2018, p. 305.
- Reverberi A.P., Kuznetsov N.T., Meshalkin V.P., Salerno M., Fabiano B., 2019. Systematical Analysis of Chemical Methods in Metal Nanoparticles Synthesis//*Theoretical Foundations of Chemical Engineering*. 2019. V. 50. №1. -Pp. 63–75.
- Ryzhov A.P., 2017. Fuzzy set theory and its applications. -M.: 2017. -278. ISBN: 978-3-540-70777-6
- Sabzi H.Z., 2017. Developing an intelligent expert system for streamflow prediction, integrated in a dynamic decision support system for managing multiple reservoirs: a case study// *Expert systems with applications*. 2017. V. 82. №3. -Pp.145–163.
- Sharikov Yu.V., Petrov P.A., 2018. Universal model for catalytic reforming, *Chemical and Petroleum Engineering*, 2018. V. 25. №4. -Pp. 33–47.
- Usheva N.V., Mojzes O.E., Mityanina O.E., Kuzmenko E.A., 2019. Mathematical modeling of chemical and technological processes. -Tomsk: Publishing House of Tomsk Polytechnic University, 2019. – 135 p.
- Zhi-Wen Zhao, De-Hui Wang. Statistical inference for generalized random coefficient autoregressive model // *Mathematical and Computer Modelling*, 2018, V.56. № 3. -Pp.152–166.
- Zhorov M., 2015. Calculations and modeling of thermocatalytic processes in oil refining. -M: Energoatomizdat, 2nd ed. 2015, p. 307.
- Zimmermann H.-J., 2018. Fuzzy Set Theory – and Its Applications. Springer Science+Business Media, LLC. Fifth Edition 2018. p.525. ISBN: 978-94-010-3870-6.

©**M.Zh. Kaldarova**¹, **A.S. Akanova**^{1*}, **M.G. Grif**², **U.Zh. Aitimova**¹,
A.S. Mukanova³, 2023

¹S. Seifullin Kazakh agrotechnical research university, Astana, Kazakhstan;

²Novosibirsk State Technical University, Novosibirsk, Russia;

³Astana International University, Astana, Kazakhstan.

E-mail: kmiraj82@mail.ru

ALGORITHM AND METHOD OF PROCESSING SPACE PHOTOS FOR ASSESSMENT OF SOIL

Kaldarova Mira Zhorabekkyzy — doctoral student, S. Seifullin Kazakh agrotechnical research university, 010000. Astana, Kazakhstan

E-mail: kmiraj82@mail.ru. ORCID ID: 0000-0001-7494-9794;

Akanova Akerke Saparovna — PhD, Senior lecturer, S. Seifullin Kazakh agrotechnical research university, 010000. Astana, Kazakhstan

E-mail: akerkegansaj@mail.ru. ORCID ID: 0000-0002-7178-2121;

Grif Mikhail Gennadievich — Doctor of Engineering Sciences, professor, Novosibirsk State Technical University, 630082. Novosibirsk, Russia

E-mail: grif@corp.nstu.ru. ORCID ID: 0000-0003-3016-3647;

Aitimova Ulzada Zholdasbekovna — Candidate of Physical and Mathematical Sciences, acting associate professor, S. Seifullin Kazakh agrotechnical research university, 010000. Astana, Kazakhstan

E-mail: zada@mail.ru. ORCID ID: 0000-0002-0803-7137;

Mukanova Assel Serikovna — PhD, acting associate professor, Astana International University, 010000. Astana, Kazakhstan

E-mail: assel.mukanova@aiu.edu.kz. ORCID ID: 0000-0002-8964-3891.

Abstract. This article deals with research conducted in the field of remote sensing of the Earth and the use of space images. The purpose of this study is to use the algorithm of neural networks when determining soil conditions. Several methods of processing spatial images were used during the research: logistic regression, support vector method (SVM), belonging to linear classifiers, polynomial and radial basis network RBF, method of naive Bayes and K-nearest neighbors (KNN). According to each of these methods, soil erosion is classified by soil type, based on which a model is created and the results of the classification of these tests are analyzed. Looking at the results, it was established that each method has its own peculiarities. One of the methods may be best suited for wet soils, and the other is best suited for waterlogged soils. And for some soil types, all methods showed low results. But

in terms of accuracy, the KNN method was the best among the six tested methods with an accuracy of 90 %, and the second method was 2.5 % more accurate than the polynomial SVM. The possibilities of applying the results of the processing of space images obtained by remote sensing of the earth in agricultural practice are considered. The results can be used in the field of agriculture, when analyzing the necessary data. In particular, the article is aimed at comparing traditional data collection methods with modern methods using remote sensing results in deep learning to help producers make optimized and sustainable agricultural decisions.

Keywords: algorithm, machine learning, space images, model, image processing, neural networks, vector and linear data

Financing: The study was conducted from the authors own financial resources.

Conflict of interest: The authors declare that they have no conflict of interest in relation to this research, whether financial, personal, authorial or otherwise, which could affect the research and its results presented in this article.

©М.Ж. Қалдарова¹, А.С. Аканова^{1*}, М.Г. Гриф², У.Ж. Айтимова¹, А.С. Муканова³, 2023

¹С. Сейфуллин атындағы Қазақ агротехникалық университеті, Астана, Қазақстан;

²Новосібір мемлекеттік техникалық университеті, Новосібір, Ресей;

³Астана халықаралық университеті, Астана, Қазақстан.
E-mail: kmiraj82@mail.ru

ТОПЫРАҚ ЖАҒДАЙЫН БАҒАЛАУ ҮШІН ҚОЛДАНЫЛАТЫН ҒАРЫШТЫҚ СУРЕТТЕРДІ ӨНДЕУ АЛГОРИТМДЕРІ МЕН ӘДІСТЕРІ

Қалдарова Мира Жорабекқызы — докторант. С. Сейфуллин атындағы Қазақ агротехникалық зерттеу университеті. 010000. Астана, Қазақстан
E-mail: kmiraj82@mail.ru. ORCID ID: 0000-0001-7494-9794;

Аканова Акерке Сапаровна — PhD, С.Сейфуллин атындағы Қазақ агротехникалық зерттеу университеті. 010000. Астана, Қазақстан

E-mail: akеркеgansaj@mail.ru. ORCID ID: 0000-0002-7178-2121;

Гриф Михаил Геннадьевич — техника ғылымдарының докторы, профессор, Новосібір мемлекеттік техникалық университеті. 630082. Новосібір, Ресей

E-mail: grif@corp.nstu.ru. ORCID ID: 0000-0003-3016-3647;

Айтимова Улдада Жолдасбековна — физика-математика ғылымдарының кандидаты, С.Сейфуллин атындағы Қазақ агротехникалық зерттеу университеті. 010000. Астана, Қазақстан

E-mail: zada@mail.ru. ORCID ID: 0000-0002-0803-7137;

Муканова Асель Сериковна — PhD, доцент м.а., Астана халықаралық университеті. 010000. Астана, Қазақстан

E-mail: assel.mukanova@aiu.edu.kz. ORCID ID: 0000-0002-8964-3891.

Аннотация. Берілген мақалада Жерді қашықтықтан зондтау мен ғарыштық суреттерді қолдану саласында жүргізілетін зерттеулер қарастырылған.

Осы зерттеудің мақсаты болып, топырақ жағдайын анықтауда нейрондық желілердің алгоритмін қолдану болып табылады. Зерттеу барысында ғарыштық суреттерді өңдеу бойынша бірнеше әдіс қолданылған: логистикалық регрессия, сызықтық классификаторларға жататын тірек векторлар әдісі (SVM), полиномиялық және радиалды-негізгі функциялар желісі RBF, Naive Bayes және К-ең жақын көрші (KNN) әдісі. Осы әдістердің әрқайсысы бойынша жердің тозуы, яғни топырақтың түрлеріне қарай жіктеп, оларға сәйкес модель жасалып, тест деректерін жіктеу нәтижелері талданған. Нәтижелерге қарап, әр әдістің өзіндік ерекшеліктері бар екені анықталды. Бір әдіс ылғалды топырақтар үшін жақсы нәтиже берсе, екіншісі сулы топырақтар үшін жақсы нәтиже көрсетті. Ал кейбір топырақ түрлері үшін барлық әдіс, төмен нәтиже көрсетті. Бірақ дәлдік тұрғысынан KNN әдісі сыналған алты әдістің ішіндегі ең жақсысы болып шықты, 90 % дәлдікпен полиномиялық SVM әдісіне қарағанда 2,5 % дәлірек болды. Жерді қашықтан зондтау арқылы алынған ғарыштық суреттерді өңдеу нәтижелерін ауылшаруашылық тәжірибесінде қолдану мүмкіндіктері қарастырылған. Нәтижелерді ауылшаруашылық саласында қолдануға, қажетті деректерді талдауда қолдануға болады. Бұл осы сала мамандары үшін өз жерлерінің жағдайын анықтауда, соған сәйкес қажетті шараларды қолдануына мүмкіндік береді.

Атап айтқанда, мақалада дәстүрлі деректерді жинау әдістерін өндірушілерге оңтайландырылған және тұрақты ауылшаруашылық шешімдерін қабылдауға көмектесу үшін қашықтықтан зондтау нәтижелерін терең оқытуда қолдана отырып, заманауи әдістермен салыстыруға бағыттайтыны анықталған.

Түйін сөздер: алгоритм, машиналық оқыту, ғарыштық суреттер, модель, суреттерді өңдеу, нейрондық желілер

Қаржыландыру: Зерттеу авторлардың жеке қаржысынан жүргізілді.

Мүдделер қақтығысы: Авторлар зерттеуге және осы мақалада келтірілген нәтижелерге әсер етуі мүмкін қаржылық, жеке, авторлық немесе басқа зерттеуге қатысты мүдделер қақтығысы жоқ екенін айтады.

©М.Ж. Калдарова¹, А.С. Аканова^{1*}, М.Г. Гриф², У.Ж. Айтимова¹,
А.С. Муканова³, 2023

¹Казахский агротехнический университет им. С. Сейфуллина,
Астана, Казахстан;

²Новосибирский государственный технический университет,
Новосибирск, Россия;

³Международный университет Астана, Астана, Казахстан.
E-mail: kmiraj82@mail.ru

АЛГОРИТМЫ И МЕТОДЫ ОБРАБОТКИ КОСМИЧЕСКИХ СНИМКОВ ДЛЯ ОЦЕНКИ СОСТОЯНИЯ ПОЧВ

Калдарова Мира Жорабеккызы — докторант, Казахский агротехнический университет имени С.Сейфуллина. 010000. Астана, Казахстан

E-mail: kmiraj82@mail.ru. ORCID ID: 0000-0001-7494-9794;

Аканова Акерке Сапаровна — PhD, Казахский агротехнический университет имени С. Сейфуллина. 010000. Астана, Казахстан

E-mail: akerkegansaj@mail.ru. ORCID ID: 0000-0002-7178-2121;

Гриф Михаил Геннадьевич — доктор технических наук, профессор, Новосибирский государственный технический университет. 630082. Новосибирск, Россия

E-mail: grif@corp.nstu.ru. ORCID ID: 0000-0003-3016-3647;

Айтимова Улзада Жолдасбековна — кандидат физико-математических наук, ассоциированный профессор, Казахский агротехнический университет имени С. Сейфуллина. 010000. Астана, Казахстан

E-mail: zada@mail.ru. ORCID ID: 0000-0002-0803-7137;

Муканова Асель Сериковна — PhD, ассоциированный профессор, Международный университет Астана. 010000. Астана, Казахстан

E-mail: assel.mukanova@aiu.edu.kz. ORCID ID: 0000-0002-8964-3891.

Аннотация. В данной статье рассматриваются исследования, проводимые в области дистанционного зондирования Земли и использования космических снимков. Целью данного исследования является использование алгоритма нейронных сетей при определении почвенных условий. В ходе исследования использовались несколько методов обработки пространственных изображений: логистическая регрессия, метод опорных векторов (SVM), принадлежащий линейным классификаторам, полиномиальная и радиальная базисная сеть RBF, метод наивного Байеса и K-ближайших соседей (KNN). По каждому из этих методов эрозия почв классифицируется по типам почв, по ним создается модель и анализируются результаты классификации данных испытаний. Глядя на результаты, было установлено, что каждый метод имеет свои особенности. Один из методов может лучше всего подходить для влажных почв, а другой лучше подходит для заболоченных почв. А для некоторых типов почв все методы показали низкие результаты. Но по точности метод KNN оказался лучшим среди шести протестированных методов с точностью 90 %, а второй метод оказался на 2,5 % точнее полиномиального SVM. Рассмотрены возможности применения результатов обработки космических изображений,

полученных дистанционным зондированием земли, в сельскохозяйственной практике. Результаты могут быть использованы в области сельского хозяйства, при анализе необходимых данных. В частности, статья направлена на сравнение традиционных методов сбора данных с современными методами с использованием результатов дистанционного зондирования в глубоком обучении, чтобы помочь производителям принимать оптимизированные и устойчивые сельскохозяйственные решения.

Ключевые слова: алгоритм, машинное обучение, космические снимки, модель, обработка изображений, нейронные сети, векторные и линейные данные

Финансирование: исследование проводилось из собственных финансовых средств авторов.

Конфликт интересов: авторы заявляют, что у них нет конфликта интересов в отношении данного исследования, будь-то финансового, личного, авторского или иного, который мог бы повлиять на исследование и его результаты, представленные в этой статье.

Кіріспе

Жер үсті объектілерін мониторингілеу мақсатында жерді қашықтықтан зондтау деректерін қолдану, ауыл шаруашылығы жерлерінің, ландшафттың т.б жерлердің жай-күйі туралы ақпарат алуға ықпал етеді, ал геоақпараттық жүйелер ақпараттың барлық жиынтығын бақыланатын аумақтың электрондық (цифрлық) картасы, сондай-ақ әртүрлі агроландшафт объектілерінің байқалатын қасиеттерінің жай-күйі туралы ақпаратты енгізу, сақтау, өңдеу және шығару бағдарламалық құралдары мен әдістемелері түрінде ұсынуға мүмкіндік береді.

Қарастырылып отырған тақырыптың өзектілігі бір жағынан контекстуализациялау болса, екінші жағынан ғарыштық суреттермен жұмыс жасауға арналған балама бағдарлама мен алгоритмдердің тиімділігін зерттеу болып табылады. Жұмыста қашықтықтан зондтау, ауылшаруашылық тәжірибесінің қазіргі жағдайын қалай жақсартуға болатындығы қарастырылады. Атап айтқанда, топырақ, жер мен өсімдік түрлерін талдай отырып, баяндамада белгілі бір маусымда отырғызу дақылдарының маңызды қасиеттерін анықтау үшін қашықтықтан зондтауды қалай қолдануға болатындығы талқыланады.

Бағдарламаға мәліметтер жиынтығын жүктеу, сонымен қатар талдау негізінде пиксельдер ұсынатын жер суреттерінің ықтимал түрлерін жіктеу және жіктеу алгоритмінің жалпы нәтижелерінен басқа, алгоритм мәліметтер жиынтығын дәл және тиімді жіктей алатын әртүрлі машиналық оқыту модельдерін бағалау осы зерттеудің мақсаты болып табылады.

Қазақстан Республикасының Жер кодексінде қарастырылған міндеттер (ҚР жер кодексі N 442, өзг. 2023) қатарына алынған суреттерден ҚР

топырақтарының деректер базасын құру және Landsat жер серігінің суреттері деректерінде қандай модельдер ең тиімді анықтау, жіктеу, классификациялау модельдерін және бағалау мәселелерінің тиімділігін арттыру жатады.

Бұл ғарыш қызметтерінің техникалық мүмкіндіктерін ғана емес, сонымен қатар қашықтықтан зондтау деректері мен реттеуші ақпарат негізінде объектіні объективті түрде бағалауға қабілетті жасанды интеллекттің заманауи әдістерін қолдана отырып, деректерді тиісті түрде өңдеу мүмкіндігін қарастыру қажеттілігін ұсынады (Potarov et al., 2022: 17; Ma et al., 2021: 11). Зерттеулерде осы объектіні бақылау міндеттеріне қатысты экономикалық негізделген ұсыныстарды тұжырымдау қажеттілігін атап өту қажет.

Жіктеу әдістерінен басқа, бұзылған топырақты анықтауға және оның кеңістіктегі пайда болуын картаға түсіруге мүмкіндік беретін топырақ пен спектрлік шағылыстың арасында тікелей байланыс бар. Осылайша, жердің тозуы оның қалыптасу деңгейлері бойынша, мысалы, қабаттар, ағындар немесе жыралар сияқты, жерсеріктік деректерді қашықтықтан зондтау арқылы бейнелеудің қиындығы олардың спектрлік айырмашылықтарында жатыр. Топырақтың эрозия деңгейінің спектрлік белгісі болып оның минералды құрамы, топырақ құрылымы, ылғал және органикалық заттар әсер етеді (Spadoni et al., 2020: 100419). Топырақтың спектрлік шағылысуынан басқа, өсімдік жамылғысы қалыптасқан аумақтар, егістік алқаптары, су объектілері және тағы да басқалары сияқты беткейлік белгілердің спектрлік бөлінгіштігі, эрозияға ұшыраған топырақты спектральді түрде ажыратуға ықпал етеді. Өйткені өсімдіктер мен алқапты жерлер топыраққа қарағанда күрделі спектрлік қасиеттерге ие. Жақын инфрақызыл (NIR) аймақтағы сипаттамалардың спектрлік реакциясы өсімдіктерді тірі емес белгілерден ажыратуды жеңілдетеді. Келесі автор еңбегінде Landsat TM жолақтарының, әсіресе 4-жолағының (NIR) шағылыстыру мәндері мен эрозия жылдамдығы арасында өзара байланысты ашып көрсетілген (Shoko et al., 2016: 12). Осы тұжырымдарды қолдай отырып, ғалымдар спутниктік деректерді қолдана отырып, эрозияға ұшыраған, тұрақты және тұндырылған аймақтарды табысты анықтады. Спутниктік суреттер негізінде алынған өсімдік индекстері (вегетациялық индекс - ВИ) жер бетінің шағылысуы туралы ақпарат алудың негізгі қашықтық әдістерінің бірі болып табылады. Индекс топырақ эрозиясын картаға түсіру үшін белгілерді алудың қарапайым және жылдам әдісі ретінде қолданылады. Атап айтқанда, ВИ-тер топырақ эрозиясын зерттеуде жақсы жағынан танылды.

Жұмыста қашықтықтан зондтаудың қажеттілігін көрсетуде, балама жүйе мен терең оқыту құралдары мен жер, топырақ үлгілерін жіктеу үшін ұсынылған әдістерді қолданумен байланысты болғандықтан, қашықтықтан зондтауды жүйелеу және нақты қалай қолданылатынын түсіну, пайдалы әрі іргелі білім түрі болып табылады. Зерттеу дерекқоры спутниктің көмегімен өлшеуге болатын агрономиялық факторлардың кең спектрін қамтиды, оның

ішінде жасыл жамылғы, өсімдік температурасы және топырақ ылғалдылығы жатады. Сонымен қатар ұшқышсыз ұшу аппараттарынан алынған суреттер арқылы қажетті дерекқорларды жинақтау нәтижесінде ауылшаруашылық жерлерінің кадастрлық шекараларын анықтау үшін терең оқытуды қолдануға болады (Crommelinck et al., 2019). Осындай мәліметтер жиынтығын талдау үшін эмпирикалық әдістердің үлкен жиынтығын қолдана отырып, жұмыста мұндай деректерді талдау, жерді пайдалану мен нақты егіншілікті болжауда кеңінен қолдануға қалай әкелетіні көрсетілген. Зерттеу нәтижелері жерді қашықтықтан зондтау жүйелерінің сенімділігі мен қолданылуы ауылшаруашылық қоғамдастығының болашағына айтарлықтай әсер ететінін көрсетеді. Жұмыс ауыл шаруашылығында Deep Learning пайдалану және қашықтықтан зондтау саласында жүргізілген зерттеулердің жалғасы болып табылады. Атап айтқанда, осы жұмыста дәстүрлі деректерді жинау әдістерін өндірушілерге оңтайландырылған және тұрақты ауылшаруашылық шешімдерін қабылдауға көмектесу үшін қашықтықтан зондтау терең оқытуды қолдана отырып, заманауи әдістермен салыстыруға бағыттайды. Сонымен қатар орманды алқаптар жайлы ақпараттар алу, оларды белгілі бір органдарға беру мақсатында және онда болатын өрт жағдайларын анықтау үшін ғарыштық суреттерді сегменттеуде терең оқыту, яғни U-Net архитектурасын қолдану мүмкіндіктері жоғары (Wyniawskij et al., 2019: 5; Khryashchev et al., 2020: 5). Бұл жұмыс, жерді қашықтықтан зондтау аясында, машиналық оқыту мен спутниктік бейнелерді қолдану саласында жүргізілген зерттеулердің жалғасы болып табылады. Нақтырақ айтсақ, өндірушілерге оңтайландырылған және тұрақты шешімдер қабылдауға көмектесу үшін қашықтықтан зондтау / терең оқытуды қолдана отырып, деректерді жинаудың заманауи әдістерін қолдануға бағытталған. Қазіргі уақытта ҚазҒарыш Сапары немесе Egistik сияқты көптеген ұлттық компаниялар мен жеке ұйымдар бар, олар оңтайлы дақылдардың өнімділігін талдау үшін қашықтықтан зондтау спутниктерінен алынған мәліметтер мен суреттерге сүйенеді. Бұл ұйымдардың мақсаттары, осы жұмыстың мақсатына ұқсас. Яғни, топырақты жіктеу алгоритмін қолдану суреттелген жердің қай учаскелері егістік немесе ауыл шаруашылығына жарамды екенін анықтауға көмектеседі. Құнарлы жер барған сайын пайдалы тауарға айналатын үнемі өсіп келе жатыр және ауыл шаруашылығының тиімділігін арттырудың мұндай әдістері Қазақстан аумағының халқы үшін егіннің тұрақты өсуі, жерді тиімді пайдалану аясында қамтамасыз ету үшін қажет.

Жер мәселелеріу талдау үшін, деректерді алдын-ала өңдеудің келесі әдістері жиі қолданылады: суреттердің өлшемдерін бір өлшемге келтіру (256x256, 224x224, 128x128, 96x96), ұсынылған классқа жатпайтын фонды алып тастау, NDVI индексінің мәні белгілі бір шекті мәннен төмен пиксельдерді алып тастау, қосымша арналар ретінде оқыту мысалдарына қосылу арқылы, өсімдіктер мен ылғалдылықтың қосымша индекстерін есептеу, бір жылдағы

мәндерге сәйкес осы пиксельдерді орташа бір өлшемді векторға келтіріледі. Жер үсті объектілерін спутниктік суреттерден жіктеу тапсырмаларын шешу үшін бірнеше жылдық мәліметтер жиынтығы алынады және түрлі жүйелердің (Modis т.б.) деректері де алынуы мүмкін. Жоғары кеңістіктік ажыратымдылық маңызы бар деректерді бірнеше жылдық, оның ішінде бірнеше күндік түсіру жиілігінің деректерін қолдануға болады (Zou et al., 2020: 108985).

Сол сияқты келесі деректе әртүрлі айнымалылар жиынтығында көрінетін жолақтарды, текстуралық ақпаратты және спектрлік және морфометриялық индекстерді пайдалана отырып, сегменттеуге қарастырылған. Машиналық оқыту арқылы (бес алгоритммен) анықтауды құруда көрінетін ортофотографиялар мен фотограмметриялық тығыз нүкте бұлттарының тиімділігін анықталған. Кездейсоқ орман (RF) ең жақсы (99,8%), ал ішінара ең кіші квадраттар ең нашар жалпы дәлдікке (~ 60%) ие болды. > 95% дәлдікке тіпті сынып деңгейінде де қол жеткізуге болатынын анықталған. Рекурсивті белгілерді жою (RFE) айнымалыларды таңдаудың тиімді құралы болған, оның алты айнымалы нәтижесі барлық қол жетімді 31 айнымалыны қолданумен бірдей болған. Морфометриялық көрсеткіштер өндірушінің 82% дәлдігіне және пайдаланушының 85% дәлдігіне ие болып (сәйкесінше PA және UA) және спектрлік және текстуралық индекстермен үйлескенде, бұл жақсартуға үлкен үлес қосқан. Морфометриялық көрсеткіштер әрдайым қол жетімді бола бермейді, бірақ қызыл-жасыл-көк жолақтарға (RGB) текстуралық және спектрлік индекстерді қосқанда, PA 12% - ға, ал UA 6% - ға жақсарғанын байқаған. Жерді қашықтықтан зондау негізінде модельдеу эмпирикалық және физикалық тәсілдер оны жердің тозуы аспектісінде кеңінен қолданған (Schlosser et al., 2020:2397). Деректерді жүктегеннен кейінгі келесі қадам-оларды алдын-ала өңдеу және аугментациялау. Тапсырма үшін алдын-ала өңдеудің келесі әдістерін қолдану туралы шешім қабылданды: деректерді 0-ден 1-ге дейін қалыпқа келтіру және деректерді нөлдік ортаға келтіру. Бұл әдістер деректердің үлкен таралуын жоюға және деректерді бір түрге келтіруге мүмкіндік береді. Бұл түрлендірулер математикалық модельді оқыту процесін бастағанға дейін және бүкіл үлгі үшін класс болжамын алғанға дейін қолданылады.

Жіктеу алгоритмдерінің параметрлік немесе параметрлік емес түрлері бар, екеуі де суреттерді өңдеуде классификациялау оқу мәліметтерін енгізуді қажет етеді және топырақ эрозиясын анықтауда белгілерді сәйкестендіру және картаға түсіруде қолданылады. Сонымен қатар жіктеудің бірнеше тәсілдері бар, мысалы, бақыланатын, бақыланбайтын немесе гибриді (бақыланатын және бақыланбайтын жіктеулердің жиынтығы) әдістері қарастырылған (Bhatnagar et al., 2020: 2505). Бұл жіктеу әдістері әдетте математикалық алгоритмдермен бірге қолданылады, олар мыналарды қамтиды: интерактивті өзін-өзі ұйымдастыратын деректерді талдау әдісі (ISO DATA), максималды ықтималдылықты жіктеу (MLC), тірек векторлық Машиналар (SVM) және

анық емес жиындарды жіктеу логикасы. Статистикалық ықтималдық функцияларына негізделген жіктеуіштер, әдетте, жіктеуге (объект жіктеуіштері) көмектесу үшін бір пиксельдің спектрлік епекшелігін (пиксель жіктеуіштері) және пиксельдің айналасындағы кеңістіктік контексті пайдаланады. Дәстүрлі қашықтықтан зондтау ғалымдарының ғана емес, сонымен қатар дәстүрлі суперкомпьютерлерді немесе ауқымды бұлтты есептеу ресурстарын пайдалану үшін қажетті техникалық мүмкіндіктері жоқ әлдеқайда кең аудиторияның мүмкіндіктерін кеңейтуге арналған интеграцияланған платформа ретінде бірегей (Gorelick et al., 2017: 10).

SVM жіктеуіштері қашықтықтан зондтау мультиспектрлі деректеріне SVM жіктеуіштері топырақ жамылғысының басқа түрлерімен салыстырғанда топырақ эрозиясын анықтауға арналған перспективтік параметрлік емес жіктеу әдісін ұсынады. SVM –ның басты күші болып, бұл модель әр кластың мүшелерінен оңтайлы жазықтыққа дейінгі қашықтықты максималды ететін гипержазықтықтарды орналастыру арқылы кластарды бөлу қабілетінде. Бір классты жіктеу үшін, SVM-ді пайдалану кезінде алынған параметрлер мен дәлдік деңгейлері туралы ақпарат өте аз. Алғашында, SVM орман өрттерін анықтау және SPOT 5 жерсеріктік суреттерінде қалалық аймақтарды анықтау үшін қолданылған. Қашықтан бейнелеуді қолдана отырып, жоспарланбаған қалалық елді мекендегі жеке ғимараттардың автоматты сипаттамасын және күрделі ландшафттар мен елді мекендердің тығыздығына байланысты зондтау арқылы U-net негізіндегі қалалық ауылдарды картаға түсіру парадигмасы ұсынылған. Зерттеу мақсаттары ретінде кеңістіктік ажыратымдылығы 0,5 м және сегіз панорамалық жолағы бар worldview спутниктік суреті пайдаланылған. SVM-ді Landsat — TM көмегімен басқа алгоритмдермен салыстырып, дәлдігі 86 % жеткен (Pan et al., 2020: 1574).

Соңғы жылдары конволюциялық нейрондық желі (CNN) кескінді өңдеу үшін машиналық оқытуда басым парадигмаға айналды. Келесі зерттеуде U-Net модификацияланған архитектурасын спутниктік кескіндерден контекстік ақпаратты алу үшін қолданған. Зерттеу DenseNet архитектурасын енгізу арқылы, қайталанатын белгілер карталарын пайдалануды және желі арқылы ақпараттың таралуын күшейтеді. Субдискретизация мен күшейтілген іріктеу арасында ұзақ қашықтықты өткізіп жіберуге байланысты байланыстар жүзеге асырылған. Ұсынылған архитектура сәйкесінше сандық салыстырудың қиылысумен біріктірілуі 73,02 % және жалпы дәлдігі 96,02 % көрсеткен (Soni et al., 2019: 13).

Көптеген зерттеулер NDVI өнімдері топырақтың түсіне, ылғалдылығына, атмосфералық жағдайларға және т.б. байланысты тұрақсыз болуы мүмкін екенін растайды. Топырақ эрозиясының кеңістіктік таралу карталарын жасау кезінде Landsat (NDVI, SAVI және SARVI) көмегімен алынған үш ВИС қолданылды және олар алынған топырақ эрозиясының VI карталарының дәлдігін бағалады және топырақ эрозиясының белгілерін су жинау деңгейінде анықтау үшін ең жақсы мәні VI екенін анықтады (Pan et al., 2019), ал Rep

Н., Zhou G және т.б. авторлар өз еңбектерінде қызыл және NIR жолақтары «топырақ сызығы» тұжырымдамасын қолдану арқылы топырақ эрозиясын анықтауға болатынын ашып көрсетті (Ren et al., 2018:7).

Суреттердің сапасы мен қасиеттеріне ғарыш кемесі орбитасының пішіні мен биіктігі әсер етеді, яғни, орбита шеңбері, жер беті түсірілімінің бірдей биіктігі. Түсіру барысында да қозғалып кету нәтижесінде суреттерде қателіктер пайда болады. Сондай суреттердің бірі 2-ші суретте көрсетілген, яғни түсірілім кезіндегі істен шыққан пиксельдерді өңдеу нәтижесі. (Soloviev et al., 2016: 7).

Сонымен, ғылыми зерттеулерді қарастырып топырақ жағдайын бағалау үшін ғарыштық суреттерді өңдеуде қолданылатын бірнеше әдістерді қарастырып олардың ішіндегі тиімдісің таңдау өзекті мәселелердің бірі боллатыны айқындалды.

Әдістер мен материалдар

Ғарыштық суреттерді өңдеу үшін зерттеу барысында келесі әдістер қолданылды: талдау, салыстыру, деректерді визуализациялау, бағдарламалау.

Қазіргі таңда ғарыштық суреттерді өңдеу барысында келесі кезеңдер қарастырылады:

- жобалық құжаттаманы ескере отырып, ғарыштық түсіру деректерін қолдану аясында, объектінің техникалық сипаттамаларын аламыз;
- нормативтік деректер бойынша қажетті материалдардың санын және жұмыстың еңбек сыйымдылығын есептеу;
- ғарыштық түсіру және деректерді шығару деректерін пайдалана отырып, материалдардың қажетті көлемінің және еңбек сыйымдылығының сәйкестігін талдау;
- ақпараттық-талдау жүйесі негізінде орындаушылардың техникалық деңгейі мен технологиялық дайындығын ескере отырып, тапсырыс берушілер үшін орындаушыларды оңтайлы таңдау бойынша ұсыныстар қалыптастыру;
- автоматты сараптама жүйесін пайдалана отырып және сараптама жүйесін есептеу хаттамаларын ұсына отырып, материалдарды жеткізушілер және жұмыстарды орындаушыларды таңдау туралы шешім қабылдау. Материалдарды жеткізушілермен және жұмыстарды орындаушылармен шарттар жасасу;
- ЖҚЗ деректерін және есептік құжаттаманы пайдалану және жұмыстардың орындалу барысының үздіксіз мониторингін жүргізу;
- Data Mining / Data Science әдістері мен деректерді іздеу негізінде мониторинг нәтижелерін үнемі бақылау;
- жобаны орындауды аяқтау және сараптама жүйесінің нәтижелерімен расталған құжаттама жиынтығын қалыптастыру;
- есептеудің қарапайымдылығына байланысты ең көп қолданылатыны-қалыпқа келтірілген вегетациялық индекс (Normalized Difference Vegetation Index, NDVI).

Жерді пайдалануды оңтайландыру саласындағы ғылыми және практикалық міндеттерді шешу үшін ғарыштық түсірілім материалдарын тану мен дешифрлеудің сапалық және сандық сипаттамаларын арттыру міндетін шешудің өзектілігі ақпарат алудың жаңа әдістерін іздеудің, сондай-ақ қазіргі бар әдістердің тиімділігін арттырудың өзектілігін айқындайды.

Ғарыштық бейне түсірілімдерін талдау үшін 7-ші және 8-ші буындағы Landsat спутниктерінен материалдар алынады. Олардың кеңістіктік рұқсаты жеке орман аудандарын, ауылшаруашылық жерлерін, елді мекендердің аудандарын зерттеуге және әртүрлі санаттағы жерлердің геокеңістіктік құрамын ескеруге мүмкіндік береді. Осы спутниктерден алынған мәліметтердің спектрлік ажыратымдылығы көрінетін, жақын және қысқа толқындық инфрақызыл диапазондардағы геофизикалық қабықтардың қасиеттері туралы ақпарат береді.

Кеңістіктік объектілер мен жердің жай-күйінің неғұрлым ақпараттық кеңістіктік және уақытша сипаттамаларын алу үшін, әлемдік тәжірибе бойынша әдеттегідей аумақтың спектрлік параметрлеріне ғана емес, сонымен қатар геофизикалық объектіні (айналаны) қоршап тұрған аумақтың инвариантты және динамикалық қасиеттерін сипаттайтын мәліметтерді жүйелі талдауға сүйену керек. Бағдарлама Amazon Web Services-арқылы Landsat 8 спутнигінен суреттер және басқа да файлдарды жүктей алады. 11 жолақты қай аймақпен жұмыс жасағыңыз келетініне байланысты сұраныс бойынша жүктеуге болады. Таңдалған көрініспен жұмыс жасау үшін қажет барлық файлдар жүктеледі. Файлдар жергілікті дискіге сақталғаннан кейін көріністер жасалады. Жолақтарды белгілі бір есептеу немесе композиттеу (бірнеше бөліктен немесе элементтері біріктіру үшін қандай жолақтар қажет екеніне байланысты дербес немесе автоматты түрде жүктеуге болады.

WRS (Worldwide Reference System) — бұл Landsat спутниктерінен алынған деректерді индекстеудің жаһандық жүйесі. Іс жүзінде сурет алынған жер бетінің учаскесін көрсетеді. Ол үшін екі параметр қолданылады: Path және Row. Path параметрі батыстан шығысқа қарай "аймақты" белгілейді; Row параметрі солтүстік — оңтүстік бағытта "аймақ" нөмірін анықтайды. "c1" — 1-топтаманы білдіреді, "L8" — Landsat 8-ді білдіреді, "139" — аймақ жолы, "045" - аймақ жолын білдіреді, ал соңғы каталог келесі атау туралы келісімді қолданатын өнім идентификаторына сәйкес келеді (сурет-1):

L = Landsat

X = Сенсор

SS = Спутник

PPP = WRC жолы

RS = WRC жол

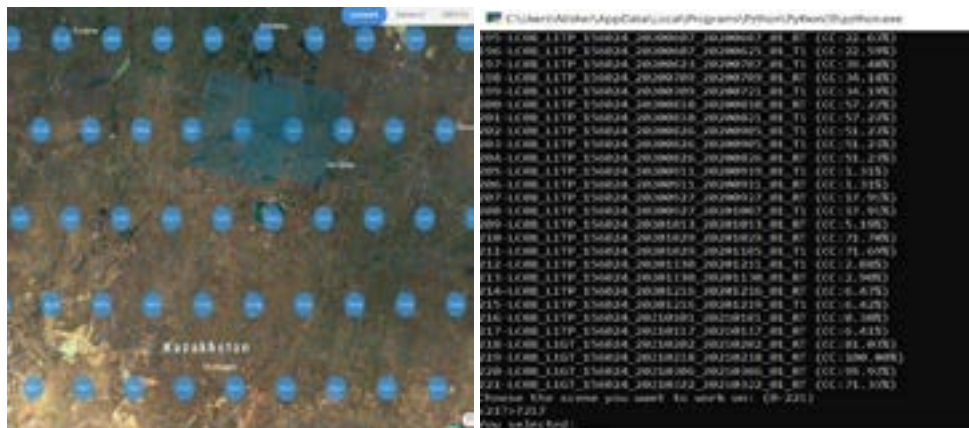
ЖЖЖЖ ММ КК = түсірілген күні

ууууmdd = өндеу күні

CC = коллекция нөмірі

TX = коллекция санаты

LC08_L1TP_156024_20210117_20210117_01_RT --bands NDVI



Сур.1. WRS (Worldwide Reference System) жүйесі мен файлдарды жүктеу барысы
(Fig.1. WRS (Worldwide Reference System) System and file download progress)

Деректер базасы бейнеленген көріністің кішкене аймағынан алынған көп спектрлі пиксель мәндерінен тұрады. Мұндай пикселдер осы спутниктік кескіннің шегінде 3x3 төңірегінен алынады. Сонымен қатар, дерекқордағы әр пиксель 0-ден 255-ке дейінгі мәнге сәйкес келеді. Сонымен қатар, пиксельдің әр жолы үшін 3 3x3 жіктеу белгісі бар. Содан кейін 9 пиксельдің әр айналасы үшін төрт спектрлік диапазон бар, олар мәліметтер жиынтығында 37 бағанды құрайды (36 пиксель және 1 жіктеу белгісі). Атап өткендей, 36 пиксельде тәуелді айнымалылар болады, ал жіктеу белгісі эксперименттің тәуелсіз айнымалысы болып табылады. Мәселенің сипаттамасын қайта құру үшін қазіргі мақсат 36 пиксельді ескере отырып, жіктеу нөмірін болжау болып табылады. Ең алдымен, бұл мәліметтер жиынтығы оқыту және тест жиынтықтарына бөлінді: 4435 данадан тұратын оқыту және 2000 данадан тұратын тест мәліметі. Деректер жиынтығында алты түрлі топырақ кластары бар, олардың белгілері төменде көрсетілген:

1. Қызыл топырақ
2. Мақта дақылдары
3. Сұр топырақ
4. Ылғалды Сұр топырақ
5. Орылған егістік
6. Өте ылғалды сұр топырақ

Бұл жұмыс үшін негізгі міндет-кескіннің кішкене ауданы үшін пиксель мәндері бойынша топырақ түрін сәтті анықтау, сондықтан талдау процесін автоматтандыру үшін жіктеу моделін жасау қажет. Жіктеу — бұл машиналық оқытудың кең қолданысы және жіктеу модельдерінің бірнеше түрлері бар.

Жобада төрт әдіс қолданылды: логистикалық регрессия, тірек векторлық машиналар (SVM), Naive Bayes және ең жақын К-көрші (K-NN). Қазіргі уақытта әртүрлі стратегиялардың нәтижелерін талдау және қайсысы ең қолайлы нәтиже беретінін анықтау өте маңызды. Әрбір жіктеу моделін талдағаннан кейін олардың екеуі алдын-ала анықталған критерийлер негізінде таңдалады және мәліметтер жиынтығындағы іздеу нәтижелеріне қосымша талдау жасалады. Есептің осы бөлімінде модельдерге негізделген негізгі тенденциялар визуалды элементтер мен мәліметтер арқылы бағаланады.

Сонымен қатар, мұндай әдістер деректер арасындағы заңдылықтардың табиғаты туралы көп түсінік бермейді, бірақ тиімділігі жоғары және дәл мәліметтер береді деп күтуге болады. Осылайша, KNN әдісі нақты қарапайым классификатор құру мақсатында жарайды, дегенмен бұдан басқа қосымша ақпарат бере алмайды. Осы себептен, соңғы көрсетілген нәтижелер бұрын айтылған әдістердің нәтижелеріне сүйенеді деп күтуге болады. KNN моделі ешқандай гиперпараметрлерді қажет етпейді. Бұл әдістің алгоритмі алдыңғы үш модельден біршама ерекшеленеді. Ең жақсы өнімділік параметрлерін қамтамасыз ететін k мәнін таңдау маңызды болғандықтан, алдымен ең жоғары дәлдік көрсеткіші бар k мәні табылғаннан кейін модель жасалады және жаттығу деректерімен салыстырылып келістіріледі, содан кейін тест деректерімен болжамдар жасалады. KNN моделінде жұмыс уақыты бойынша қиындықтар бар, өйткені k мәнін табудың итеративті әдісі басқа модельді бірнеше рет құруды талап етеді. Бұл бағдарламаның жалпы орындалуына үлкен уақыт қосуы мүмкін. Осы мәселелерді жеңілдету үшін біз k мәнінің итеративті функциясын тек бір рет іске қоса аламыз немесе деректерді масштабтауға болады, осылайша келістіру мен болжауға аз уақыт кетеді.

SVM (Polynomial) шатастыру матрицасы, бұл жағдайда модельдің суретте көрсетілген логистикалық регрессия әдісіне қарағанда біршама дәлдікке ие (шамамен 3 %). екенін көрсетеді. Сонымен қатар, сол SVM топтарының алдыңғы әдісімен салыстырғанда, сызықтық модель-көпмүшелік әдіс әлдеқайда жақсы дәлдікке ие екені анықталды. Сонымен қатар, әрі қарай қарастыру кезінде SVM моделінің жалғыз қателігі қайтадан ылғалды сұр топырақты жіктеумен байланысты болды. Мұндай жағдайларда модель нақты ылғалды сұр топырақтың жартысына жуығы дұрыс болады деп болжады, ал екінші жартысы өте ылғалды сұр топырақ санатына жатқызылды. Логистикалық модельге қарағанда біршама дәл болумен қатар, жұмыс уақыты әлдеқайда жылдам (4,95 с-қа қарағанда 350 мс). Осылайша, осы екі көрсеткішті де қарастырғанда SVM көпмүшесі анағұрлым оңтайлы болады.

Радиалды базистік функцияның (RBF) шатастыру матрицаларын SVM полиномиялық әдістерімен салыстырған кезде олар бірдей болатындығын ескеру қажет. Деректер жиынтығына кіретін мәндер басқа жиынтықтар сияқты күрделі болғандықтан, осы екі әдіс арасындағы ұқсастықты күтуге болады. Басты айырмашылықтардың бірі болып, RBF әдісінің жұмыс уақыты

бойынша шамамен 12,6 % жылдам. Осылайша, оңтайлы шешімді таңдағанда, RBF әдісі жақсы болар еді, өйткені көпмүшелік SVM есептеу уақытын ұзартылуы классификатор үшін зиян болып табылады. SVM алгоритмінің негізгі параметрлері келесі кестеде көрсетілген (1-кесте).

1 – кесте. Берілген әдістердің гиперпараметрлері

Әдісі	Қателік	Inverse of Regularization Strength (c) (Қайта реттеу күші)	Degree	Max Iterations (итерациялардың максималды саны)
Логистикалық регрессия	16	1.0	-	2950
SVM (Linear)	24	1.0	64	500
SVM (Polynomial)	13	1.0	2	1000
SVM (RBF)	13	1.0	2	1000

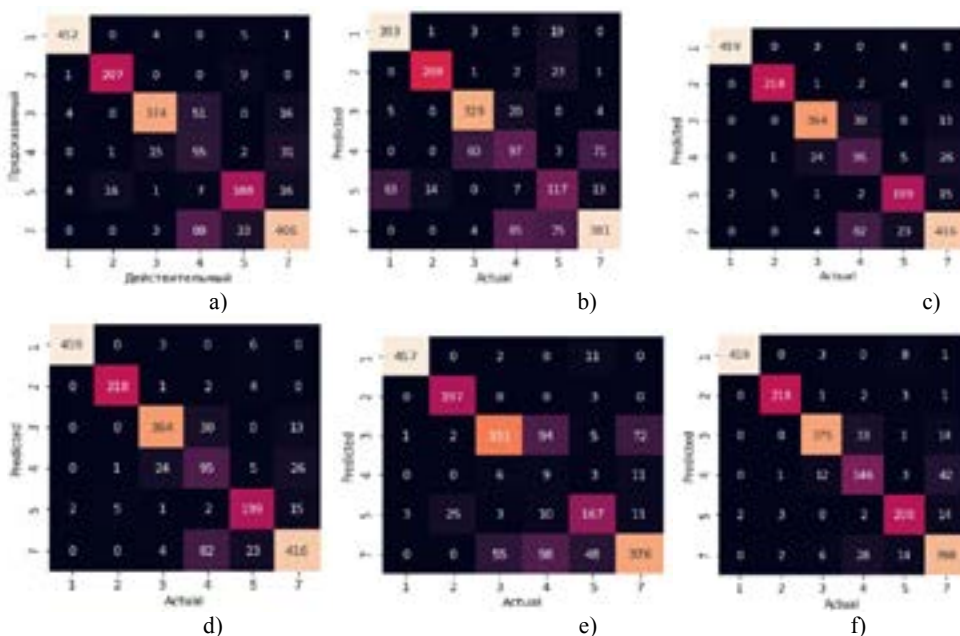
Naïve Bayes моделінің шатастыру матрицасы бойынша жұмыс уақыты мен дәлдігі кестеде көрсетілген. SVM-нің басқа әдістерімен салыстырған қатты айырмашылық жоқ. Сондықтан бұл әдісті де қолдануға болады.

Келесі суреттерде (сурет – 2) шатастыру матрицалары құрылып, әрбір модельдің өзінің жіктелуіне сәкес мәліметтер алынды. Көптеген жіктеу белгілері үшін логистикалық регрессия әдісі салыстырмалы түрде баяу уақыт кезеңіндегі нақты белгімен салыстырғанда берілген сынақ жиынтығын дәл болжай алды (2 а – сурет).

Келесі шатастыру матрицасы осы модельдің өзінің жіктелуінде логистикалық регрессия әдісіне қарағанда дәлірек емес (шамамен 8 %) екенін көрсетеді (2 б – сурет). Алайда, әрі қарай қарастырған кезде сызықтық SVM сұр топырақты дәл жіктеп қана қоймай, сонымен қатар қызыл топырақ пен өсімдіктері бар топырақ жағдайларын да жіктеді. Қызыл топырақ жағдайлары үшін шамамен 10 % - ы өсімдігі бар топырақ ретінде дәл болжанбаған. Содан кейін, өсімдіктердің нақты жағдайлары үшін модель олардың жартысына жуығы өте ылғалды сұр топырақ болады деп нақты болжаған жоқ. Сонымен, модель 76 % дәлдіктің салыстырмалы түрде жақсы бағасына ие болса да, топырақтың белгілі бір түрлері үшін жеке деңгейде мәселелер бар. Бұл алдыңғы логистикалық регрессия матрицасынан айырмашылығы, тек сұр топырақ жағдайында қиындықтар болған сияқты.

K-жақын көршілер (KNN) әдісінің тестілеу және оқыту нәтижелері K-нің әртүрлі мәндері үшін жасалды, сондықтан ол k=4 мәні ең дәл деп анықталды, өйткені оқыту мен тестілеу балдары арасындағы дисперсия минималды болды және соған сәйкес шатасу матрицасы құрылды. Дәлдік тұрғысынан KNN әдісі сыналған алты әдістің ішіндегі ең жақсысы болып шықты. 90 % дәлдікпен ол екінші әдіс полиномиялық SVM ге қарағанда 2,5 % дәлірек болды. Келесіде, екі шатастыру матрицасын салыстырған кезде, KNN сәтті болатын жетілдірудің негізгі бағыты - оның сұр топырақты дәл жіктеу қабілеті. Жоғарыда айтылғандай, полиномиялық SVM тек ылғалды сұр топырақты 52

% дәлдікпен болжады, бұл басқа жеке белгілерге қарағанда әлдеқайда аз. Мұнда KNN 70 % ылғалды сұр топырақ үшін дәлдік көрсеткішін шығарды, ол орташа дәл болса да, бес әдіске қарағанда әлдеқайда жақсы. Жоғары дәлдігіне қарамастан, KNN әдісі алдыңғы әдістер арасында ең баяу жұмыс уақытына ие. Миллисекундтар ішінде жұмыс уақыты бар SVM әдістерімен салыстырғанда KNN әдісі әлдеқайда баяу болды. Сонымен, салыстырмалы түрде кішігірім мәліметтер жиынтығы үшін модель өзін жақсы көрсетеді. Егер бұл нәтижелер кең ландшафтты қамтитын суреттер үшін әлдеқайда үлкен мәліметтер жиынтығына экстраполяцияланған болса, KNN ұзақ жұмыс уақытын зиянды факторға айналдыруы мүмкін.



Сур.2. Өңдеу әдістерінің шатастыру матрицалары (а – логистикалық регрессия, b - SVM (Linear), c - SVM (Polynomial), d - SVM (RBF), e - Naive Baye, f – KNN жақын көршілер) (Fig.2. Confusion matrices of processing methods (a-logistic regression, b-SVM (Linear), c - SVM (Polynomial), d - SVM (RBF), e - Naive Baye, f – KNN near neighbors))

Әзірленген жіктеу моделдері үшін нәтижелер мен көрсеткіштерді мұқият талдағаннан кейін ең оңтайлы екеуі таңдалады. Енді әртүрлі жіктеу әдістерінің тиімділігін зерттеуге таңдалып алынған Landsat мәліметтер жиынтығын бағаланады. Таңдалған екі әдісті қолдана отырып, топырақтың болжамды классификациясының гистограммасы құрылады. Бұл диаграмма нәтижелердің дәлдігін және топырақтың қандай түрлері модельдер үшін қателіктер тудырғанын жақсы көруге мүмкіндік береді.

Тұтастай алғанда, көп спектрлі деректерді талдауға жасанды интеллектті

енгізу өте маңызды, өйткені топырақ кескіндерін жіктеуге жеткілікті оқытылған деректерсіз таза математикалық алгоритмді жасау өте қиын болуы мүмкін; түсірілген кескіндер әрқашан әртүрлі болып келеді. Топырақ әрдайым спутниктік суреттер үшін толықтай ашық және мөлдір бола бермейді, ал кемшіліктер әрқашан айқын көрінеді. Сондықтан топырақтың күрделі және бұлыңғыр кескіндерін анықтау үшін оқытудың үлкен көлемін пайдаланатын жасанды интеллект классификациясының моделін жасау қажет.

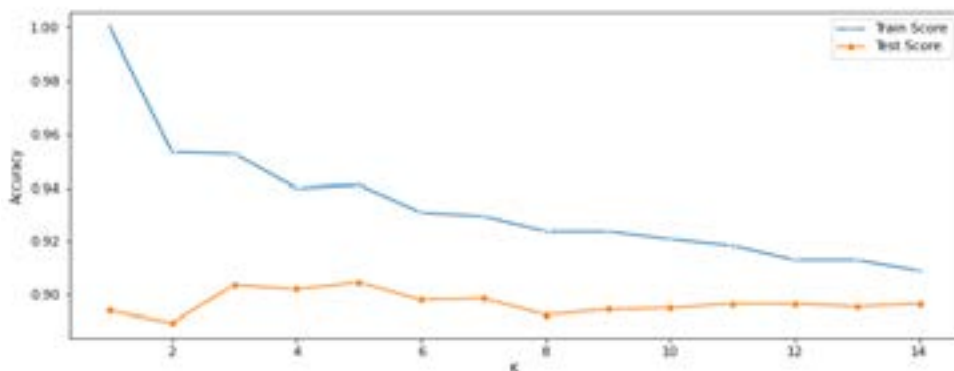
Нәтижелер

Бірдей мәліметтер жиынтығын қолдана отырып, әртүрлі жіктеу модельдерінің нәтижесін келесі кестеден (2-кесте) көруге болады. Бұл жердегі мақсат жинақтылық, дәлдік және жауап беру уақыты тұрғысынан жақсы жұмыс істейтін оңтайлы әдісті анықтау болды.

2 - кесте. Әдістердің орындалу уақыты мен дәлдігі

Көрсеткіштер/ әдістер	Логистикалық регрессия	SVM (Linear)	SVM (Polynomial)	SVM (RBF)	Naive Bayes метрикасы	KNN жақын көрші
Орташа орындалу уақыты	4.96 сек	0.219 сек	0.393 сек	0.35 сек	0.049 сек	11.73 сек
Дәлдігі	84.2	76 %	87.6 %	87.5 %	77 %	90 %

Дәлдік тұрғысынан KNN әдісі сыналған алты әдістің ішіндегі ең жақсысы болып шықты. 90% дәлдікпен ол екінші әдіс полиномиялық SVM ге қарағанда 2,5 % дәлірек болды. Келесі суретте (сурет -3) KNN әдісінің тестілеу нәтижесі берілген.



Сур.3. Жақын көршілер үшін тестілеу және оқыту нәтижелері (Fig.3. Test and training results for the nearest neighbors)

Полиномиялық SVM тек ылғалды сұр топырақты 52 % дәлдікпен болжады, бұл басқа жеке белгілерге қарағанда әлдеқайда аз. Мұнда KNN 70 % ылғалды сұр топырақ үшін дәлдік көрсеткішін шығарды, ол орташа дәл болса да, бес

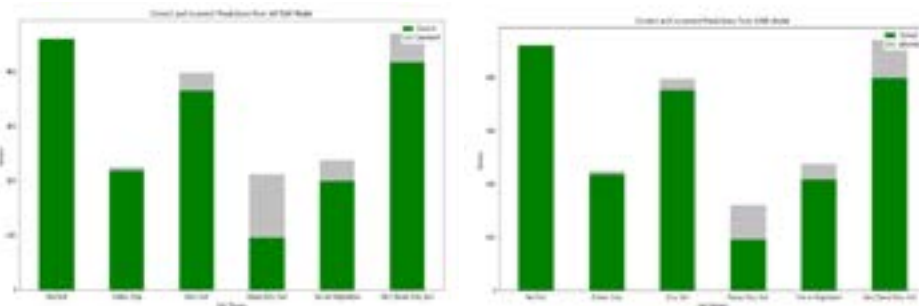
әдіске қарағанда жоғары нәтиже. Жоғары дәлдігіне қарамастан, KNN әдісі алдыңғы әдістер арасында ең баяу жұмыс уақытына ие. Миллисекундтар ішінде жұмыс уақыты бар SVM әдістерімен салыстырғанда KNN әдісі әлдеқайда баяу болды. Егер бұл нәтижелер кең ландшафтты қамтитын суреттер үшін әлдеқайда үлкен мәліметтер жиынтығына экстраполяцияланған болса, KNN ұзақ жұмыс уақытын зиянды факторға айналдыруы мүмкін (3-кесте).

3 – кесте. Модельдердің жылдамдық уақыты

Атауы	Logistic Regression	Linear	Polynomial	Naive Bayes	RBF	KNN
1.Қызыл топырақ (Краснозём)	0.99	0.89	0.99	0.98	0.99	0.98
2.Мақта дақылдары (Хлопковая культура)	0.94	0.91	0.97	0.93	0.97	0.97
3.Сұр топырақ (Серая Почва)	0.90	0.88	0.91	0.73	0.91	0.91
4.Дымқыл топырақ (Влажная Серая Почва)	0.36	0.44	0.52	0.08	0.52	0.70
5.Орылған егістік (Почва с растительной стерней)	0.80	0.52	0.86	0.73	0.86	0.89
6.Сулы топырақ (Очень Влажная Серая Почва)	0.82	0.76	0.84	0.72	0.84	0.87

Әдістер полиномиялық SVM, RBF SVM және KNN болды. Сонымен қатар, KNN моделі көп жағдайда дәл болды. Атап айтқанда, KNN ылғалды сұр топырақты басқа белгілерге қарағанда жоғары пайызбен жіктеуге қабілетті болды. Осылайша, дәлдік тұрғысынан K-NN моделі өзінің жалпы жоғары дәлдігімен ғана емес, сонымен қатар сұр топырақтың түрлерін дұрыс анықтау қабілетімен де оңтайлы болып табылады. Бұл тақырып келесі ары қарай толығырақ түсіндіріледі, бірақ айта кететін жайт, KNN басқа модельдерге қарағанда 3% дәлірек болса да, сұр топырақты жіктеудің едәуір жақсаруы KNN-нің жоғары дәлдігін растайды. Егер модельдердің нәтижелері нақты ортада қарастырылса, онда оңтайлы модельді таңдағанда жұмыс уақытын ескеру маңызды болады. KNN кез-келген модельге қарағанда дәлірек болғанымен, оның ең баяу жұмыс уақыты — 11,73 секунд. Жіктеу алгоритмдері ноутбукта жұмыс істегенімен, уақыт мәселесі жақсы компьютерлерде әлі де қиындықтар тудыруы мүмкін. Осылайша, жалпы оңтайлы модель RBF SVM әдісі ең тиімді болуы мүмкін. Мысалы, нақты әлем модельдерінде көптеген кескіндердің тұтастығын бұзатын әлдеқайда үлкен мәліметтер жиынтықтары болады; сондықтан жалпы жіктеу мақсатында RBF SVM топырақ түрлеріне дәл жіктеу үшін пайдалы болуы мүмкін. Алайда, пайдаланушы аймақтың ең дәл талдауын алғысы келетін жағдайда, KNN жақсы жұмыс істейді, өйткені бұл жағдайда жұмыс уақыты міндетті түрде маңызды рөл атқармайды.

Жоғарыда келтірілген талдаудан біз SVM (RBF ядросы) және K-NN алгоритмдері ең тиімді екені анықталды.



Сур.4. Топырақты SVM және KNN модельдері арқылы анықтаудың қате және дұрыс тұжырымдары
(Fig.4. incorrect and correct conclusions for determining the soil using SVM and KNN models)

Графикте (сурет-4) осы жіктеу алгоритмдерінің көмегімен алынған нәтижелерді байқауға болады. Талдау барысында көптеген модельдер сұр топырақтың үш санатын анықтауда қиындық тудыратыны байқалды. Қарапайым "сұр топырақ" санаты үшін екі модель де болжамдарының дәлдігі бойынша ұқсас жұмыс жасады. SVM әдісі "ылғалды сұр топырақ" класы үшін беретін болжамдарының көпшілігінде қателік кездеседі, бірақ "өте ылғалды сұр топырақ" классификациясында KNN-нен асып түседі. KNN моделі үшін де болжамдар қателігі болды. SVM моделімен салыстырғанда KNN моделі "ылғалды сұр топырақ" классификациясында қателік деңгейі аздау екендігін көрсетті. Алынған нәтижелерді графиктен байқауға болады. KNN 3 % дәлірек болса да, сұр топырақты жіктеудің едәуір жақсаруы, KNN-нің жоғары дәлдігін растайды. KNN кез-келген модельге қарағанда дәлірек болғанымен, оның ең тез жұмыс уақыты 11,73 секундты құрады.

Талқылау

Зерттеу барысында RBF SVM әдісі және KNN модельдері пиксель деректерін жіктеудің ең жақсы әдістері екендігі анықталды. Қандай әдіс қолданылатындығын әр жағдайға байланысты таңдалуы керек. Жалпы алғанда, RBF SVM сапасы мен жылдамдығы бойынша ең жақсы модельдер болды. Егер модельдерді пайдаланушы жердің үлкен ауданын сканерлеп, топырақ түрін анықтап ажыратуда қолданса, онда RBF SVM моделі үлкен аумақтар үшін тиімді болады. Ал топологиялық кеңістіктің макетін алуда KNN баяу жылдамдығына қарамастан, оңтайлы әдіс ретінде қолданылуы мүмкін.

Нәтиже шешімдері нақтырақ болу үшін, барлық модельдерге көбірек оқыту деректерін енгізу қажет. Бұл оқыту деректері топырақтың ұқсас түрлерін саралауға ғана емес, сонымен қатар жүйенің дәлдігін арттыруға көмектеседі. Егер "дымқыл сұр" және "өте ылғалды сұр" топырақтар мәліметтері біріктірілсе, онда жіктеу модельдері үшін қате болжамдар аз болар еді. Әрине бұл болжамдар енгізілген деректер параметрлері мен

мақсаттарына тікелей байланысты болдаы, сондықтан кескін деректері бойынша жіктеуді жақсарту үшін желіні қайта құрып оқыту қажет. Орындау уақыты көптеген шарттарға байланысты өзгеріп отыратынына қарамастан, бұл мән модельдің жалпы жұмысының көрсеткіші деп есептеуге болады. RBF SVM классификациясының нәтижелері топырақ типтерінің дәл жуықтамасын алуға көмегін тигізеді деп айтуға болады. Ал егер пайдаланушы аймақты дәл талдағысы келсе, KNN қолданған жөн, өйткені жұмыс уақытын маңызды санамай нақты дәлдікке жақындауға мүмкіндік береді.

Қорытынды

Жұмыстың мақсаты көрсетілген топологиялық мәліметтердің осы жиынтығын талдау және берілген пиксельді ұсынылған топырақтың ықтимал түрімен жіктеу болды. Сонымен қатар, бұл жұмыста деректерді жіктеудің дәл және тиімді болатын оңтайлы әдіс тәсілдерін іздеу зерттелген. Қатысқан әрбір модельдің орындалу сапасын бағалау үшін бірқатар аналитикалық әдістер жасалды;

1. Пиксельді жіктеудің ең оңтайлы әдістері SVM (Radial Basis Function Kernel) және KNN модельдері болды

2. Қателер негізінен сұр топырақ класының ұқсастығына байланысты пайда болды, сондықтан осы үш топтың бөлінуін қайта қарастыру керек.

Яғни осы тақырып бойынша қойылған есептер шығарылып нәтижесінде пикселдің ең оңтайлы әдісі анықталды және сұр топырақ класын қарастыру барысындағы туындалған қателіктер өзге жолдарды немесе басқа деңгейлерді таңдауға мүмкіндік берді.

ӘДЕБИЕТТЕР

Қазақстан Республикасының Жер кодексі 2003 жылғы 20 маусымдағы N 442 Кодексі // 01.01.2023 ж өзгертілген

Bhatnagar S., Gill L., Ghosh B., 2020 — *Bhatnagar S., Gill L., Ghosh B.* Drone image segmentation using machine and deep learning for mapping raised bog vegetation communities. *Remote Sensing*, 12(16). P. 2602, <https://doi.org/10.3390/rs12162602>

Crommelinck S., Koeva M., Yang M.Y., Vosselman G., 2019 — *Crommelinck S., Koeva M., Yang M.Y., Vosselman G.* Application of deep learning for delineation of visible cadastral boundaries from remote sensing imagery. *Remote sensing*, 11(21). P. 2505, <https://doi.org/10.3390/rs11212505>

Gorelick N., Hancher M., Dixon M., Ilyushchenko S., Thau D., Moore R., 2017 — *Gorelick N., Hancher M., Dixon M., Ilyushchenko S., Thau D., Moore R.* Google Earth Engine: Planetary - scale geospatial analysis for everyone. *Remote sensing of Environment*, 202. Pp 18–27, <https://doi.org/10.1016/j.rse.2017.06.031>

Khryashchev V., Larionov R., 2020 — *Khryashchev V., Larionov R.* Wildfire segmentation on satellite images using deep learning. In *2020 Moscow Workshop on Electronic and Networking Technologies (MWENT) IEEE*. Pp. 1–5, <https://doi.org/10.1109/MWENT47943.2020.9067475>.

Ma Z., Liu C., Xue H., Li J., Fang X., Zhou J., 2021 — *Ma Z., Liu C., Xue H., Li J., Fang X., Zhou J.* Identification of Winter Wheat by Integrating Active and Passive Remote Sensing Data Based on Google Earth Engine Platform. *Trans. Chin. Soc. Agric. Mach.*, 52. Pp.195–205.

Pan S., Zhao X., Yue Y., 2019 — *Pan S., Zhao X., Yue Y.* Spatiotemporal changes of NDVI and correlation with meteorological factors in northern china from 1985-2015. *E3S Web of Conferences*. – EDP Sciences, 131, <https://doi.org/10.1051/e3sconf/201913101040>

Pan Z., Xu J., Guo Y., Hu Y., Wang G., 2020 — *Pan Z., Xu J., Guo Y., Hu Y., Wang G.* Deep

learning segmentation and classification for urban village using a worldview satellite image based on U-Net. *Remote Sensing*, 12(10), P. 1574, <https://doi.org/10.3390/rs12101574>

Potapov V.P., Oparin V.N., Mikov L.S., Popov S.E., 2022 — *Potapov V.P., Oparin V.N., Mikov L.S., Popov S.E.* Information Technologies in Problems of Nonlinear Geomechanics. Part I: Earth Remote Sensing Data and Lineament Analysis of Deformation Wave Processes. *Journal of Mining Science*, 58(3). Pp. 486–502, <https://doi.org/10.1134/S1062739122030164>

Ren H., Zhou G., Zhang F., 2018 — *Ren H., Zhou G., Zhang F.* Using negative soil adjustment factor in soil-adjusted vegetation index (SAVI) for aboveground living biomass estimation in arid grasslands. *Remote Sensing of Environment*, 209. Pp. 439–445, <https://doi.org/10.1016/j.rse.2018.02.068>.

Schlosser A.D., Szabó G., Bertalan L., Varga Z., Enyedi P., Szabó S., 2020 — *Schlosser A.D., Szabó G., Bertalan L., Varga Z., Enyedi P., Szabó S.* Building extraction using orthophotos and dense point cloud derived from visual band aerial imagery based on machine learning and segmentation. *Remote Sensing*, 12(15). P. 2397, <https://doi.org/10.3390/rs12152397>

Shoko C., Mutanga O., Dube T., 2016 — *Shoko C., Mutanga O., Dube T.* Progress in the remote sensing of C3 and C4 grass species aboveground biomass over time and space. *ISPRS Journal of Photogrammetry and Remote Sensing*, 120. Pp. 13–24, <https://doi.org/10.1016/j.isprsjprs.2016.08.001>

Soloviev A.A., Zharkikh J.I., Krasnoperov R.I., Nikolov B.P., Agayan S.M., 2016 — *Soloviev A.A., Zharkikh J.I., Krasnoperov R.I., Nikolov B.P., Agayan S.M.* GIS-oriented solutions for advanced clustering analysis of geoscience data using ArcGIS platform. *Russian Journal of Earth Sciences*, 16(6). Pp. 1–6., doi:10.2205/2016ES000587.

Soni A., Koner R., Villuri V.G.K., 2020 — *Soni A., Koner R., Villuri V.G.K.* M-UNet: modified U-Net segmentation framework with satellite imagery. Proceedings of the Global AI congress. Springer, Singapore. Pp. 47–59.

Spadoni G.L., Cavalli A., Congedo L., Munafò M., 2020 — *Spadoni G.L., Cavalli A., Congedo L., Munafò M.* Analysis of Normalized Difference Vegetation Index (NDVI) multi-temporal series for the production of forest cartography. *Remote Sensing Applications: Society and Environment*, 20. P. 100419, <https://doi.org/10.1016/j.rsase.2020.100419>

Wyniawskij N.S., Napiorkowska M., Petit D., Podder P., Marti P., 2019 — *Wyniawskij N.S., Napiorkowska M., Petit D., Podder P., Marti P.* Forest monitoring in guatemala using satellite imagery and deep learning. In *IGARSS 2019–2019 IEEE International Geoscience and Remote Sensing Symposium* Pp. 6598–6602.

Zou J., Ding J., Welp M., Huang S., Liu B., 2020 — *Zou J., Ding J., Welp M., Huang S., Liu B.* Using MODIS data to analyse the ecosystem water use efficiency spatial-temporal variations across Central Asia from 2000 to 2014. *Environmental research*, 182. P. 108985, <https://doi.org/10.1016/j.envres.2019.108985>

REFERENCES

Bhatnagar S., Gill L., Ghosh B., 2020 — *Bhatnagar S., Gill L., Ghosh B.* Drone image segmentation using machine and deep learning for mapping raised bog vegetation communities. *Remote Sensing*, 12(16). P. 2602, <https://doi.org/10.3390/rs12162602>

Crommelinck S., Koeva M., Yang M.Y., Vosselman G., 2019 — *Crommelinck S., Koeva M., Yang M.Y., Vosselman G.* Application of deep learning for delineation of visible cadastral boundaries from remote sensing imagery. *Remote sensing*, 11(21). P. 2505, <https://doi.org/10.3390/rs11212505>

Gorelick N., Hancher M., Dixon M., Ilyushchenko S., Thau D., Moore R., 2017 — *Gorelick N., Hancher M., Dixon M., Ilyushchenko S., Thau D., Moore R.* Google Earth Engine: Planetary-scale geospatial analysis for everyone. *Remote sensing of Environment*, 202. Pp 18–27, <https://doi.org/10.1016/j.rse.2017.06.031>

Khryashchev V., Larionov R., 2020 — *Khryashchev V., Larionov R.* Wildfire segmentation on satellite images using deep learning. In *2020 Moscow Workshop on Electronic and Networking Technologies (MWENT) IEEE*. Pp. 1–5, <https://doi.org/10.1109/MWENT47943.2020.9067475>.

Land Code of the Republic of Kazakhstan Code of the Republic of Kazakhstan № 442 of June 20, 2003 // amended on 01.01.2023

Ma Z., Liu C., Xue H., Li J., Fang X., Zhou J., 2021 — *Ma Z., Liu C., Xue H., Li J., Fang X., Zhou J.* Identification of Winter Wheat by Integrating Active and Passive Remote Sensing Data Based on Google Earth Engine Platform. *Trans. Chin. Soc. Agric. Mach.*, 52. Pp.195–205.

Pan S., Zhao X., Yue Y., 2019 — *Pan S., Zhao X., Yue Y.* Spatiotemporal changes of NDVI and correlation with meteorological factors in northern china from 1985–2015. E3S Web of Conferences. – EDP Sciences, 131, <https://doi.org/10.1051/e3sconf/201913101040>

Pan Z., Xu J., Guo Y., Hu Y., Wang G., 2020 — *Pan Z., Xu J., Guo Y., Hu Y., Wang G.* Deep learning segmentation and classification for urban village using a worldview satellite image based on U-Net. *Remote Sensing*, 12(10). P. 1574, <https://doi.org/10.3390/rs12101574>

Potapov V.P., Oparin V.N., Mikov L.S., Popov S.E., 2022 — *Potapov V.P., Oparin V.N., Mikov L.S., Popov S.E.* Information Technologies in Problems of Nonlinear Geomechanics. Part I: Earth Remote Sensing Data and Lineament Analysis of Deformation Wave Processes. *Journal of Mining Science*, 58(3). Pp. 486–502, <https://doi.org/10.1134/S1062739122030164>

Ren H., Zhou G., Zhang F., 2018 — *Ren H., Zhou G., Zhang F.* Using negative soil adjustment factor in soil-adjusted vegetation index (SAVI) for aboveground living biomass estimation in arid grasslands. *Remote Sensing of Environment*, 209. Pp. 439–445, <https://doi.org/10.1016/j.rse.2018.02.068>

Schlosser A.D., Szabo G., Bertalan L., Varga Z., Enyedi P., Szabo S., 2020 — *Schlosser A.D., Szabo G., Bertalan L., Varga Z., Enyedi P., Szabo S.* Building extraction using orthophotos and dense point cloud derived from visual band aerial imagery based on machine learning and segmentation. *Remote Sensing*, 12(15). P. 2397, <https://doi.org/10.3390/rs12152397>

Shoko C., Mutanga O., Dube T., 2016 — *Shok, C., Mutanga O., Dube T.* Progress in the remote sensing of C3 and C4 grass species aboveground biomass over time and space. *ISPRS Journal of Photogrammetry and Remote Sensing*, 120. Pp. 13–24, <https://doi.org/10.1016/j.isprsjprs.2016.08.001>

Soloviev A.A., Zharkikh J.I., Krasnoperov R.I., Nikolov B.P., Agayan S.M., 2016 — *Soloviev A.A., Zharkikh J.I., Krasnoperov R.I., Nikolov B.P., Agayan S.M.* GIS-oriented solutions for advanced clustering analysis of geoscience data using ArcGIS platform. *Russian Journal of Earth Sciences*, 16(6). Pp. 1–6., doi:10.2205/2016ES000587

Soni A., Koner R., Villuri V.G.K., 2020 — *Soni A., Koner R., Villuri V.G.K.* M-UNet: modified U-Net segmentation framework with satellite imagery. Proceedings of the Global AI congress. Springer, Singapore. Pp. 47–59.

Spadoni G.L., Cavalli A., Congedo L., Munafo M., 2020 — *Spadoni G.L., Cavalli A., Congedo L., Munafo M.* Analysis of Normalized Difference Vegetation Index (NDVI) multi-temporal series for the production of forest cartography. *Remote Sensing Applications: Society and Environment*, 20. P. 100419, <https://doi.org/10.1016/j.rsase.2020.100419>

Wyniawskij N.S., Napiorkowska M., Petit D., Podder P., Marti P., 2019 — *Wyniawskij N.S., Napiorkowska M., Petit D., Podder P., Marti P.* Forest monitoring in guatemala using satellite imagery and deep learning. In *IGARSS 2019–2019 IEEE International Geoscience and Remote Sensing Symposium*. Pp. 6598–6601.

Zou J., Ding J., Welp M., Huang S., Liu B., 2020 — *Zou J., Ding J., Welp M., Huang S., Liu B.* Using MODIS data to analyse the ecosystem water use efficiency spatial-temporal variations across Central Asia from 2000 to 2014. *Environmental research*, 182. P. 108985, <https://doi.org/10.1016/j.envres.2019.108985>

NEWS OF THE NATIONAL ACADEMY OF SCIENCES OF THE REPUBLIC OF KAZAKHSTAN
PHYSICO-MATHEMATICAL SERIES

ISSN 1991-346X

Volume 2. Number 346 (2023). 193–207

<https://doi.org/10.32014/2023.2518-1726.193>

MFTAP 29.01.45

© **K. Kelesbaev¹, Sh. Ramankulov¹, M. Nurizina^{2*}, A. Pattaev¹,
N. Mussakhan¹, 2023**

¹Khoja Ahmet Yassawi International Kazakh-Turkish University,
Turkestan, Kazakhstan;

²Sarsen Amanzholov East Kazakhstan University, Ust-Kamenogorsk, Kazakhstan.
E-mail: makpal.nurizina@gmail.com

FEATURES OF STEAM PROJECT TRAINING IN THE PREPARATION OF FUTURE SPECIALISTS IN PHYSICS

Kelesbaev K. — PhD student, Khoja Ahmet Yassawi International Kazakh-Turkish University,
Turkestan, Kazakhstan

E-mail: kazhymukan.kelesbayev@ayu.edu.kz. ORCID ID: <https://orcid.org/0009-0008-9605-7138>;

Ramankulov Sh. — Assoc. Prof., PhD. Department of Natural Science, Khoja Ahmet Yassawi
International Kazakh-Turkish University, Turkestan, Kazakhstan

E-mail: sherzod.ramankulov@ayu.edu.kz. ORCID ID: <https://orcid.org/0000-0002-4786-942X>;

Nurizina M. — author for correspondence, PhD student, Higher School of IT and Natural Science.
Sarsen Amanzholov East Kazakhstan University, Ust-Kamenogorsk, Kazakhstan

E-mail: makpal.nurizina@gmail.com. ORCID ID: <https://orcid.org/0000-0001-8319-4928>;

Pattaev A. — PhD student, Khoja Ahmet Yassawi International Kazakh-Turkish University,
Turkestan, Kazakhstan

E-mail: aminjon_26@mail.ru. ORCID ID: <https://orcid.org/0009-0004-8385-2096>;

Mussakhan N. — PhD student, Khoja Ahmet Yassawi International Kazakh-Turkish University,
Turkestan, Kazakhstan

E-mail: nurken.mussakhan@ayu.edu.kz. ORCID ID: <https://orcid.org/0000-0003-4154-0356>.

Abstract. Creativity in education is a set of indicators that allow future specialists to apply their knowledge in the field of natural sciences, technology, engineering and mathematics (STEM) in product development. However, STEM claims from previous studies that the absence of work is associated with the formation of key competencies of future education professionals in relation to creative training, education and content. In addition, the educational programs for the training of future physicists (students) show a small number of disciplines related to the industrial sphere, and problems in its effective implementation. Developed countries are updating the technologies of teaching this discipline at all stages of education, taking into account the significant influence of physics as a discipline in the field of science and technology, the development of new engineering products. The results of the preliminary analysis of scientific papers and educational programs showed that along with developed countries, in the Republic of Kazakhstan, special attention should be paid to updating the scientific literacy of students in physics

education, the development of creative indicators. This article defines the features of project-based STEM learning in the effective implementation of education. The results of the study allow us to look for solutions to urgent world-class problems, such as the impact of the 4th Industrial Revolution on education systems.

Keywords: STEM education, STEAM project, future physics specialist, teaching physics, creativity

© **К. Келесбаев¹, Ш. Раманқулов¹, М. Нуризинова^{2*}, А. Паттаев¹,
Н. Мұсахан¹, 2023**

¹Қожа Ахмет Ясауи атындағы Халықаралық қазақ-түрік университеті
Түркістан, Қазақстан;

²Сәрсен Аманжолов атындағы Шығыс Қазақстан университеті,
Өскемен, Қазақстан.

E-mail: makpal.nurizanova@gmail.com

STEM ЖОБАЛЫҚ ОҚЫТУДЫҢ БОЛАШАҚ ФИЗИКА МАМАНДАРЫН ДАЯРЛАУДАҒЫ ЕРЕКШЕЛІКТЕРІ

Келесбаев К. — Қожа Ахмет Ясауи атындағы Халықаралық қазақ-түрік университетінің докторанты

Email: kazhymukan.kelesbayev@ayu.edu.kz. ORCID ID: <https://orcid.org/0009-0008-9605-7138>;

Раманқулов Ш. — Қожа Ахмет Ясауи атындағы Халықаралық қазақ-түрік университетінің қауымдастырылған профессор, PhD.

E-mail: sherzod.ramankulov@ayu.edu.kz. ORCID ID: <https://orcid.org/0000-0002-4786-942X>;

Нуризинова М. — корреспонденция үшін автор, Сәрсен Аманжолов атындағы Шығыс Қазақстан университеті, Физика және технологиялар кафедрасының докторанты

E-mail: makpal.nurizanova@gmail.com. ORCID ID: <https://orcid.org/0000-0001-8319-4928>;

Паттаев А. — Қожа Ахмет Ясауи атындағы Халықаралық қазақ-түрік университетінің докторанты

E-mail: aminjon_26@mail.ru. ORCID ID: <https://orcid.org/0009-0004-8385-2096>;

Мұсахан Н. — Қожа Ахмет Ясауи атындағы Халықаралық қазақ-түрік университетінің докторанты

E-mail: nurken.mussakhan@ayu.edu.kz. ORCID ID: <https://orcid.org/0000-0003-4154-0356>.

Аннотация. Білім берудегі креативтілік болашақ мамандардың жаратылыстану, технология, инженерия және математика (STEM) саласындағы білімдерін өнім жасауда қолдана алу көрсеткіштерінің жиыны болып табылады. Дегенмен, алдыңғы зерттеулерден STEM білім берудегі болашақ мамандардың креативтілік дайындығы, білім және мазмұн бойынша негізгі құзыреттіліктерін қалыптастыруға байланысты еңбектердің жетіспеушілігі алға тартады. Сонымен қатар, болашақ физика мамандарын (студенттерді) даярлаудың білім беру бағдарламаларында өндірістік саламен байланысты пәндердің аздығы және оны тиімді жүзеге асырудағы проблемалар көрініс алады. Дамыған елдер физиканың пән ретінде, ғылым мен техника саласына, инженерлік бағыттағы жаңа өнімдердің дамуына елеулі ықпалын ескеріп, білім берудің барлық кезеңдерінде осы пәнді оқытудың технологияларын

жаңартып отырады. Ғылыми еңбектерге, білім беру бағдарламаларына жүргізілген алдын ала талдау нәтижелері дамыған елдермен қатар, Қазақстан Республикасында да физика бойынша білім беруде студенттердің ғылыми сауаттылығын жаңартуға, креативтілік көрсеткіштерін дамытуға ерекше назар аударуды қажет ететіндігін көрсетті. Бұл мақалада білім беруді тиімді жүзеге асыруда STEM жобалық оқытудың ерекшеліктері айқындалды. Зерттеу барысында Зерттеу нәтижелері 4-ші өнеркәсіптік революцияның білім беру жүйелеріне тигізетін әсері сияқты әлемдік деңгейдегі өзекті мәселелердің шешімін іздеуге мүмкіндік береді.

Түйін сөздер: STEM білім беру, STEM жоба, болашақ физика маманы, физиканы оқыту, креативтілік

Алғыс. Бұл зерттеуді Қазақстан Республикасы Ғылым және жоғары білім министрлігінің Ғылым комитеті қаржыландырды (ЖТН «AP19579398»).

© К. Келесбаев¹, Ш. Раманкулов¹, М. Нуризинова^{2*}, А. Паттаев¹,
Н. Мұсахан¹, 2023

¹Международный казахско-турецкий университет им. Ахмеда Ясави,
Туркестан, Казахстан;

²Восточно-Казахстанский университет имени Сарсена Аманжолова,
Усть-Каменогорск, Казахстан.

E-mail: makpal.nurizanova@gmail.com

ОСОБЕННОСТИ ПРОЕКТНОГО ОБУЧЕНИЯ STEM В ПОДГОТОВКЕ БУДУЩИХ СПЕЦИАЛИСТОВ ПО ФИЗИКЕ

Келесбаев К. — докторант Международного казахско-турецкого университета имени Ходжи Ахмеда Ясави

E-mail: kazhymukan.kelesbayev@ayu.edu.kz. ORCID ID: <https://orcid.org/0009-0008-9605-7138>;

Раманкулов Ш. — ассоциированный профессор, PhD Международного казахско-турецкого университета имени Ходжи Ахмеда Ясави.

E-mail: sherzod.ramankulov@ayu.edu.kz. ORCID ID: <https://orcid.org/0000-0002-4786-942X>;

Нуризинова М. — автор корреспонденции, докторант кафедры физики и технологий Восточно-Казахстанского университета имени Сарсена Аманжолова

E-mail: makpal.nurizanova@gmail.com. ORCID ID: <https://orcid.org/0000-0001-8319-4928>;

Паттаев А. — докторант Международного казахско-турецкого университета имени Ходжи Ахмеда Ясави,

E-mail: aminjon_26@mail.ru. ORCID ID: <https://orcid.org/0009-0004-8385-2096>;

Мұсахан Н. — докторант Международного казахско-турецкого университета имени Ходжи Ахмеда Ясави

E-mail: nurken.mussakhan@ayu.edu.kz. ORCID ID: <https://orcid.org/0000-0003-4154-0356>.

Аннотация. Креативность в образовании — это набор показателей, позволяющих будущим специалистам применять свои знания в области естественных наук, технологий, инженерии и математики (STEM) при разработке продуктов. Однако из предыдущих исследований STEM утверждает, что

отсутствие труда связано с формированием ключевых компетенций будущих специалистов в области образования в отношении творческой подготовки, образования и содержания. Кроме того, в образовательных программах подготовки будущих специалистов-физиков (студентов) проявляется малочисленность дисциплин, связанных с производственной сферой, и проблемы в ее эффективной реализации. Развитые страны актуализируют технологии преподавания данной дисциплины на всех этапах образования, учитывая значительное влияние физики как дисциплины, в области науки и техники, разработки новых продуктов инженерного направления. Результаты предварительного анализа научных трудов, образовательных программ показали, что наряду с развитыми странами, в Республике Казахстан, особое внимание необходимо уделять обновлению научной грамотности студентов в образовании по физике, развитию креативных показателей. В данной статье определены особенности проектного обучения STEM в эффективной реализации образования. Результаты исследования позволяют искать решения актуальных проблем мирового уровня, таких как влияние 4-й промышленной революции на системы образования.

Ключевые слова: STEM-образование, STEM-проект, будущий специалист по физике, преподавание физики, креативность

Кіріспе

Қазіргі таңда Индустрия 4.0 жағдайында цифрлық құралдардың көптеп пайда болуына байланысты заманауи жаратылыстану ғылымдарының әр түрлі салаларындағы физиканың қолданбалылығы қарқынды түрде арта бастады. ХХ-ғасырдың соңында өнеркәсіп орындарындағы техникалық құралдардың автоматтандырылуы және жұмыс орындарының азаюына байланысты адам өзінің қызмет саласын өзгертуге мәжбүр болды. Сондықтан, политехникалық оқыту идеясы қолға алына бастады. Бұл ұғым қазіргі орта және жоғары оқу орындарында бейіндік оқыту ұғымымен кең таралды.

Физиканың қолданбалы ғылым екендігін ескерсек, оның пәні, мақсаты мен міндеттері тәжірибе арқылы тағайындалатынына көз жеткіземіз. Физикалық білім қазіргі таңда шексіз салада қолданылады: ғылым мен техникада, инженерия мен өндірісте және т.б. салаларда. Сондықтан да осы салада жаңа идеялар мен жетістіктерге қол жеткізу тек креативті мамандардың іс-әрекеттерімен жүзеге асады.

Соңғы жылдары жоғары рейтингті журналдарда орын алған мақалаларға метаанализ жасау нәтижесінде физика саласындағы білімді креативтілік ұғымымен байланыстыра отырып оқытудың мәселелерін зерттеген бірнеше ғалымдардың еңбектерін негізге алуға болады.

Сусилавати С., Дойен А., Муляди Л., Або К., К. Пинедра және тағы басқалар Индустрия 4.0 жағдайларын ескеріп, ғылыми процесс дағдыларын және физика саласында болашақ оқытушы студенттердің креативтілігін жақсартуда

PhET виртуалды медиасы арқылы сұрау үлгісін пайдалана отырып, заманауи физика оқыту құралдарын әзірлеудің тиімділігін тексерумен айналысқан. Осы оқыту құралдарын әзірлеу кезінде 4D әзірлеудің анықтау, жобалау, әзірлеу және тарату сияқты зерттеу моделі қолданылған (Susilawati et al., 2022).

Ғылыми еңбектерден әлемдік аренада зерттеу тақырыбы бойынша алынған нәтижелер жалпылама ұсыныстар ретінде қолданылып, нақты жүйелі түрде, белгілі бір физикалық пәндер аясында ұйымдастырылмағаны көрініс табады. Сонымен қатар, отандық ғалымдардың еңбектерінен креативтілік ұғымының көбінесе шығармашылық ұғымы ретінде қолданылатындығын, техникалық пәндерді оқыту салаларында бұл ұғымды қолдану мәселелері бойынша зерттеулердің аздығы алға тартады.

Біздің алдын ала жүргізген зерттеулерінің нәтижесі болашақ физика мамандарының креативтілігін қалыптастырудың қажеттілігі мен оны дамытудың мүмкіндіктерін айқындауға ықпал етеді. Студенттердің креативтілігін қалыптастырудың мәселелерін тек физиканың белгілі бір тарауларын оқыту мысалында зерттегенімен, бұл зерттеулер аталған сала бағыты бойынша жүйелі ізденістерді жасауға мүмкіндік туғызары анық.

Жоғарыда келтірілген еңбектер мен тұжырымдамалардан креативтіліктің физика саласында білімгерлердің бойында болуы қажет болған қасиеттердің бірі екендігіне көз жеткізуге болады. Алайда, болашақ физика мамандарының креативтіліктерін нақты бір пән аясында жүйелі түрде дамытудың технологиялары, оны тиімді іске асырудың механизмдері айқындалмаған. Білімнің қазіргі бұл жеткіліксіздігін шешудің, яғни, студенттердің физика саласындағы білімдерін индустрия 4.0 жағдайында қалыптастырудың және олардың креативтілік көрсеткіштерін дамытудың озық технологиясы ретінде STEM білім беруді қолданамыз.

2019 жылдың 27 желтоқсанындағы Қазақстан Республикасы Үкіметінің №988 «Білім және ғылымды дамытудың 2020–2025 жылдарға арналған мемлекеттік бағдарламасын» бекіту туралы қаулыда индустрия 4.0 талаптарына қол жеткізудің бір жолы ретінде — химия, биология, физика пәндерін STEM-кабинеттермен жабдықтау негізінде оқыту мәселелері қарастырылған (Memleketтік бағдарламасы, 2017).

Жоғары оқу орындарында физиканы оқыту үдерісінде STEM тәсілі білім алушылардың белсендірек болуы және ғылым мен технологияны қолдануға негізделген идеяны іске асыру құралы ретінде қызмет етеді. Болашақ физика мамандарын STEM негізіндегі оқытуды іске асырудың бірнеше артықшылықтар бар: ақпараттық даму қоғамына дайын тұратын ұлттың болашағын дайындау мүмкіндігі, қолданыстағы жаңа инновацияларды дамытуға үлес қосу, білімгерлердің STEM мамандығына деген қызығушылығын ояту, білімді күнделікті өмірмен байланыстыру, өзін-өзі бағалаудың тиімділігін арттыру, білімгерлердің STEM сауаттылығын дамыту және т.б. Осы артықшылықтарды және физиканың ғылым саласы ретінде

кеңінен қолданылатындығын ескеріп, жоғары оқу орнында STEM негізінде физиканың қолданбалы аспектілерін креативтілік арқылы жүзеге асырудың маңызды екендігін түсінеміз (Ramankulov et al., 2022).

STEM негізінде оқытудың негізгі принциптерінің бірі — болашақ физика мамандарына өндірістік тәжірибені және әдістемелік тұрғыдан күрделі зерттеу әдістерін меңгеруге үйрету. Зерттеушілер STEM тәсілін игеруден алынған білім мен тәжірибе кез келген басқа тәжірибелерден маңыздырақ екенін көрсетті (Badeleh, 2021). Екінші маңызды мәселе — өнеркәсіптік бизнесті дамыту, яғни өндірістік міндеттерге жауап ретінде жаңа идеяларға қол жеткізу.

Алайда, индустрия 4.0 талаптарын орындауда болашақ физика мамандарының креативтілік көрсеткіштерін дамытуға STEM технологияның әлеуетін бағалау әлемдік деңгейдегі өзекті мәселе ретінде қалып отыр.

Болашақ физика мамандарының креативтілігін, креативтілік сипаттамаларын зерттеген көптеген авторлардың ойларына, креативтілік кәсіби іс-әрекеттерде және физикадан алған білімдерін жаңа нәтижелерге қол жеткізуде көрінетін тұлғаның қасиеттерінің жиынымен сипатталады (Shute et al., 2021; Kazbekova et al., 2022).

Дегенмен, жоба тақырыбына сәйкес болашақ физика мамандарының креативтілігін дамытуда STEM технологиясының мүмкіндіктерін қолданудың әдістемелік жүйесі, іске асыру механизмдері бойынша зерттеулер жеткіліксіз болып табылады. Нақтырақ, физика білім беру бағдарламасындағы пәндерді оқытуда студенттердің креативтілігін дамыту саласындағы дидактикалық модельдердің жасалынбағандығы, STEM технологиясының «креативтілік» ұғымымен ұштаспауы, білім беру бағдарламаларында индустрия 4.0 жолындағы тұжырымдарды қамтитын және сол арқылы болашақ физика мамандарын өнеркәсіптік қызметтерде креативті тұлға ретінде қалыптасуына ықпал ететін пәндердің жоқтығы, пәндер бойынша оқу-әдістемелік кешен және оқу материалдарының жеткіліксіздігі алға тартады.

Жоғарыда аталған қажеттіліктерді орындаудың алғашқы қадамы STEM білім берудің болашақ физика мамандарын даярлаудағы оқу пәндерін тиімді іске асыруға әсерін анықтау болып табылады. Сондықтан біздің зерттеуіміздің алғашқы қадамы ғылыми әдебиеттерге жүйелі талдау жасау арқылы STEM оқытудың ерекшеліктерін айқындау болды.

Физикадан білім беруде STEM технологиясын қолданып студенттердің креативтілігін дамытуға негізделген STEM жобаларды әзірлеу және оны оқу үдерісіне енгізудің қажеттілігі мен нақты бір пән бойынша STEM өнімдерді жасау арқылы студенттердің даярлығын жетілдіруді тиімді іске асыру бойынша ізденістердің аздығы арасында қарама-қайшылықтар орын алады. Осындай қарама-қайшылықтардың шешімін айқындау әлемдік және отандық ғылым саласына үлкен серпіліс береді. Егер, білім беру мекемелерінде STEM технология негізінде болашақ физика мамандарының креативтілігін дамытуға

оқытудың әдістемесі теориялық негізделсе, креативтілік көрсеткіштерді дамытуда STEM технологияларды пайдалану ғылыми-әдістемелік тұрғыдан қамтамасыз етілсе, онда STEM білім берудің нақты физикалық пәндерде қолданудың ерекшеліктері айқындалады және бұл болашақта оқу үдерісінде студенттердің STEM өнімдерді жасау арқылы индустрия 4.0 талаптарын орынауға даярлығын жетілдіруге мүмкіндік береді.

Зерттеу әдістері

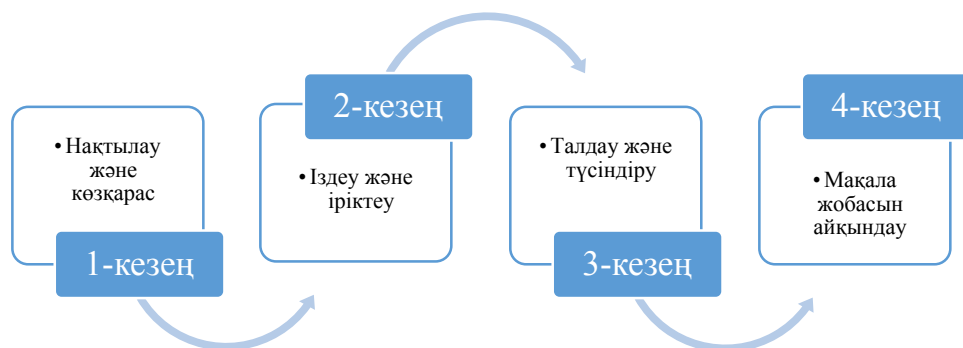
Зерттеу барысында ғылыми еңбектерді іріктеу нақты анықталған критерийлер негізінде жүйелі шолу жасау арқылы жүргізілді. Ғылыми мақалаларды іріктеу процесі «PRISMA» декларациясына сәйкес әзірленді. Сонымен қатар, әдебиеттерді іздеуге және топтастыруға ағылшын, қазақ, орыс тілдерінде жазылған журнал мақалаларды, «STEM education», «STEM project», «physics teaching», «creativity» терминдері тақырыпта, аннотацияда немесе кілт сөздерде көрсетілуін, креативтілікті бағалау әдістерінің көрсетілуін, STEM білімінің креативтілікке әсері туралы қорытындыларды ұсынуын критерийлер ретінде таңдап алынды. Әдебиеттерді жинақтау 2023 жылдың бірінші жартысында Web of Science және Elsevier's Scopus дерекқорлары, ҒЖБССҚҚ ұсынған рецензияланатын отандық басылымдар арқылы жүргізілді. Әдебиеттерді жинақтауда 2019–2023 жылдар аралығындағы соңғы бес жылдықтағы басылымдар есепке алынды. Шолу үшін таңдалған журналдардың қысқаша сипаттамасын 1-кестеден көруге болады (кесте -1).

Кесте 1

Журналдардың атауы	Мақала саны	Индекстелу бойынша
International Journal of STEM Education	5	Scopus
EURASIA Journal of Mathematics Science and Technology Education	3	Scopus
Education Sciences	5	Scopus
STEAM-H: Science, Technology, Engineering, Agriculture, Mathematics and Health	5	Scopus
Journal of Nusantara Studies-Jonus	3	WoS
European Journal of STEM Education	6	Scopus
International Journal of STEM Education for Sustainability	2	J-Gate
International Journal of Advanced Research in Engineering and Technology (IJARET)	4	Scopus
International Journal of Educational Science and Research (IJESR)	2	Copernicus

1-кестеге сәйкес, талдау үшін таңдалған мақалалар 9 беделді халықаралық журналдардан тұрады. Олар 6 «Scopus» базасында индекстелген журналдардан, 1 «Web of Science» журналынан (WoS), 2 «Copernicus» та индекстелген журналдардан және 1 «J-Gate» те индекстелген журналдан тұрады. Осылайша, жоғары рейтингті журналдардан таңдалған мақалалар сапалы деп есептеуге болады.

Таңдап алынған 35 мақалалар бойынша зерттеулер физикадан білім берудегі STEM-жобалық оқытудың жалпы сипаттамалары мен тиімділігіне негізделген мәселелерді талдауға бағытталады. Деректер жиналып, кестелер мен сызбалар түрінде нақтыланып, нәтижесінде сипаттамалық түрде талданды. Деректерді талдау нәтижелері сандық және пайыздық көрсеткіштермен беріледі. Әдебиетке шолу зерттеуінде 1-суретке сәйкес бірнеше кезеңдер қолданылды.



1-сурет. Мақалаларды талдау кезеңдері.

Зерттеу үдерісінің кезеңдері келесідей:

Нақтылау және көзқарас кезеңі STEM-ді физика біліміне біріктіретін жобалық оқытудың тиімділігіне қатысты мақалаларға шолу жасаудың негіздемесін әзірлеуге, зерттеу сұрақтарын, мақала критерийлерін анықтауға және зерттеу құрылымын қалыптастыруға бағытталды.

Іздеу және іріктеу кезеңдерінде физика бойынша білім берудегі жобаларға негізделген STEM оқыту тәсілінің тиімділігін түсіндіретін мақалалар жинақталды. Журналдарды іздеу «Google Scholar» арқылы жүргізілді. Рецензияланған мақалалардың сапасын қамтамасыз ету үшін авторлар мақалаларды тек беделді халықаралық журналдардан таңдаумен шектелді. Негізгі кілт сөздер ретінде «STEM жобасы», «STEM жобаларына негізделген оқыту», «физиканы оқыту» қолданылды. Нәтижесінде авторлар физикалық білім беруде STEM-PJBL (STEM – жобаға бағытталған оқыту)-ге қатысты жүзден аса мақаланы тапты. Алайда авторлар шолу үшін тек 30 мақаланы таңдап алды. Талдау және түсіндіру кезеңінде авторлар зерттеу сипаттамаларының көрінісін талдады. Содан кейін деректерді талдау нәтижелері кестелер мен сызбалар арқылы сипатталды. Соңғы кезеңде деректерді талдау нәтижелері мақала жобасы түрінде әзірленді. Бұл зерттеулердің нәтижелері зерттеу тақырыбына сәйкес білім беруді тиімді жүзеге асыруда STEM жобалық оқытудың ерекшеліктері айқындауға бағдарланды.

Зерттеу нәтижелері және талқылау

Бұл зерттеу жарияланым түрін, жарияланған жылын, зерттеу тәсілін

және білім деңгейін қоса алғанда, оның жалпы сипаттамаларына сәйкес зерттеуді ұсынады. Алынған мәліметтерге сүйене отырып, жүйелі талдау үшін таңдалған мақалалар 100% халықаралық журналдардан алынған деген қорытынды жасауға болады. Бұл зерттеуде физиканы зерттеудегі STEM-PJBL тиімділігі автор, мақала атауы және мақала нәтижелері бойынша бөлінеді. Физиканы оқытудағы STEM-PJBL тиімділігі бойынша кейбір негізгі әдебиеттер бойынша талдау нәтижелері 2-кестеде келтірілген.

Кесте 2

Мақаланың атауы (қазақ тілінде)	Алынған нәтижелер	Авторлар	Баспа атауы мен жарияланым мерзімі
1	2	3	4
Физикада STEM-PBL интеграцияланған модулінің студенттердің қызығушылығы, ойлауы және күш-жігері тұрғысынан тиімділігі	STEM – PBL интеграцияланған физикалық модулі студенттердің жеке қызығушылығын, сондай-ақ интервенциядан кейінгі ойлау мен күш-жігерді айтарлықтай арттырғанын көрсетті.	Ф. Сулейман, Дж. Розалес, Л. Кен (Sulaiman, et al., 2023).	Journal of Baltic Science Education, 2023
Физика саласындағы ғылыми сауаттылықты арттыру үшін STEM негізіндегі басқарылатын сұраныстарды оқытудың тиімділігі	STEM негізіндегі басқарылатын сұраныстар арқылы оқыту моделін қолданудың тиімділігі студенттердің ғылыми сауаттылығының жоғары көрсеткіштерімен дәлелденеді.	В. Сапутро, Е. Васис, Т. Прастово (Saputro Wasis & Prastowo, 2023).	Studies in Learning and Teaching, 2023
Бакалаврларға арналған бейінді физиканы STEM біліміне біріктіру	Мақалада STEM курстың мазмұнын, соның ішінде ұқсас курсты оқытуға қызығушылық танытқан басқа оқытушыларға жол ашуға көмектесетін өте танымал практикалық сабақтар сипатталады.	Б. Скалеттер, Дж. Абни (Scalettar & Abney, 2023).	Biophysical Journal, 2023
Орта мектеп оқушыларының STEM-дегі мансапқа деген қызығушылығы және оның Қазақстандағы жынысымен, бағаларымен және отбасы мөлшерімен байланысы	Зерттеу нәтижелері студенттердің STEM мансабына деген қызығушылығын дамыту ішкі факторлар мен сыртқы факторлар арасындағы күрделі өзара әрекеттесуді қамтитын эпигенетикалық құбылыс екенін көрсетеді.	Н. Балта, Н. Жапашов, А. Мансурова, К. Цафилкоу, А. Оливейра, В. Лагрон (Balta et al., 2023).	Science Education, 2023
STEAM PJBL моделі арқылы негізінде физика мұғалімдерінің жоғары деңгейлі ойлау қабілеттерін арттыру	STEM-PJBL моделін қолдана отырып оқыту дәстүрлі оқытуға қарағанда жоғары деңгейлі ойлау дағдыларын жақсартуға алады.	Р. Харьяди, Х. Пуджиастуч (Haryadi & Pujiastuti, 2022).	International Journal of STEM Education for Sustainability, 2022

21 ғасырдағы студенттердің дағдыларын жақсарту үшін интеграцияланған физика бойынша электрондық STEM модульдерін әзірлеу	Зерттеу нәтижелері 21 ғасырдағы студенттердің дағдыларын жақсарту үшін STEM integrated physics электронды модульдерін әзірлеуді білдіреді, олар іс жүзінде оқуда қолдануға болатын етіп жасалған.	Н. Назифа, А. Асризал (Nazifah & Asrizal, 2022).	Jurnal Penelitian Pendidikan, 2022
STEM негізіндегі физиканы оқытудың студенттердің сыни ойлау дағдыларына әсерін бағалау: әдебиетке жүйелі шолу.	Оптикалық аспаптар, термодинамика, температура және жылу шығару, динамикалық сұйықтықтар және дыбыс толқындарын қоса алғанда, сыни ойлау дағдыларын жақсарту үшін STEM әдісін қолдана отырып физика тұжырымдамаларын үйретеді.	Д. Сулисворо, Н. Калиаппан, А. Винарти, Д. Сулисворо, Н. Калиаппен (Sulisworo et al., 2021).	Indonesian Review of Physics, 2021
Жобаға негізделген STEM қосымшаларының (PjBL-STEM) 21 ғасырдағы дағдыларды дамытуға әсері	Қосымшалар студенттердің 21 ғасырдағы көптеген дағдыларына оң әсер етті. Мысалы, қарым-қатынас және ынтымақтастық, мәселелерді шешу, шығармашылық, сыни ойлау, жауапкершілік, қоршаған орта туралы хабардар болу және ақпараттық технологиялар сауаттылығы.	М. Баран, М. Баран, Ф. Каракоюн, А. Маскан (Baran et al., 2021).	Journal of Turkish Science Education, 2021
Жаратылыстану ғылымдарын зерттеудегі гендерлік айырмашылықтарға негізделген студенттердің STEM ' сауаттылығын талдау	STEM сауаттылығы гендерлік айырмашылықтарға байланысты әр түрлі бағаланады. Ұлдар қыздарға қарағанда жаратылыстану ғылымдарын, әсіресе негізгі физикадағы зертханалық жұмыстарды зерттеу кезінде STEM сауаттылығын жақсы біледі.	Нуррамадхани (Nurramadhani, 2020).	JHSS (Journal of humanities and social studies), 2020
STEM негізіндегі оқыту кезінде студенттің тұрақты ток электр мәселелерін шешу қабілетін талдау	STEM-оқытуда проблемалық және жобалық оқытудың интеграциялық моделін қолдану арқылы студенттердің проблемаларды шешу қабілеті артты.	Р. Априяни, Т. Рамалис, И. Суварма (Apriyani et al., 2019).	Journal of Science Learning, 2019
STEM жобаларына негізделген оқыту арқылы студенттердің шығармашылығын арттыру	STEM жобаларына негізделген оқыту студенттердің шығармашылығына жақсы әсер етеді.	С. Ханиф, А. Виджая, Н. Винарно (Hanif et al., 2019).	Journal of Science Learning, 2019

2-кестеден PjBL-STEM пайда болатын білім беру мәселелерін шешу үшін, әсіресе физиканы оқыту кезінде қолданылатынын көруге болады. Қарастырылған 36 зерттеу мақаласының нәтижелері STEM PjBL таным

аспектілеріне, дағдыларға және физикалық көзқарасқа оң әсер ететін ғылымды, технологияны, инженерияны және математиканы пайдалана отырып, оқытудың ең жаңа әдістері ретінде пайдаланылуы мүмкін деген қорытындыға келеді.

Қарастырылған 36 мақаланың барлығында сапалы, сандық және аралас әдістер сияқты әртүрлі зерттеу әдістері қолданылған. Әдебиеттерде ең көп қолданылатын зерттеу тәсілі — сандық әдістер болып табылады. Ал аз қолданылатын зерттеу тәсілі — сапалық әдістер. Көптеген ғалымдардың еңбектерінен зерттеу тәсілдері ретінде аралас әдістерді қолданғанын байқадық. Зерттеушілер деректерді жинау үшін тесттер, сұхбаттар, сауалнамалар, бақылау, аудио/бейне жазу құралдарын және т.б. қолданған. Талдау нәтижелері физиканы оқытуда STEM-PJBL көмегімен зерттеулерді енгізу университет деңгейінде әлі де өте шектеулі екенін көрсетеді. Бұл зерттеу сонымен қатар физиканы зерттеудегі STEM-PJBL тиімділігін талдайды. Алдыңғы әдебиеттерге шолу тек STEM білімінде жобаға бағытталған оқыту модельдерін (PJBL) енгізудің тиімділігін қарастырды. Бұл зерттеу алдыңғы зерттеулерге қарағанда нақтырақ қамтумен талданады, атап айтқанда физиканы зерттеудегі STEM-PJBL тиімділігі зерттеледі. Бұл зерттеу автордың сипаттамаларына, мақала атауына және мақала нәтижелеріне негізделген STEM-PJBL тиімділігін, ерекшеліктерін түсіндіреді.

Ғылыми әдебиеттерді талдай келе, болашақ физика мамандарын STEM-ді жобалау негізінде даярлауда келесі аспектілерді қолданудың қажеттігін есепке алдық:

Білім беру бағдарламасына сәйкес негізгі білімді меңгерудің мақсатты деңгейін нақтылау.

Физикалық құбылыстанмен байланысты тақырыптар бойынша STEM-ге негізделген сабақ жоспарын әзірлеу кезінде мұғалімдер алдымен курстың мақсатын және оқушылардың оны меңгеру дәрежесін нақты анықтауы керек. Оқу мақсаттарын игеру үшін мұғалімдер оқу бағдарламасын құруда әртүрлі қиындықтар туғызу міндетін қарастыруы керек. Яғни, жеңіл меңгерілетін оқу мақсаттарына аз сағат бөліп, ал тереңірек меңгерілетін оқу мақсаттарындағы негізгі ойларды бөлектеу үшін оқу бағдарламасында бірнеше рет қойылуы керек.

Оқыту нәтижелерінің мақсаттарына жету үшін жеңіл меңгерілетін STEM жобаларды әзірлеу немесе өмірмен байланысты кейстерді құру.

STEM сабақтарының дәстүрлі оқытудан айырмашылығы, негізінен жобалық оқыту түрінде жүзеге асырылады. STEM оқу үдерісінде жобаларға негізделген, бірақ негізінен білім алушылар жобаны бір-екі аптада аяқтауы немесе практикалық мәселені шешуі үшін жеңіл меңгерілетін микро жобаларды әзірлеу үшін оның қолжетімді құралдар мен нысандарды пайдаланады.

3. STEM білім беруге негізделген бағалау механизмін құру.

STEM сабақтарды жоспарлау кезінде қорытынды бағалау механизмін құруға назар аударылуы керек. Оқу тақырыбының мазмұнына сәйкес, білімгер жобаны жүзеге асыру туралы есеп беруі, өндірістік нәтижелерді ұсына алуы, тіпті сәтсіздіктердің қысқаша мазмұнын жаза білуі керек.

STEM білім берудің мәнін терең зерттеу идеясы проблемалық және жобалық оқытуды қарастыруға байланысты кумулятивті әсерге негізделеді. Осылайша, болашақ физика мамандарының алдыңғы бөлімдерді зерделеу кезінде жинаған білімдері, дағдылары мен тәжірибелері білім беру бағдарламаларына сәйкес пәндерді оқу кезінде меңгерген білімдермен үйлесіп, олардың оқуға деген ынтасын күшейтеді. Болашақ физика мамандарының жинаған бұл білімдері, дағдылары мен тәжірибелері жай ғана қолданылып қана қоймай, оқу іс – әрекетінің сапалы жаңа нәтижесін – STEM өнімді әзірлеуге үйретеді. Бұл үдеріс STEM білім берудің формалары мен әдістерін зерттеу арқылы кезең-кезеңімен іске асырылады.

STEM білім берудің физиканы оқытудағы ерекшеліктерін айқындауға негіз болған STEM білім берудің формалары мен әдістеріне тоқталайық.

STEM case – пәнаралық сипаттағы нақты экономикалық, экологиялық және әлеуметтік мәселелерге негізделген нақты жағдаят.

STEM-ойын – пәнаралық тәсіл негізінде білімгерлердің оқу-танымдық іс-әрекетін эмоционалды саланы дамыту арқылы ұйымдастырудың құралы.

STEM-жоба – пәнаралық қағидаларға негізделген және белгіленген мәселені шешуге бағытталған өзара байланысты компоненттердің жүйесі.

STEM – зертхана - STEM-білім беруді енгізу мен дамытуға бағытталған эксперименттік ғылыми-зерттеу, оқу-зерттеу жұмыстарын жүргізетін ғылыми мекеме бөлімі.

Физиканың білім беру бағдарламаларын, STEM білім беру ресурстарын, ғылыми әдебиеттерді талдай келе, физиканы оқытудағы STEM білім берудің, оның болашақ физика мамандарын даярлаудағы ерекшеліктерін айқындадық.

Болашақ физика мамандарын даярлаудағы STEM білім берудің ерекшеліктері төмендегідей:

- Физика білім беру бағдарламасының жеке пәндері бойынша емес, оның мазмұнына сәйкес тақырыптар бойынша біріктірілген оқытуды тиімді іске асырудың мүмкіндігі;

- Физикадан алған ғылыми-техникалық білімдерін нақты өмірде қолдану;

- Болашақ физика мамандарының сыни тұрғыдан ойлау және проблемаларды шешу дағдыларын дамыту;

- Болашақ физика мамандарының өз қабілеттеріне деген сенімділіктерін қалыптастыру;

- STEM негізінде белсенді байланыс және топтық жұмысты тиімді іске асыру;

- Болашақ физика мамандарының техникалық пәндерге қызығушылықты дамыту;

- STEM жобаларға креативті және инновациялық тәсілдерді сабақтастырудың мүмкіндігін арттыру;

- Әр білім алушының жас және жеке ерекшеліктерін ескере отырып, олардың іс-әрекеттері арқылы техникалық шығармашылыққа деген ынтаны дамыту;

- Кәсіби бағдар беру;

- Болашақ физика мамандарын өмірдің технологиялық инновацияларына дайындау.

Жалпы қорыта келгенде, аталған бағытта ғылыми зерттеулерді іске асыру барысында алынған нәтижелер 4-ші өнеркәсіптік революцияның білім беру жүйелеріне тигізетін әсері сияқты әлемдік деңгейдегі өзекті мәселелердің шешімін іздеуге мүмкіндік береді.

Қазақстан Республикасындағы жоғары оқу орындарында физикалық және химиялық ғылымдар даярлау бағыты бойынша болашақ физика мамандарын даярлаудың білім беру бағдарламаларына STEM-ге негізделген жаңа пәндер әзірленіп оқу үдерісіне ендірілу болашақтың еншісіндегі өзекті мәселе болып табылады.

Қорытынды

Бұл зерттеу жұмысының нәтижелері болашақ физика мамандарының арасында STEM технологиясы ұғымымен байланысты ғылыми ойлауды дамытуға бағытталған іс-шаралармен байытуға мүмкіндік беретіндігін көрсетті. Біздің алдын ала жүргізген зерттеулеріміз, сонымен қатар, ғылыми әдебиеттердегі зерттеулердің нәтижелеріне сүйене отырып, болашақ физика мамандарын даярлауға STEM жобалық оқытудың әсерін зерттеу кезінде STEM өнімдерін көмекші оқу материалдары ретінде қамтамасыз ету көбірек қызығушылық танытатындығын байқауға болады.

Демек, STEM мазмұнға негізделген пәндерді ұйымдастыра отырып, бұл жалпы пәндердің интеграциясына ықпал етуін, проблемаларды табу және шешу қабілеттерін дамытуын байқауға болады. Зерттеуден алынған мәліметтер жоғары және орта білім беру ұйымдарында қолданысқа ұсынылады және бұл нәтижелер физиканың білім беру саласы бойынша жаңа бағыттағы зерттеулерге негіз болады.

Бұл мақаладағы алынған мәліметтер алдағы уақытта зерттеу тақырыбының шеңберінде зерттелінуді қажет ететін жаңа ғылыми және әдіснамалық проблемаларды сипаттауға мүмкіндік береді: Яғни, креативтілік білім мен STEM білім беру арасындағы алшақтық, индустриясы 4.0 элементтерін білім беру жүйесіне енгізудегі және оны еңбек нарығында дамытудағы кедергілер; білім беру бағдарламаларында өндірістік білімге бағдарланған пәндердің аздығы; болашақ физика мамандарының креативтік көрсеткіштерін өндірістік дамытуға бағдарлаудың қажеттілігі мен оны жүзеге асырудың әдістемедік жүйесінің жасалынбағандығы; болашақ физика мамандарын даярлауда STEM пәндерді енгізу мәселелері және оқытудағы критикалық

ойлау, өмір бойына білім алу мен ақпаратты басқару сияқты іс-әрекеттерді дамыту мәселелерін шешуге мүмкіндік береді.

ӘДЕБИЕТТЕР

Apriyani R., Ramalis T.R. & Suwarma I.R., 2019 – *Apriyani R., Ramalis T.R. & Suwarma I.R.* Analyzing Student's Problem Solving Abilities of Direct Current Electricity in STEM-based Learning. *Journal of Science Learning*, 2(3): 85–91. <https://doi.org/10.17509/jsl.v2i3.17559>. (in Eng.).

Baran M., Karakoyun F. & Maskan A., 2021 – *Baran M., Baran M., Karakoyun F. & Maskan A.* The Influence of Project-Based STEM (PjBL-STEM) Applications on the Development of 21st-Century Skills. *Journal of Turkish Science Education*, 18(4): 798–815. <https://doi.org/10.36681/tused.2021.104>. (in Eng.).

Badeleh A., 2021 – *Badeleh A.* The effects of robotics training on students' creativity and learning in physics. *Education and Information Technologies*, 26(2): 1353–1365. <https://doi.org/10.1007/s10639-019-09972-6> (in Eng.).

Казбекова Г.Н., Исмагулова Ж.С., 2022 – *Казбекова Г.Н., Исмагулова Ж.С.* Инновациялық STEM-білім беру тәсілін қалыптастыру// Ясауи университетінің хабаршысы. 3(125): 200–210. <https://doi.org/10.47526/2022-3/2664-0686.17>. (in Kaz.).

Nazifah N. & Asrizal A., 2022 – *Nazifah N. & Asrizal A.* Development of STEM Integrated Physics E-Modules to Improve 21st Century Skills of Students. *Jurnal Penelitian Pendidikan IPA*, 8(4): 2078–2084. <https://doi.org/10.29303/jppipa.v8i4.1820>. (in Eng.).

Nurramadhani A., 2020 – *Nurramadhani A.* Analysis of Students' STEM Literasy Based on Gender Differences in Science Learning. *JHSS (Journal of humanities and social studies)*, 4(1): 21–25. <https://doi.org/10.33751/jhss.v4i1.1903>. (in Eng.).

Ramankulov Sh., Choruh A., Polatuly S., 2022 – STEAM technology as a tool for developing creativity of students: on the example of a school physics course// Ясауи университетінің хабаршысы. –4(126): 200–211. <https://doi.org/10.47526/2022-4/2664-0686.17>. (in Eng.).

Susilawati S., Doyan A., Mulyadi L., Abo C.P. & Pineda C.I.S., 2022 – *Susilawati S., Doyan A., Mulyadi L., Abo C.P. & Pineda C.I.S.* The Effectiveness of Modern Physics Learning Tools Using the PhET Virtual Media Assisted Inquiry Model in Improving Cognitive Learning Outcomes, Science Process Skills, and Scientific Creativity of Prospective Teacher Students. *Jurnal Penelitian Pendidikan IPA*, 8(1): 291–295. <https://doi.org/10.29303/jppipa.v8i1.1304>. (in Eng.).

Shute V.J. & Rahimi S., 2021 – Shute V.J. & Rahimi S. Stealth assessment of creativity in a physics video game. *Computers in Human Behavior*, 116. <https://doi.org/10.1016/j.chb.2020.106647>. (in Eng.).

Sulaiman F., Rosales J.J. & Kyung L.J., 2023 – *Sulaiman F., Rosales J.J. & Kyung L.J.* The effectiveness of the integrated stem-pbl physics module on students' interest, sensemaking and effort. *Journal of Baltic Science Education*, 22(1): 113–129. <https://doi.org/10.33225/jbse/23.22.113>. (in Eng.).

Saputro V.C.E., Wasis & Prastowo T., 2023 – *Saputro V.C.E., Wasis & Prastowo T.* The Effectiveness of STEM-Based Guided Inquiry Learning to Train Science Literacy of Physics. *Studies in Learning and Teaching*, 3(3): 141–148. <https://doi.org/10.46627/silet.v3i3.179>. (in Eng.).

Sulisworo D., Kaliappen N., History A., Winarti W., Sulisworo D. & Kaliappen N., 2021 – *Sulisworo D., Kaliappen N., History A., Winarti W., Sulisworo D. & Kaliappen N.* Evaluation of STEM-based physics learning on students' critical thinking skills: a systematic literature review. *Indonesian Review of Physics*, 4(2), 23–31. Retrieved from <http://journal2.uad.ac.id/index.php/irip>. (in Eng.).

Haryadi R. & Pujiastuti H. 2022 – *Haryadi R. & Pujiastuti H.* Enhancing Pre-service Physics Teachers' Higher Order Thinking Skills Through STEM-PjBL Model. *International Journal of STEM Education for Sustainability*, 2(2): 156–171. <https://doi.org/10.53889/ijses.v2i2.38>. (in Eng.).

«Цифрлық Қазақстан» мемлекеттік бағдарламасы. Қазақстан Республикасы Үкіметінің 2017 жылғы 12 желтоқсандағы № 827 қаулысы. <http://adilet.zan.kz> 20.04.2023. (in Kaz.).

REFERENCES

- Apriyani R., Ramalis T.R. & Suwarma I.R., 2019 – *Apriyani R., Ramalis T.R. & Suwarma I.R.* Analyzing Student's Problem Solving Abilities of Direct Current Electricity in STEM-based Learning. *Journal of Science Learning*, 2(3). Pp. 85–91. <https://doi.org/10.17509/jsl.v2i3.17559>. (in Eng.).
- Baran M., Karakoyun F. & Maskan A., 2021 – *Baran M., Karakoyun F. & Maskan A.* The Influence of Project-Based STEM (PjBL-STEM) Applications on the Development of 21st-Century Skills. *Journal of Turkish Science Education*, 18(4). Pp.798–815. <https://doi.org/10.36681/tused.2021.104>. (in Eng.).
- Badeleh A., 2021 – *Badeleh A.* The effects of robotics training on students' creativity and learning in physics. *Education and Information Technologies*, 26(2). Pp. 1353–1365. <https://doi.org/10.1007/s10639-019-09972-6> (in Eng.).
- Haryadi R. & Pujiastuti H. 2022 – *Haryadi R. & Pujiastuti H.* Enhancing Pre-service Physics Teachers' Higher Order Thinking Skills Through STEM-PjBL Model. *International Journal of STEM Education for Sustainability*, 2(2). Pp. 156–171. <https://doi.org/10.53889/ijses.v2i2.38>. (in Eng.).
- Kazbekova G.N., Ismagulova Zh.S., 2022 – *Kazbekova G.N., Ismagulova Zh.S.* formation of innovative STEM-educational approach// *Bulletin of Yasavi University*. 3(125). Pp. 200–210. <https://doi.org/10.47526/2022-3/2664-0686.17>. (in Kaz.).
- Nazifah N. & Asrizal A., 2022 – *Nazifah N. & Asrizal A.* Development of STEM Integrated Physics E-Modules to Improve 21st Century Skills of Students. *Jurnal Penelitian Pendidikan IPA*, 8(4). Pp. 2078–2084. <https://doi.org/10.29303/jppipa.v8i4.1820>. (in Eng.).
- Nurramadhani A., 2020 – *Nurramadhani A.* Analysis of Students' STEM Literasy Based on Gender Differences in Science Learning. *JHSS (Journal of humanities and social studies)*, 4(1). Pp. 21–25. <https://doi.org/10.33751/jhss.v4i1.1903>. (in Eng.).
- Ramankulov Sh., Choruh A., Polatuly S., 2022 – STEAM technology as a tool for developing creativity of students: on the example of a school physics course// *Ясауи университетінің хабаршысы*. –4(126). Pp. 200–211. <https://doi.org/10.47526/2022-4/2664-0686.17>. (in Eng.).
- Susilawati S., Doyan A., Mulyadi L., Abo C.P. & Pineda C.I.S., 2022 – *Susilawati S., Doyan A., Mulyadi L., Abo C.P. & Pineda C.I.S.* The Effectiveness of Modern Physics Learning Tools Using the PhET Virtual Media Assisted Inquiry Model in Improving Cognitive Learning Outcomes, Science Process Skills, and Scientific Creativity of Prospective Teacher Students. *Jurnal Penelitian Pendidikan IPA*, 8(1). Pp. 291–295. <https://doi.org/10.29303/jppipa.v8i1.1304>. (in Eng.).
- Shute V.J. & Rahimi S., 2021 – Shute V.J. & Rahimi S. Stealth assessment of creativity in a physics video game. *Computers in Human Behavior*, 116. <https://doi.org/10.1016/j.chb.2020.106647>. (in Eng.).
- Sulaiman F., Rosales J. J., & Kyung L. J., 2023 – *Sulaiman F., Rosales J. J., & Kyung L. J.* The effectiveness of the integrated stem-pbl physics module on students' interest, sensemaking and effort. *Journal of Baltic Science Education*, 22(1). Pp. 113–129. <https://doi.org/10.33225/jbse/23.22.113>. (in Eng.).
- Saputro V.C.E., Wasis & Prastowo T., 2023 – *Saputro V.C.E., Wasis & Prastowo T.* The Effectiveness of STEM-Based Guided Inquiry Learning to Train Science Literacy of Physics. *Studies in Learning and Teaching*, 3(3). Pp. 141–148. <https://doi.org/10.46627/silet.v3i3.179>. (in Eng.).
- The state program «Digital Kazakhstan», 2017 – Resolution of the Government of the Republic of Kazakhstan dated December 12: 827. <http://adilet.zan.kz> 20.04.2023. (in Kz.).

© **A.E. Kulakayeva¹, Y.A. Daineko¹, A.Z. Aitmagambetov¹,
A.T. Zhetpisbaeva², B.A. Kozhakhmetova^{1*}, 2023**

¹JSC «International Information Technologies University», Almaty, Kazakhstan;

²Kazakh Agrotechnical University named after S.Seifullin, Astana, Kazakhstan.

E-mail: kozhahmetova.ba@gmail.com

ABOUT THE INFLUENCE OF THE ORBIT CHARACTERISTICS OF A SMALL SPACECRAFT ON THE PARAMETERS OF THE SATELLITE RADIO MONITORING SYSTEM

Kulakayeva Aigul — PhD, associate professor International Information Technology University.
050040. Almaty, Kazakhstan

E-mail: aigul_k.pochta@mail.ru. ORCID: <https://orcid.org/0000-0002-0143-085X>;

Daineko Yevgeniya — PhD, associate professor. International Information Technology University.
050040. Almaty, Kazakhstan

E-mail: yevgeniyadaineko@gmail.com. ORCID: <https://orcid.org/0000-0001-6581-2622>;

Aitmagambetov Altai — candidate of technical sciences, professor. International Information
Technology University. 050040. Almaty, Kazakhstan

E-mail: altayzf@mail.ru. ORCID: <https://orcid.org/0000-0002-7808-5273>;

Zhetpisbayeva Ainur — PhD, associate professor. Kazakh Agrotechnical University named after
S.Seifullin. 010000. Astana. Kazakhstan

E-mail: aigulji@mail.ru. ORCID: <https://orcid.org/0000-0002-4525-5299>;

Kozhakhmetova Bagdat — master, senior lecturer. International Information Technology University.
050040. Almaty, Kazakhstan

E-mail: kozhahmetova.ba@gmail.com. ORCID: <https://orcid.org/0000-0002-9566-3629>.

Abstract. Within the framework of the existing ground-based radio monitoring facilities of the Republic of Kazakhstan, it is impossible to qualitatively perform the functions and tasks of radio monitoring. For countries with a large territory, such as the Republic of Kazakhstan, to increase the efficiency of the radio monitoring system for the use of the radio frequency spectrum, it is proposed to use low-orbit small spacecraft as radio monitoring stations. In this paper, an analysis of the energy budget of radio lines of various sources of radio emissions was performed, which showed the possibility of using low-orbit small spacecraft for satellite radio monitoring. The characteristics of the trajectory of a low-orbit small spacecraft were determined and the coverage (radio availability) of a satellite radio monitoring system based on one low-orbit small spacecraft was analyzed, and the area of radio

monitoring of the territory of the Republic of Kazakhstan was determined. The proposed system will allow for more effective control of radio-electronic means and increase the level of management of the use of the radio frequency spectrum in the Republic of Kazakhstan. Also in this work, such programs as «TRASSA-OMIR» and Satellite Tool Kit were used. With the help of these programs, the motion of a small spacecraft in orbit was simulated, and the most optimal measurement modes were determined.

Keywords: radio monitoring, low orbit, small spacecraft, radio electronic means, radio frequency spectrum, radio control, SNR

This work was financially supported by the Science Committee of the Ministry of Science and Higher Education of the Republic of Kazakhstan (grant №AP08857146, 2020–2022).

© А.Е. Кулакаева¹, Е.А. Дайнеко¹, А.З. Айтмагамбетов¹,
А.Т. Жетписбаева², Б.А. Кожаметова^{1*}, 2023

¹«Халықаралық ақпараттық технологиялар университеті» АҚ,
Алматы, Қазақстан;

²С. Сейфуллин атындағы Қазақ агротехникалық университеті,
Астана, Қазақстан.

E-mail: kozhahmetova.ba@gmail.com

ШАҒЫН ҒАРЫШ АППАРАТЫ ОРБИТАСЫНЫҢ СИПАТТАМАЛАРЫНЫҢ СПУТНИКТИК РАДИО МОНИТОРИНГ ЖҮЙЕСІНІҢ ПАРАМЕТРЛЕРІНЕ ӘСЕРІ ТУРАЛЫ

Кулакаева Айгүль — PhD, қауымдастырылған профессордың ж.а. Халықаралық ақпараттық технологиялар университеті. 050040. Алматы, Қазақстан

E-mail: aigul_k.pochta@mail.ru. ORCID: 0000-0002-0143-085X;

Дайнеко Евгения — PhD, қауымдастырылған профессор. Халықаралық ақпараттық технологиялар университеті. 050040. Алматы, Қазақстан

E-mail: yevgeniyadaineko@gmail.com. ORCID: 0000-0001-6581-2622;

Айтмагамбетов Алтай — т.ғ.к., профессор. Халықаралық ақпараттық технологиялар университеті. 050040. Алматы, Қазақстан

E-mail: altayzf@mail.ru. ORCID: 0000-0002-7808-5273;

Жетписбаева Айнура — PhD, қауымдастырылған профессор. С. Сейфуллин атындағы Қазақ агротехникалық университеті. 010000. Астана, Қазақстан

E-mail: aigulji@mail.ru. ORCID: 0000-0002-4525-5299;

Кожаметова Бағдат — сениор-лектор. Халықаралық ақпараттық технологиялар университеті. 050040. Алматы, Қазақстан

E-mail: kozhahmetova.ba@gmail.com. ORCID: 0000-0002-9566-3629

Аннотация. Қазақстан Республикасының қолданыстағы жерүсті радиобақылау құралдары шеңберінде радиобақылаудың функциялары мен міндеттерін сапалы орындау мүмкін емес. Аумағы үлкен елдер үшін Қазақстан Республикасы ретінде радиожілік спектрін пайдаланудың радио мониторингі жүйесінің тиімділігін арттыру үшін радио бақылау станциялары

ретінде төмен орбиталық шағын ғарыш аппараттарды пайдалану ұсынылады. Бұл жұмыста әртүрлі радио сәулелену көздерінің радиолинияларының энергетикалық бюджетіне талдау жасалды, бұл спутниктік радио мониторингті жүзеге асыру үшін төмен орбиталық шағын ғарыш аппараттарын қолдану мүмкіндігін көрсетті. Төмен орбиталық шағын ғарыш аппаратының қозғалыс траекториясының сипаттамалары айқындалды және бір төмен орбиталық шағын ғарыш аппаратының базасында спутниктік радиомониторинг жүйесінің жабынына (радиоқолжетімділігіне) талдау жүргізілді, сондай-ақ Қазақстан Республикасы аумағының радиомониторинг саласы айқындалды. Ұсынылған жүйе радиоэлектрондық құралдарды неғұрлым тиімді бақылауға және Қазақстан Республикасында радиожілік спектрін пайдалануды басқару деңгейін арттыруға мүмкіндік береді. Сондай-ақ, бұл жұмыста «ТРАССА-ӨМІР» және Satellite Tool Kit сияқты бағдарламалар қолданылды. Осы бағдарламалардың көмегімен орбитадағы шағын ғарыш аппараттарының қозғалысын модельдеу жүргізілді, сондай-ақ өлшеулерді жүргізудің ең оңтайлы режимдері анықталды.

Түйін сөздер: радио мониторинг, төмен орбита, шағын ғарыш аппараты, радиоэлектрондық құрал, радиожілік спектрі, радио бақылау, SNR

Бұл жұмысты Қазақстан Республикасы Ғылым және жоғары білім министрлігінің Ғылым комитеті қаржылай қолдады (грант №АР08857146, 2020–2022).

© А.Е. Кулакаева¹, Е.А. Дайнеко¹, А.З. Айтмагамбетов¹,
А.Т. Жетписбаева², Б.А. Кожаметова^{1*}, 2023

¹АО «Международный университет информационных технологий»,
Алматы, Казахстан;

²Казахский агротехнический университет имени С.Сейфуллина,
Астана, Казахстан.

E-mail: kozhahmetova.ba@gmail.com

О ВЛИЯНИИ ХАРАКТЕРИСТИК ОРБИТЫ МАЛОГО КОСМИЧЕСКОГО АППАРАТА НА ПАРАМЕТРЫ СИСТЕМЫ СПУТНИКОВОГО РАДИОМОНИТОРИНГА

Кулакаева Айгуль — PhD, и.о. ассоциированный профессор. Международный университет информационных технологий. 050040. Алматы, Казахстан

E-mail: aigul_k.pochta@mail.ru. ORCID: 0000-0002-0143-085X;

Дайнеко Евгения — PhD, ассоциированный профессор. Международный университет информационных технологий. 050040. Алматы, Казахстан

E-mail: yevgeniyadaineko@gmail.com. ORCID: 0000-0001-6581-2622;

Айтмагамбетов Алтай — к.т.н., профессор. Международный университет информационных технологий. 050040. Алматы, Казахстан

E-mail: altayzf@mail.ru. ORCID: 0000-0002-7808-5273

Жетписбаева Айнур — PhD, ассоциированный профессор. Казахский агротехнический университет им. С. Сейфуллина. 010000. Астана, Казахстан

E-mail: aigulji@mail.ru. ORCID: 0000-0002-4525-5299;

Кожаметова Багдат — сениор-лектор. Международный университет информационных технологий. 050040. Алматы, Казахстан
E-mail: kozhahmetova.ba@gmail.com. ORCID: 0000-0002-9566-3629.

Аннотация. В рамках существующих наземных средств радиомониторинга Республики Казахстан невозможно качественно выполнять функции и задачи радиоконтроля. Для стран с большой территорией как Республика Казахстан для повышения эффективности системы радиомониторинга использования радиочастотного спектра предлагается использовать низкоорбитальные малые космические аппараты в качестве станций радиоконтроля. В данной работе был выполнен анализ энергетического бюджета радиолиний различных источников радиоизлучений, который показал возможность применения низкоорбитальных малых космических аппаратов для осуществления спутникового радиомониторинга. Определены характеристики траектории движения низкоорбитального малого космического аппарата и проведен анализ покрытия (радиодоступности) системы спутникового радиомониторинга на базе одного низкоорбитального малого космического аппарата, а также определена область радиомониторинга территории Республики Казахстан. Предлагаемая система радиомониторинга позволит более эффективно контролировать радиоэлектронные средства и повысить уровень управления использованием радиочастотным спектром в Республике Казахстан. Также в данной работе были применены такие программы как «ТРАССА-ОМИР» и Satellite Tool Kit. При помощи данных программ было произведено моделирование движения малого космического аппарата на орбите, а также были определены наиболее оптимальные режимы проведения измерений.

Ключевые слова: радиомониторинг, низкая орбита, малый космический аппарат, радиоэлектронное средство, радиочастотный спектр, радиоконтроль, SNR

Работа была выполнена при финансовой поддержке Комитета науки Министерства науки и высшего образования Республики Казахстан (грант №AP08857146, 2020–2022).

Благодарность. Работа выполнена при финансовой поддержке КН МОН РК по программе грантового финансирования научных исследований на 2020-2022 гг., грант №AP08857146.

Введение

Системы радиомониторинга играют ключевую роль в управлении с использованием радиочастотного спектра (РЧС). Однако выполнить все функции и задачи радиомониторинга на территории даже небольшой страны без применения современной системы радиоконтроля практически невозможно. Известно (Handbook on Spectrum Monitoring, Radiocommunication Bureau, 2011), что уровень развития телекоммуникационной отрасли и методы регулирования РЧС в разных странах отличаются. В настоящее время пока

не существует универсальной методики и системы для построения системы радиомониторинга. Каждая страна создает систему радиомониторинга исходя от уровня развития телекоммуникационной отрасли, а также из своих финансовых возможностей и из потребностей других ведомств, которые в той или иной степени связаны с использованием РЧС (Navarro и др., 2013; Souryal и др., 2015; Zhang и др., 2020; Chen и др., 2012; Konkin 2015; Dokuchaev и др., 2018).

При радиомониторинге особое место занимает определение местоположения источников радиоизлучения (ИРИ), что является сложной задачей. Для определения с максимальной точностью местоположения исследуемых ИРИ необходимо применение автоматизированного комплекса специального оборудования с заданными техническими характеристиками. В системах радиомониторинга для точного определения местоположения ИРИ необходимо применение эффективных методов и алгоритмов.

Для определения местоположения ИРИ используются такие методы, как: амплитудный, угломерный, дальномерный, разностно-дальномерный, разностно-частотный и другие методы. Однако на практике для увеличения достоверности при определении местоположения ИРИ часто используют комбинированные методы на основе двух или более вышеперечисленных методов. Такой подход позволяет устранить недостатки одного метода, дополнив свойствами другого (Skolnik 2008; Cetin и др., 2016; ITU Report-R SM.2211-1, 2014; Mark и др., 2006; Buehrer и др., 2019; Gel и др., 2008).

Для создания системы спутникового радиомониторинга на начальном этапе необходимо произвести оценку уровней сигналов на входе бортового измерительного приемника от различных наземных ИРИ. Такую оценку можно провести на основе методики расчета энергетического бюджета различных радиоканалов связи в соответствии с рекомендациями Международного Союза Электросвязи (МСЭ) (Aitmagambetov и др., 2021; Aitmagambetov и др., 2019).

Для этих целей были выбраны номинальные значения характеристик оборудования и сигналов реальных действующих наземных радиоэлектронных средств (РЭС), функционирующих на территории Республики Казахстан в диапазоне частот от 94 МГц до 14 ГГц (ОВЧ, УВЧ и СВЧ диапазонов).

Материалы и основные методы

Для расчета и анализа энергетического бюджета в качестве ИРИ для примера были выбраны: станции мобильной связи различных стандартов GSM-900 МГц, UMTS-1800 МГц, LTE-2100 МГц; станция наземного цифрового телевидения стандарта DVB-T2 (TV) (474 МГц) 21 телевизионный канал; наземное звуковое вещание (FM) 94 МГц; земные станций спутниковой связи в диапазонах X -7 ГГц и Ku-14 ГГц (Aitmagambetov и др., 2021; Aitmagambetov и др., 2019).

Анализ уровней сигналов на входе приемника системы радиомониторинга

показал, что для рассматриваемых наземных РЭС отношение сигнал/шум (SNR) больше 10 дБ, что приемливо для осуществления спутникового радиомониторинга на базе низкоорбитального малого космического аппарата (МКА). Поэтому такие системы возможно использовать в целях повышения эффективности системы радиомониторинга использования радиочастотного спектра в качестве станций радиоконтроля.

Для определения координат ИРИ предлагается использовать разработанный способ (Aitmagambetov и др., 2021), основанный на угломерном методе, с применением на борту МКА двух сканирующих антенн типа активная фазированная антенная решетка (АФАР) (рисунок 1), луч одной из которых сканирует в меридианальном направлении для определения пеленга на источников радиоизлучения в момент фиксации сигнала методом равносигнальной зоны (для определения широты источников радиоизлучения), а луч второй антенны сканирует в ортогональном направлении движению МКА для определения пеленга на источники радиоизлучения (для определения долготы источников радиоизлучения), а местоположение источников радиоизлучения определяются на основе анализа геометрических соотношений расстояний и углов между МКА, источников радиоизлучения и центром масс Земли, пеленгами на источники радиоизлучения с использованием итераций и решения уравнения 4-ой степени.

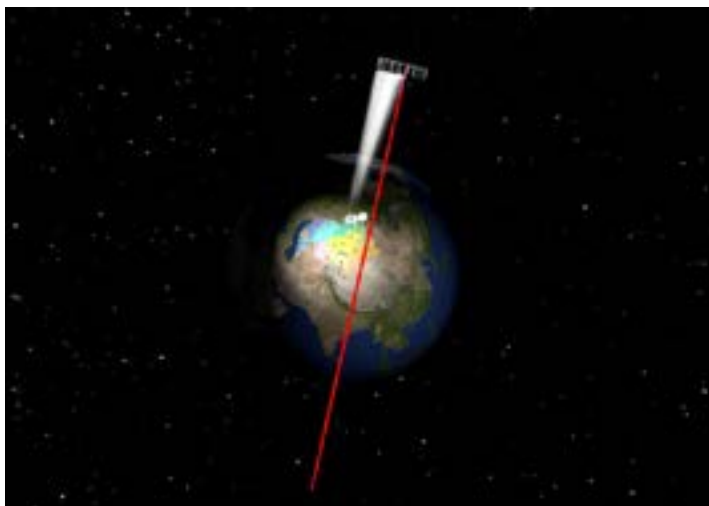


Рисунок 1. Определение местоположения источника радиоизлучения

Область радиомониторинга (ОРМ) территории Республики Казахстан находится в диапазоне значений северной широты от $\varphi = 40^\circ$ до $\varphi = 56^\circ$ (среднее значение ОРМ $\varphi = 48^\circ$), а также в диапазоне значений восточной долготы от $\nu = 45^\circ$ до $\nu = 87^\circ$. Как отмечалось ранее, на сегодняшний день

в рамках существующих наземных средств радиомониторинга невозможно качественно выполнить функции и задачи радиоконтроля, учитывая столь большую территорию нашей страны и трудоемкость проведения процедур радиоконтроля с помощью наземного радиоизмерительного оборудования в сложных климатических условиях и в условиях сложного рельефа местности. Для стран с большой территорией как Республика Казахстан для повышения эффективности системы радиомониторинга использования радиочастотного спектра целесообразно использовать низкоорбитальные малые космические аппараты в качестве станций радиоконтроля, так как такие системы охватывают значительные по площади территории и не зависят от рельефа местности, климатических условий (Aitmagambetov и др., 2020; Aitmagambetov и др., 2016).

Результаты

Для определения радиуса круговой орбиты МКА R_0 требуется вычислить радиус эллипсоида (линейное расстояние от центра масс Земли до поверхности), зависящий от широты:

$$Re(\varphi = 48^\circ) = a * b / [(b^2 + a^2 * tg^2 48^\circ)^{0.5} * cos 48^\circ] = 6366299 \text{ м.} \quad (1)$$

где $a = 6378136$ м – экваториальный радиус Земли; $b = 6356751$ м – полярный радиус Земли [18].

Далее определим ряд параметров орбиты МКА:

Радиус круговой орбиты МКА:

$$R_0 = Re(48^\circ) + h = 7016299 \text{ м.} \quad (2)$$

Линейную скорость движения МКА:

$$V = \sqrt{G * M / R_0} = 7542 \text{ м/с.} \quad (3)$$

где $M = 5,98 * 10^{24}$ кг масса Земли и гравитационная постоянная $G = 6,6743 * 10^{-11}$ м³*кг⁻¹*с⁻² [19].

Время одного витка МКА определяется как:

$$T = 2 * \pi * R_0 / V = 5842,264 \text{ с} = 97,37 \text{ мин} \quad (4)$$

Время нахождения МКА над территорией Республики Казахстан:

$$t = T * (56^\circ - 40^\circ) / 360^\circ = 4,33 \text{ мин} \quad (5)$$

Для наглядного примера на рисунке 2 приведена зона радиовидимости МКА.



Рисунок 2. Зона радиовидимости

Выполнение функций радиомониторинга на начальном этапе желательно начать с использованием одного низкоорбитального МКА, так как это является наиболее привлекательным с экономической точки зрения. При проектировании системы радиомониторинга на базе одного низкоорбитального МКА необходимо начинать с выбора параметров орбиты. Для этого предлагается рассмотреть круговую полярную орбиту (наклонение $i = 92^\circ$), и высоту орбиты МКА $h = 650$ км. МКА с подобными характеристиками используются в спутниковых системах ДЗЗ дистанционного зондирования Земли (ДЗЗ) (Aver'janov и др., 2021)

Зона покрытия одного спутника представляет собой круговую область (рисунок 3) на поверхности Земли, на которой виден спутник. При этом зона покрытия будет зависеть от высоты орбиты МКА. На рисунке 4 приведена траектория движения одного низкоорбитального МКА над территорией РК за сутки.

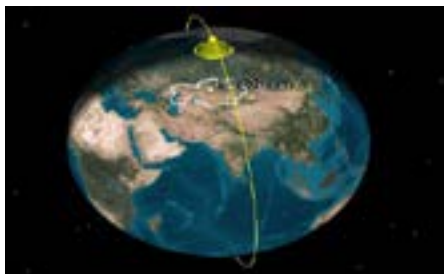


Рисунок 3. Радиомониторинг на базе одного низкоорбитального МКА



Рисунок 4. Траектория движения спутника за сутки

Зона покрытия спутника на Земле зависит от параметров орбиты. При этом земная станция обработки информации радиомониторинга может осуществлять связь со спутником только тогда, когда наземная станция находится в зоне покрытия (след спутника), как показано на рисунке 5. Продолжительность видимости и, следовательно, продолжительность связи составляет несколько минут, поскольку спутник движется над Землей с большой скоростью. Вместе со спутником перемещается и след от него и, следовательно, теряется связь с МКА (рисунок 6).



Рисунок 5. Зона покрытия одного низкоорбитального МКА



Рисунок 6. Перемещение МКА

Необходимо рассмотреть, как изменится долгота спутника на второй позиции по отношению к первой за счет вращения Земли. Для этого надо найти угловую скорость вращения Земли ω_3 .

Угловую скорость Земли определяется как:

$$\omega_3 = \frac{2\pi}{T} \Rightarrow \frac{360^\circ}{T}$$

где T – период обращения Земли вокруг своей оси.

Так как в сутках 24 часа, следовательно, можно предположить, что период обращения Земли вокруг своей оси T составит также 24 часа. Но так как Земля

вращается еще вокруг Солнца, то период обращения её вокруг собственной оси будет немного короче привычных нам солнечных суток и составит 23 часа 56 минут и 4 секунды. Это так называемые звездные сутки. В пересчете на секунды мы получаем: $T=86164$ с.

Получается:

$$\omega_3 = \frac{360^\circ}{86164} = 0,00417^\circ/\text{с}$$

Чтобы найти число витков в сутки, необходимо среднее количество минут в сутках разделить на время одного витка МКА, тогда получается:

$$P_v = \frac{t_{\text{ср}}}{T} = \frac{1440}{97,37} = 14,8 \approx 15 \text{ витков в сутки}$$

где – это среднее количество минут в сутках.

Изменение долготы на следующем витке составит:

$$\Theta_{\text{изм}} = \omega_3 * T = 0,00417 * 5842,264 = 24,343^\circ$$

т.е. смещение за сутки будет происходить на относительно начала суток.

При этом угловая скорость движения МКА ω_c за один час составит:

$$\omega_c = 360/5842,264=0,06162^\circ/\text{с}$$

Число витков в час:

$$P_v/24 = 14,8/24=0,616$$

На рисунке 7 приведена траектория движения МКА над территорией РК, и можно увидеть влияние вращения Земли. Из-за вращения Земли расстояние между витками составит 24,343 градусов (около 2700 км).



а) первый виток



б) второй виток

Рисунок 7. Траектория движения МКА

Далее рассмотрим вопросы определения координат и длительности нахождения низкоорбитального спутника над территорией Республики Казахстан и интервалы его появления, а также методы ориентации МКА с целью создания системы радиомониторинга на основе низкоорбитального спутника для определения координат наземных ИРИ. Исследования проводились с помощью программы по моделированию движения спутника на орбите Земля «ТРАССА-ОМИР» (Institute of space technique and technology, 2022), которая может производить моделирование движения спутника на орбите Земли. Программа создана специалистами ДТОО «Институт космической техники и технологий» на основе изучения различных существующих зарубежных аналогов.

Данная программа позволяет для заданного временного интервала с заданным шагом по времени рассчитывать положение спутника на орбите (высоты, широты и долготы подспутниковой точки), а также ряд сопутствующих параметров – L-оболочки, освещенности спутника, прямой видимости спутника из точки на поверхности Земли, параметров геомагнитного поля и др.

С помощью данной программы были промоделированы все точки пролета над территорией Республики Казахстан за период с 01.12.2020 по 08.12.2020 (см. рисунок 8).



Рисунок 8. Моделирование движения МКА системы спутникового радиомониторинга

Обсуждение

В программе была применена область считывания по территории Республики Казахстан. Это область находится в диапазоне значений координат от 40° до 55° северной широты и от 45° до 87° восточной долготы. Была

создана таблица границ территории РК и выписаны точки пролета внутри границы.

Так, например, 01.12.2020 спутник находился над Усть-Каменогорском с 05:19:30 до 05:20:45 (75 секунд), с 06:55:00 до 06:57:15 (135 секунд) пролетал от Костаная в направлении Актобе, с 16:13:00 до 16:14:00 (60 секунд) был над Усть-Каменогорском, с 17:48:45 до 17:50:15 (90 секунд) в направлении от Кызылорды в Актобе. В этот день спутник сначала пролетал над восточной частью Казахстана и через 1 час 35 минут он снова пролетит над РК, но только над северной частью. Через 9 часов 18 минут спутник снова пролетит над ВКО. И через 1 час 35 минут он будет над Кызылординской областью. Итоговое время 6 мин. Таким образом, спутник в среднем будет 4 раза по 90 секунд над РК. За это время необходимо провести максимальное количество измерений параметров сигналов и необходимых углов для более точного определения местонахождения искомым ИРИ.

Заключение

Таким образом, учитывая большую территорию страны для повышения эффективности системы радиомониторинга предлагается применение низкоорбитального МКА в качестве станции радиоконтроля.

Предлагаемая спутниковая система позволит значительно расширить зону радиомониторинга и ускорить получение данных, необходимых для проведения анализа использования радиочастотного спектра и может интегрироваться с существующей наземной сетью радиомониторинга Республики Казахстан и с международной сетью радиомониторинга Международного Союза Электросвязи (МСЭ). В работе были применены программы «ТРАССА-ОМИР» и Satellite Tool Kit, с помощью которых было произведено моделирование движения малого космического аппарата на орбите и были определены наиболее оптимальные режимы проведения измерений.

REFERENCES

Handbook on Spectrum Monitoring. Radiocommunication Bureau, 2011. Geneva, Electronic Publication

Navarro A., Restrepo J., 2013 — *Navarro A., Restrepo J.* A new method for spectrum monitoring networks design. 2013 IEEE Antennas and Propagation Society International Symposium (APSURSI). Pp. 2018–2019. DOI:10.1109/aps.2013.6711667 (in Eng.)

Souryal M. et al., 2015 — *Souryal M. et al.* Real-time centralized spectrum monitoring: Feasibility, architecture, and latency. 2015 IEEE International Symposium on Dynamic Spectrum Access Networks (DySPAN). Pp. 106–112. DOI:10.1109/dyspan.2015.7343894 (in Eng.)

Zhang J. et al. 2020 – *Zhang J. et al.* Spectrum Knowledge and Real-Time Observing Enabled Smart Spectrum Management. IEEE Access. 8: 44153-44162. DOI:10.1109/access.2020.2978005. (in Eng.)

Chen D. et al. 2012 - *Chen D. et al.* Spectrum occupancy analysis based on radio monitoring network. 2012 1st IEEE International Conference on Communications in China (ICCC). Pp.739–744. DOI: 10.1109/iccchina.2012.6356981. (in Eng.)

Konkin V.V. 2015 - *Konkin V.V.* Options for building a unified radio monitoring system of the Moscow region [Varianty postroeniya edinoj sistemy radiokontrolja Moskovskogo regiona]. T-Comm-Telecommunications and Transport.9:6. (in Rus.)

Dokuchaev V.A., Pavlov S.V. 2018 - *Dokuchaev V.A., Pavlov S.V.* Methodological foundations for building a model of radio monitoring equipment [O razrabotke metodologicheskikh osnov postroeniya modeli tehniceskikh sredstv radiomonitoringa] T-Comm, 12(7): Pp. 48–51. (in Rus.)

Skolnik M., 2008 - *Skolnik M.* Radar Handbook. Third Edition. McGraw-Hill Education. 1351 p. ISBN: 978-0-071-48547-0

Cetin E., Dempster A.G., 2016 - *Cetin E., Dempster A.G.* Interference Localization for Satellite Navigation Systems. Proceedings of the IEEE, 104(6): Pp.1318–1326. [7439734]. DOI:10.1109/JPROC.2016.2530814 (in Eng.)

ITU Report-R. SM. 2211-1 (06/2014). Comparison of methods for determining the geographical location of the signal source based on the difference in the time of arrival and the angle of arrival of the signal

Mark J.W., Yao G., Zhao L., 2006 - *Mark J.W., Yao G., Zhao L.* Mobile positioning based on relaying capability of mobile stations in hybrid wireless networks. IEEE Proc.-Commun, 153 (5)

Buehrer R.M., Zekavat A.S., 2019 - *Buehrer R.M., Zekavat A.S.* Handbook of position location theory, practice, and advances. Second Edition. IEEE Press Wiley. 1376 p. ISBN 978-1-119-43458-0

Gel V.E., Gretsova V., Potapov I.A., 2008 - *Gel V.E., Gretsova V.P., Potapov I.A.* A brief analysis of known methods for determining the location of radio emission sources. Doppler system for measuring the coordinates of radio emission sources [Kratkij analiz izvestnykh sposobov opredeleniya mestopolozheniya istochnikov radioizluchenij. Doplerovskaja sistema izmereniya koordinat istochnikov radioizluchenij]. Information and space. 3: Pp. 40–47. (in Rus.)

Aitmagambetov A., Butuzov Y., Tikhvinskiy V., Kulakayeva A., Ongenbayeva Z., 2021 - *Aitmagambetov A., Butuzov Y., Tikhvinskiy V., Kulakayeva A., Ongenbayeva Z.* Energy budget and methods for determining coordinates for a radiomonitoring system based on a small spacecraft. Indonesian Journal of Electrical Engineering and Computer Science, 21(2): Pp. 945–956 (in Eng.)

Aitmagambetov A.Z., Kulakaeva A.E., Kozhakhmetova B.A., Zhaksylyk A.Zh., 2019 - *Aitmagambetov A.Z., Kulakaeva A.E., Kozhakhmetova B.A., Zhaksylyk A.Zh.* Assessment of the energy budget for a radio monitoring system based on low-orbit satellites [Ocenka jenergeticheskogo byudzhetu dlja sistemy radiomonitoringa na baze nizkoorbital'nyh sputnikov]. Bulletin of the AUES.4 (47):88. (in Rus.)

Aitmagambetov A.Z. et al., 2020 - *Aitmagambetov A.Z. et al.* Analysis of the possibility of using a low-orbit satellite for radio monitoring [Analiz vozmozhnosti primeneniya nizkoorbital'nogo sputnika dlja radiomonitoringa]. Questions of technical and physical-mathematical sciences in the light of modern research. Pp.57–64. (in Rus.)

Aitmagambetov A.Z., Butuzov Yu.A., Kulakayeva A.E., 2016 - *Aitmagambetov A.Z., Butuzov Yu.A., Kulakayeva A.E.* Mathematical models for determining the location of radio emission sources in radio monitoring systems on the basis on low-orbit satellites. T-Comm. 10(1): Pp. 73–76.

Parameters of the Earth in 1990 (PZ-90.11)., 2014 Reference document. Military Topographic Directorate of the General Staff of the Armed Forces of the Russian Federation. 52 p. (in Rus.)

CODATA Recommended values of the fundamental physical constants (2019): 2018 NIST SP 961.

Aver'janov A.V., Kaljuzhnyj A.V. & Kuznecov V.V., 2021 - *Aver'janov A.V., Kaljuzhnyj A.V. & Kuznecov V.V.* Methodology for substantiating the functional distribution of small spacecraft in the cluster of remote sensing of the Earth [Metodika obosnovaniya funkcional'nogo raspredeleniya malyh kosmicheskikh apparatov v klasteru distancionnogo zondirovaniya Zemli]. Izvestiya vysshih uchebnykh zavedenij. Priborostroenie, 64 (8): Pp.620–625. (in Rus.)

Institute of space technique and technology (2022) Official website. URL: <https://istt.kz/eng>

NEWS OF THE NATIONAL ACADEMY OF SCIENCES OF THE REPUBLIC OF KAZAKHSTAN
PHYSICO-MATHEMATICAL SERIES

ISSN 1991-346X

Volume 2. Number 346 (2023). 221–236

<https://doi.org/10.32014/2023.2518-1726.195>

UDC 28.23.25

© **A.E. Nazyrova***, **G.T. Bekmanova**, **A.S. Mukanova**, **N. Amangeldi**,
M.Zh. Kaldarova, 2023

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

E-mail: nazyrova_aye_1@enu.kz

DEVELOPMENT OF AN AUTOMATED SYSTEM FOR EDUCATIONAL PROGRAMS

Bekmanova Gulmira Tleuberdievna — candidate of technical sciences, PhD, associate professor of the Department of artificial intelligence, L.N. Gumilyov Eurasian National University, Astana, Kazakhstan

E-mail: gulmira-r@yandex.kz. ORCID ID: <https://orcid.org/0000-0002-9162-6791>;

Nazyrova Aizhan Esbolovna — doctoral student of the Department of artificial intelligence, L.N. Gumilyov Eurasian National University, Astana, Kazakhstan

E-mail: nazyrova_aye_1@enu.kz. ORCID ID: <https://orcid.org/0000-0001-8554-7627>;

Amangeldy Nurzada — senior lecturer of the Department of artificial intelligence, L.N. Gumilyov Eurasian National University, Astana, Kazakhstan

E-mail: nurzadaamangeldy@gmail.com. ORCID ID: <https://orcid.org/0000-0002-4669-9254>;

Mukanova Asel Serikovna — associate professor, dean of the Higher School of Information Technologies and engineering, Astana International University, Astana, Kazakhstan

E-mail: assel.mukanova@aiu.edu.kz. ORCID ID: <https://orcid.org/0000-0002-9162-6791>;

Kaldarova Mira Zhorabekkyzy — senior lecturer, of the Higher School of Information Technologies and engineering, Astana International University, Astana, Kazakhstan

E-mail: kmiraj82@mail.ru. ORCID ID: 0000-0001-7494-9794.

Abstract. This article proposes a method of structuring data for automatic design systems of an intelligent EP designer in order to form a list of necessary competencies for students in the specialty being studied. Saving time and reducing shortcomings when creating an EP (Education Program) allows you to improve the process of training specialists at the university with the help of special software capable of processing information about subjects to identify inconsistencies and contradictions in the formation of students' competencies. The reliability of the state educational standard should be taken into account when solving the main tasks of managing educational processes, such as developing an educational program and ensuring the quality of training specialists. The development of the general capabilities of graduates in the future will be the focus of the employer's attention. When organizing the training of graduates, competence passports are prepared. When developing an educational program, the degree of competence demand should be taken into account, and the number of credit units should be distributed among

courses in accordance with the importance of each discipline in the development of competencies. Such procedures as determining the content of a discipline, choosing the best order for studying related disciplines, choosing a department for a specific course and taking into account the requirements of the employer when developing a work program in subjects are interrelated and require regulation. All this is directly related to the growing size and complexity of the document management system, which creates an additional burden on the heads of departments and divisions. For the successful and effective work of many departments, automated information systems currently exist in all universities. However, there are currently no procedures that could help instructors in their work by updating and developing lesson plans in accordance with changing requirements. Such systems should provide quick access to the necessary information from various sources, gaining experience in courses taken in specially organized databases, scheduling work on subjects, minimizing errors in scheduling and simplifying internal work. The purpose of the article: data structuring optimization to identify the most important competencies, build them in a logical hierarchy and develop an effective curriculum so that students can achieve the desired goals and objectives of the course. The data structure presented in the framework of this study can serve as a prerequisite for creating an automatic EP constructor for all university specialties, and the prerequisites of disciplines take into account the correctness of the choice.

Keywords: educational program, intelligent systems, prerequisites, ER (Entity Relationship) diagram, ontological model

© **А.Е. Назырова***, **Г.Т. Бекманова**, **А.С. Муканова**, **Н. Амангелді**,
М.Ж. Калдарова, 2023

Л.Н. Гумилев атындағы Еуразия Ұлттық Университеті, Астана, Қазақстан.

E-mail: nazyrova_aye_1@enu.kz

БІЛІМ БЕРУ БАҒДАРЛАМАЛАРЫ ҮШІН АВТОМАТТАНДЫРЫЛҒАН ЖҮЙЕНІ ӨЗІРЛЕУ

Назырова Айжан Есболевна — жасанды интеллект кафедрасының докторанты, Л.Н.

Гумилев атындағы Еуразия ұлттық университеті, Астана, Қазақстан

E-mail: nazyrova_aye_1@enu.kz. ORCID ID: <https://orcid.org/0000-0001-8554-7627>;

Бекманова Гүлмира Тлеубердиевна — т.ғ.к., PhD, жасанды интеллект кафедрасының

доценті, Л.Н. Гумилев атындағы Еуразия ұлттық университеті, Астана, Қазақстан

E-mail: gulmira-g@yandex.kz. ORCID ID: <https://orcid.org/0000-0002-9162-6791>;

Амангелді Нұрзада — аға оқытушы, жасанды интеллект кафедрасының доценті, Л.Н.

Гумилев атындағы Еуразия ұлттық университеті, Астана, Қазақстан

E-mail: nurzadaamangeldy@gmail.com. ORCID ID: <https://orcid.org/0000-0002-4669-9254>;

Муканова Асель Сериковна — доцент, Ақпараттық технологиялар және инженерия жоғары

мектебінің деканы, Астана халықаралық университеті, Астана, Қазақстан

E-mail: assel.mukanova@aiu.edu.kz. ORCID ID: <https://orcid.org/0000-0002-8964-3891>;

Калдарова Мира Жорабекқызы — аға оқытушы, Ақпараттық технологиялар және

инженерия жоғары мектебінің деканы, Астана халықаралық университеті, Астана, Қазақстан

E-mail: kmiraj82@mail.ru. ORCID ID: <https://orcid.org/0000-0001-7494-9794>.

Аннотация. Бұл мақалада студенттерде оқытылатын мамандық бойынша қажетті құзыреттер тізімін қалыптастыру мақсатында білім беру бағдарламасының интеллектуалды конструкторын автоматты түрде жобалау жүйелері үшін деректерді құрылымдау әдісі ұсынылады. Білім беру бағдарламасын жасау кезінде уақытты үнемдеу және кемшіліктерді азайту студенттердің құзыреттілігін қалыптастыру кезінде сәйкессіздіктер мен қайшылықтарды анықтау үшін пәндер туралы ақпаратты өңдеуге қабілетті арнайы бағдарламалық қамтамасыз ету арқылы ЖОО (Жоғарғы Оқу Орындары) — да мамандарды даярлау процесін жетілдіруге мүмкіндік береді. Мемлекеттік білім беру стандартының сенімділігі білім беру бағдарламасын әзірлеу және мамандарды дайындау сапасын қамтамасыз ету сияқты білім беру үдерістерін басқарудың негізгі міндеттерін шешу кезінде ескерілуі тиіс. Түлектердің ортақ мүмкіндіктерін дамыту болашақта жұмыс берушінің назарында болады. Түлектерді дайындауды ұйымдастыру кезінде құзыреттілік паспорттары дайындалады. Білім беру бағдарламасын жасау кезінде құзыреттілік сұранысының дәрежесі ескерілуі керек, ал кредиттік бірліктердің саны құзыреттерді дамытудағы әрбір пәннің маңыздылығына сәйкес курстар бойынша бөлінуі керек. Пәннің мазмұнын анықтау, сабақтас пәндерді оқудың ең жақсы ретін таңдау, белгілі бір курсқа кафедра таңдау және пәндер бойынша жұмыс бағдарламасын жасау кезінде жұмыс берушінің талаптарын ескеру сияқты процедуралар өзара байланысты және реттеуді қажет етеді. Мұның бәрі құжат айналымы жүйесінің өсіп келе жатқан көлеміне және күрделілігіне тікелей байланысты, бұл бөлім басшылары мен бөлімшелерге қосымша салмақ түсіреді. Көптеген кафедралардың табысты және тиімді жұмыс істеуі үшін қазіргі уақытта барлық жоғары оқу орындарында автоматтандырылған ақпараттық жүйелер бар. Дегенмен, қазіргі уақытта ауыспалы талаптарға сәйкес сабақ жоспарларын жаңарту және әзірлеу арқылы нұсқаушыларға олардың жұмысында көмектесетін процедуралар жоқ. Мұндай жүйелер әртүрлі көздерден қажетті ақпаратқа жылдам қол жеткізуге, арнайы ұйымдастырылған мәліметтер базасында өткен курстар бойынша тәжірибе жинақтауға, пәндер бойынша жұмыс кестелерін құруға, жұмыс кестесін құру қателерін азайтуға және ішкі жұмысты жеңілдетуге мүмкіндік беруі керек. *Мақаланың мақсаты:* деректерді құрылымдау ең маңызды құзыреттерді анықтауға, оларды логикалық иерархияға құруға және студенттер курстың қажетті мақсаттары мен міндеттеріне қол жеткізе алатындай тиімді оқу жоспарын жасауды оңтайландыру. Осы зерттеу аясында ұсынылатын мәліметтер құрылымы жоғары оқу орындарының барлық мамандықтары үшін автоматты ББ бағдарламасының (Білім Беру бағдарламасы) конструкторын құрудың алғышарты бола алады және пәндердің пререквизиттері таңдаудың дұрыстығын ескереді.

Түйін сөздер: білім беру бағдарламасы, зияткерлік жүйелер, пререквизиттер, ER (Entity Relationship) диаграмма, онтологиялық модель

© А.Е. Назырова*, Г.Т. Бекманова, А.С. Муканова, Н. Амангелді,
М.Ж. Калдарова, 2023

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

E-mail: nazyrova_aye_1@enu.kz

РАЗРАБОТКА АВТОМАТИЗИРОВАННОЙ СИСТЕМЫ ДЛЯ ОБРАЗОВАТЕЛЬНЫХ ПРОГРАММ

Назырова Айжан Есболовна — докторант кафедры искусственного интеллекта,

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

E-mail: nazyrova_aye_1@enu.kz. ORCID ID: <https://orcid.org/0000-0001-8554-7627>;

Бекманова Гульмира Тлеубердиевна — к.т.н., PhD, доцент кафедры искусственного интеллекта, Евразийский национальный университет им. Л.Н. Гумилева, Астана, Казахстан

E-mail: gulmira-r@yandex.kz. ORCID ID: <https://orcid.org/0000-0002-9162-6791>;

Амангелді Нұрзада — ст. преподаватель кафедры искусственного интеллекта. Евразийский национальный университет им. Л.Н. Гумилева, Астана, Казахстан

E-mail: nurzadaamangeldy@gmail.com. ORCID ID: <https://orcid.org/0000-0002-4669-9254>;

Муканова Асель Сериковна — доцент, декан Высшей школы информационных технологий и инженерии, Международный университет Астана, Астана, Казахстан

E-mail: assel.mukanova@aiu.edu.kz. ORCID ID: <https://orcid.org/0000-0002-8964-3891>;

Қалдарова Мира Жорабекқызы — ст. преподаватель Высшей школы информационных технологий и инженерии, Международный университет Астана, Астана, Казахстан

E-mail: kmiraj82@mail.ru. ORCID ID: <https://orcid.org/0000-0001-7494-9794>.

Аннотация. В данной статье предлагается метод структурирования данных для систем автоматического проектирования интеллектуального конструктора ОП (Образовательная Программа) с целью формирования у студентов списка необходимых компетенций по изучаемой специальности. Экономия времени и уменьшение недостатков при создании ОП позволяет совершенствовать процесс подготовки специалистов в вузе с помощью специального программного обеспечения, способного обрабатывать информацию о предметах для выявления несоответствий и противоречий при формировании компетенций студентов. Надежность государственного образовательного стандарта должна учитываться при решении основных задач управления образовательными процессами, таких как разработка образовательной программы и обеспечение качества подготовки специалистов. Развитие общих возможностей выпускников в будущем будет в центре внимания работодателя. При организации подготовки выпускников подготавливаются паспорта компетенций. При разработке образовательной программы следует учитывать степень компетентностного спроса, а количество кредитных единиц должно быть распределено по курсам в соответствии с важностью каждой дисциплины в развитии компетенций. Такие процедуры, как определение содержания дисциплины, выбор наилучшего порядка изучения смежных дисциплин, выбор кафедры для конкретного курса и учет требований работодателя при разработке рабочей программы по предметам,

взаимосвязаны и требуют регулирования. Все это напрямую связано с растущими размерами и сложностью системы документооборота, что создает дополнительную нагрузку на руководителей отделов и подразделений. Для успешной и эффективной работы многих кафедр в настоящее время во всех вузах существуют автоматизированные информационные системы. Однако в настоящее время нет процедур, которые могли бы помочь инструкторам в их работе, обновляя и разрабатывая планы уроков в соответствии с меняющимися требованиями. Такие системы должны обеспечивать быстрый доступ к необходимой информации из различных источников, получение опыта по курсам, пройденным в специально организованных базах данных, составление графиков работы по предметам, минимизацию ошибок при составлении графиков и упрощение внутренней работы. Цель статьи: структурирование данных оптимизация для выявления наиболее важных компетенций, построения их в логической иерархии и разработки эффективной учебной программы, чтобы студенты могли достичь желаемых целей и задач курса. Структура данных, представляемая в рамках настоящего исследования, может служить предпосылкой для создания автоматического конструктора ОП для всех специальностей вузов, а пререквизиты дисциплин учитывают правильность выбора.

Ключевые слова: образовательная программа, интеллектуальные системы, пререквизиты, ER (Entity Relationship) диаграмма, онтологиялық модель

Кіріспе

Білім беру бағдарламасын құру, мамандарды дайындау сапасын бақылау сияқты оқу процесстерін басқарудың негізгі міндеттері мемлекеттік білім беру стандартының сенімділігін ескере отырып шешілуі керек. Бітірушілер арасында стандартты құзыреттіліктерді қалыптастыру бойынша алдағы уақытта жұмыс берушіге баса назар аударылады.

Түлектерді дайындау процесін жоспарлау кезінде құзыреттілік паспорттары жасалады. Білім беру бағдарламасын қалыптастыру кезінде құзыреттіліктерге сұраныс дәрежесі ескеріліп, кредиттік бірліктердің саны пәндерге олардың құзыреттіліктерді қалыптастыру үшін маңыздылық дәрежесіне сәйкес бөлінуі керек.

Пәннің мазмұнын анықтау, сабақтас пәндерді оқудың онтайлы тәртібін анықтау, белгілі бір курсты оқу үшін кафедраны таңдау, пәндер бойынша жұмыс бағдарламасын дайындау кезінде жұмыс берушілердің талаптарын есепке алу сияқты процедуралар өзара байланысты, және оларды реттеу қажет.

Мұның бәрі құжат айналымы жүйесінің көлемінің ұлғаюымен және күрделенуімен тікелей байланысты, бұл кафедра мен бөлім басшыларына қосымша жүктеме әкеледі.

Қазіргі уақытта барлық дерлік жоғары оқу орындарында автоматтандырылған ақпараттық жүйелер бар, онсыз көптеген бөлімдердің тиімді және жедел жұмысы өте қиын, кейде мүмкін емес. Бірақ қазіргі уақытта өзгеретін талаптарға сәйкес пәндер бойынша жұмыс бағдарламаларын жаңартуға және құруға байланысты оқытушылардың ағымдағы жұмысын жеңілдететін жүйелер жоқ. Мұндай жүйелер бірыңғай көздерден қажетті ақпаратты жедел іздеуге, арнайы ұйымдастырылған дерекқорда оқылатын курстар бойынша тәжірибесін жинақтауға, пән бойынша жұмыс бағдарламасын құруға, жұмыс бағдарламаларын қалыптастыру кезінде қателіктерді азайтуға, сапа менеджменті жүйесінің талаптарына сәйкес кафедралардың ішкі аудитін жүргізуді жеңілдетуге мүмкіндік беруі керек.

Зерттеулерге шолу

Бүгінгі таңда интеллектуалдық онлайн платформаларды әзірлеу саласы қарқынды дамып келе жатқан салалардың бірі болып табылады, бірақ бұл саланың пайда болуымен платформаларды құрудағы жаңа бағыт пайда болады. Қазіргі заманғы білім беру саласындағы және оның жеке ерекшеліктерін ескере отырып, білім алушылардың барлық деңгейлері үшін білімге тең қолжетімділікті қамтамасыз етуді көздеуге тиіс. Интеллектуалдық онлайн платформаларды құрастыру кезінде осы салада осы уақытқа дейін жасалған еңбектер ескеріледі.

Мақалада талдау интеллектуалды оқыту жүйелеріне негізделген соның ішінде: сипаттамаларға, қолданбаларға және бағалау әдістеріне жүйелі шолу жасалынды (Mousavinasab et al., 2018), білім беру жүйесін құру үшін жасанды интеллект әдістеріне сәйкес жүргізілді.

COVID-19 пандемиясы жағдайында жасанды интеллект білім беру саласында қолданылатын үлкен мәртебеге ие болып отыр. Сондықтан білім беру жүйелері оқытушылар құрамын алмастыруға үміткер емес, бірақ олардың оқытушылық қызметін жеңілдететін интеллектуалды оқыту технологияларын қолданады деп мәлімдейді (García-Peñalvo et al., 2020). Бұл арнайы шығарылым жасанды интеллектте, үлкен деректерге, машиналық және терең оқытуға негізделген білім беру қолданбалары мен қызметтеріндегі түпнұсқа жетістіктер туралы мақалалар жинағын ұсынуға арналған.

Мақалада белсенді оқытуды қолдауға арналған интеллектуалды тәлімгерлік жүйесін ұсынып отыр (Castro-Schez et al., 2021). студенттерге Блум таксономиясының жоғарғы когнитивті деңгейінде жұмыс істей отырып, оларды белсенді оқуға және өзін-өзі қамтамасыз етуге шақыра отырып, практикалық тұрғыдан үйренуге мүмкіндік беретін интеллектуалды тәлімгерлік жүйесін жасады.

Қосымшаны іздеу алгоритмдеріне қатысты интерактивті мысалдар мен жаттығулар үшін қолданды (Grivokostopoulou et al., 2016).

Интеллект теориясына негізделген оқушылардың интеллект деңгейін бағалау әдістемесін қолданды (Kose et al., 2016).

Динамикалық тест генерациясының әдістері мен алгоритмдерін зерттеп, анықтады нейрондық желілерге негізделген әдістерді қолдана отырып, студенттерді талдауға негізделген. Олардың пікірінше, бұл әдістер жеке білім деңгейіне негізделген бақылау нұсқаларын бейімдеуде икемді тәсілді қамтамасыз етуге көмектеседі. Нәтижесінде, бұл процесс оқытушыға курстың зерттелген бөлімдері арқылы студенттің білімін неғұрлым репрезентативті бағалауға мүмкіндік береді (Jiang et al., 2020).

Нәтижеге негізделген білім дегенімізді - бұл оқушылардың үлгерімін нәтижелерден бағалауға баса назар аударатын оқыту моделі, ал нәтижелер білімнен, шеберліктен және мінез-құлықтан тұратынын анықтады (Miguéis et al., 2018).

Еуропада құзыреттілікке негізделген білім беру процесі өсуде. Бұл көптеген юрисдикцияларға негізделген құзыреттердің Еуропалық жүйесін енгізумен ерекшеленеді. Бұл мақалада құзыреттілікке негізделген білім берудің атрибуттары, оның салалары және құзыреттілікке негізделген тиімді оқыту мен оқытуға арналған оқыту тәсілдері қарастырылады (O'Sullivan et al., 2014).

Компьютерлік дағдылар принциптері 2000 жылдардың басынан бастап бастауыш және орта білімге енгізілді, ал оқу бағдарламаларын реформалау ақпараттық технологиялардың дамуымен байланысты. Сондықтан оқу бағдарламалары қоғамның соңғы технологиялық тенденциялары мен қажеттіліктерін көрсетуі керек. Оқу бағдарламасын әзірлеу бірнеше ұқсас құжаттардан білім алу үшін бірнеше сарапшылардың немесе профессорлардың субъективті пікірін қамтиды. Неғұрлым объективті экстракция стандартталған терминологияға негізделуі керек, ал кәсіби терминология оқу бағдарламаларын ұйымдастыруға арналған мазмұн шеңберін құруға көмектеседі. Бұл зерттеудің мақсаты-информатика саласындағы білім жиынтығынан терминдерді алу және білім салаларын ұйымдастыру үшін интеллектуалды жүйені дамыту. Алынған терминдер word2vec моделін қолдана отырып, мағыналық жағынан ұқсас білім салаларынан тұрады (Woo et al., 2018).

Жұмыстың білім беру жүйесі мен еңбек нарығының ресми және бейресми талаптарына кәсіби дайындықтың сәйкестігін қамтамасыз ету қажеттілігі өзектілігімен расталады. Университет түлектерінің еңбек нарығында жұмысқа орналасуының себебі ақпараттық технологиялар бойынша оқу бағдарламасының нәтижелерін еңбек нарығының талаптарымен үйлестіре алатын әдістер мен технологиялардың болмауы болып табылады (Kurzaeva et al., 2020).

Профессор студенттерді кәсіби оқыту барысында бағыттайтын және бағыттайтын тәлімгер ретінде танымал, бұл құзыреттілікке негізделген модельде берілген. Демек, оқытушы студенттердің қандай курстарды мақұлдағанын және келесі кезеңде қандай курстардан өтуі керек екенін анықтауы керек, сондықтан олардың академиялық дайындығы жалғасады. Оқу

жоспарлары курстардан (түйіндер түрінде ұсынылған), курстың алғышарттары арасындағы байланыстардан (жиектермен көрсетілген), сондай-ақ семестр нөмірінен (түйіннің салмағымен ұсынылған) тұратын күрделі құрылымды көрсететін етіп модельденуі керек. Мақалада студенттер үйренуі және тапсыруы керек пәндер ағынындағы ішкі және сымды шектеулерді білдіретін екі бағытты өлшенген график берілген, олар ұйымдастырылған контекст (әр пәннің салмағымен реттеледі) құра алады. нұсқаушы басшылыққа алады және бағыттайды. Студенттердің үлгерімін бақылау және студенттер мен оқытушылар арасындағы өзара әрекеттесуді қамтамасыз ету жүйесі ұсынылады, оған екі бағытты өлшенген кесте біріктірілген. Сонымен, бұл жұмыс графиктерді жасанды интеллектпен біріктіреді (Galvez et al., 2020).

Сондай-ақ, мақалада біз білім берудің барлық деңгейлерінде информатика бойынша оқу бағдарламаларын үйлестіруге арналған жартылай автоматты бағдарламалық платформа сипатталғанын көре аламыз. Сербия Республикасының информатика мұғалімдері үшін таңдалған оқу бағдарламасының моделі бағдарламалық платформаны қолдана отырып, информатика мұғалімдеріне арналған анықтамалық оқу бағдарламасының моделімен жасалды және салыстырылды. Нәтижелерді талдау бойынша ол құрылған онтологиялық модельдердің басқа мүмкін жұптары үшін алынған мәліметтермен салыстыруды қамтиды (АСМ К12 орта мектеп моделі және анықтамалық модель, орта мектеп моделі және таңдалған оқу бағдарламасының моделі). Осы мақалада келтірілген эксперименттер мұғалімдерге арналған оқу бағдарламасын жақсарту, сондай-ақ таңдаудың жаңа әдістерін қолдану туралы ойлану керек екенін көрсетеді (Mandić, 2018).

Бұл мақалада бүгінде бүкіл әлемде көптеген онлайн МООС (Massive Open Online Courses) курстары бар екендігі айтылады. Тіркелген МООС студенттерінің саны ұсынылған онлайн курстардың санымен бірге тез өсті. Көптеген университеттер екі проблемаға тап болады. 1) қолданылатын платформалардың гетерогенділігіне байланысты ресми және бейресми жүйе арасындағы байланысты енгізе алмау. 2) студенттің бейінін зерттеу үшін бейресми оқытуда студенттердің пікірлерін қолдану мүмкіндігі жоқ. Өйткені, білім беру платформаларында әртүрлі форматтарда жинақталған көптеген мәліметтер бар. Барлық деректерге қол жеткізу және шолу үшін оларды жинап, бір берік жүйеге біріктіру керек (Mrhag et al., 2020).

Алайда, келісімшарттық алаяқтық, студенттер бағалау функцияларын үшінші тараптарға беретін академиялық теріс қылықтың бір түрі, осылайша мұғалімдерді алаңдатады. Анықтау жүйелері тиімсіз болғандықтан және келісімшарттық алаяқтықты анықтаған кезде кейбір мектептер қатты сынға ұшыраған, бірақ оқушылар мен мұғалімдер ұйымдарына қымбатқа түсетін бақылау жүйелеріне жүгінеді. Осындай жүйелердің бірі, Deep Speaker-екі аудио үлгінің бір адамның дауысына ұқсайтынын жоғары дәлдікпен тексере алатын спикерді анықтау және тексеру жүйесі. Бұл мақалада Real Talk онлайн

ауызша бағалау құралын Deer Speaker-пен біріктіретін инновациялық құрал ұсынылған. Бұл ұсынылған жүйе студенттер мен оқытушылар арасындағы ұзақ пікірталастарды жеңілдетеді, сонымен бірге оқушының жеке басын минималды шығындармен және білім беру мекемелеріне әсер етумен ұзақ мерзімді тексеруді қамтамасыз етеді (Renzella et al., 2022).

Білім қорында онтология көптеген қосымшалар үшін қажетті ресурстар болып табылады. Онтологияны зерттеу кейбір тапсырмаларды автоматтандыру арқылы онтологияны қолмен құру жұмысын жеңілдетуге көмектеседі. Бұл мақалада осы саладағы заманауи жетістіктер келтірілген. Тәсілдердің әртүрлі кластары (лингвистика, статистика және машиналық оқыту), соның ішінде кейбір соңғы тәсілдер (терең оқытуға негізделген тәсілдер) қарастырылады. Сонымен қатар, онтологияны зерттеудің стратегиялары мен ендірілген әдістерін ұсынатын бірнеше байланысты шешімдер (шеңберлер) ұсынылған. Құрылған онтология компоненттерімен және автоматтандыру дәрежесімен байланысты критерийлер негізінде сипаттамалық жиынтық беріледі. Бұл сонымен қатар онтологияның сенімділігін бағалаудың күрделілігін көрсетеді. Себебі бұл салада оңай жұмыс емес. Бұл іс жүзінде тәуелсіз зерттеу саласын білдіреді (Ahlem et al., 2021).

Университеттер мен жоғары оқу орындары білімдерін үнемі жаңартып отырады. Сондықтан сіз әрқашан құрылған Академиялық және әкімшілік ақпаратты жазып алуыңыз керек. Жоғары оқу орындары басқаратын ақпараттың көптігін және бір мекемеде қатар өмір сүре алатын жүйелердің әртүрлілігін ескере отырып. Осыған байланысты білімді бөлісу үшін технологияны қолдану қажет. Онтологияның арқасында адамдар мен әртүрлі жүйелер арасында тиісті ақпарат алмасуды қамтамасыз ету арқылы білімге қол жеткізу әлдеқайда оңай. Бұл мақаланың мақсаты жоғары білім беруде онтологияны қолдану бойынша бар зерттеулерді анықтау болып табылады. Жүйелі картографиялық зерттеу оның 2792 мақаласының сериясы негізінде жүргізілді. Барлығы 52 ғылыми жұмыс қаралып, талданды. Нәтижелер онтологияның жоғары білім беруде қалай қолданылатынын, оларды әзірлеу үшін қандай технологиялар мен құралдар пайдаланылатынын және онтологиялық қолданбаларда қандай негізгі сөздіктер қайта пайдаланылатынын көрсетеді (Taria-Leon et al., 2018).

Оқыту тұрғысынан уақытты бақылау және ерте ескерту маңызды рөл атқарады, сонымен қатар қызығушылықты арттырудың тиімді құралы болып табылады. Бұл зерттеу кең және жан-жақты деректерге негізделген Уақытша бақылау мен бірнеше семантикалық функцияларды талдайды және дамытады. Біріншіден, семантикалық белгілер мен қатынастарды анықтайтын факторлар зерттеледі. Екіншіден, бірнеше семантикалық конволюциялармен (iCaNN) желіні болжау моделі оқыту мінез-құлқының күрделілігі мен ерекшелігіне негізделген. Оңтайлы эксперименттік конфигурация модельді кеңінен оқыту және тестілеу арқылы көрсетіледі. Үшіншіден, біз көп нәрсені елестетеміз семантикалық функциялар, бар мәселелерді талдаймыз және оқытудың негізгі

ережелері мен уақыт тәуекелдерін зерттейді (Xiaona et al., 2022).

Әдебиеттің теориялық талдауы мәселенің кеңінен қарастырылғанын көрсетеді. Сонымен қатар, интеллектуалдық онлайн платформалардың ерекшеліктерін ескере отырып, білім беру процесіне енгізуге мүмкіндік беретін оқытудың осындай формалары мен әдістерін қолдануды қамтиды. Бұл процесте интеллектуалдық платформалардағы талаптарға қойылатын электрондық білім беру ресурстары маңызды рөл атқара алады.

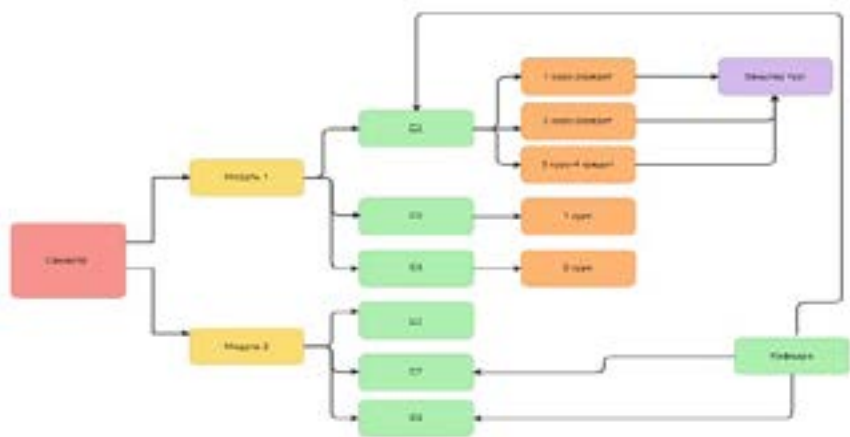
Білім беру бағдарламасының интеллектуалды конструкторы

Білім беру бағдарламасының интеллектуалды конструкторы іс жүзінде білім беру мекемелерінің басшылары үшін шешім қабылдауды қолдау жүйесі болып табылады және оқу орны басшысының ғана емес, штаттық қызметкерлердің де жұмысын едәуір жеңілдетуге, сондай — ақ білім беру процесінің сапасын арттыруға арналған. Университеттің интеллектуалды онлайн платформа жасаушылар ең алдымен университеттің негізгі бизнес — функциясын-оқу процесін автоматтандыру қажеттіліктерін ескереді. Атап айтқанда, программа әзірлеушілер студенттердің жеке оқу бағдарламасын (таңдау пәндерінің блоктары) қалыптастыру мүмкіндіктерін, оқу ағындарын қалыптастыру, сабақ кестесін қалыптастыру, сабақ өткізуді есепке алу, бақылау шараларын есепке алу, оқу үлгерімі мен оқу жүктемесін есепке алу сияқты функцияларды орындау кезінде жеке бағдарламаны есепке алады.

Аталған мүмкіндіктерге жетудің әдісі ретінде төмендегі міндеттер іске асырылуы қажет:

- Концептуалдық модельді қолданып пәндердің жіктелуін ұсыну;
- Реляциялық модельді қолдану;
- protégé 5.5.0 редакторында жүзеге асырылған онтологиясы.

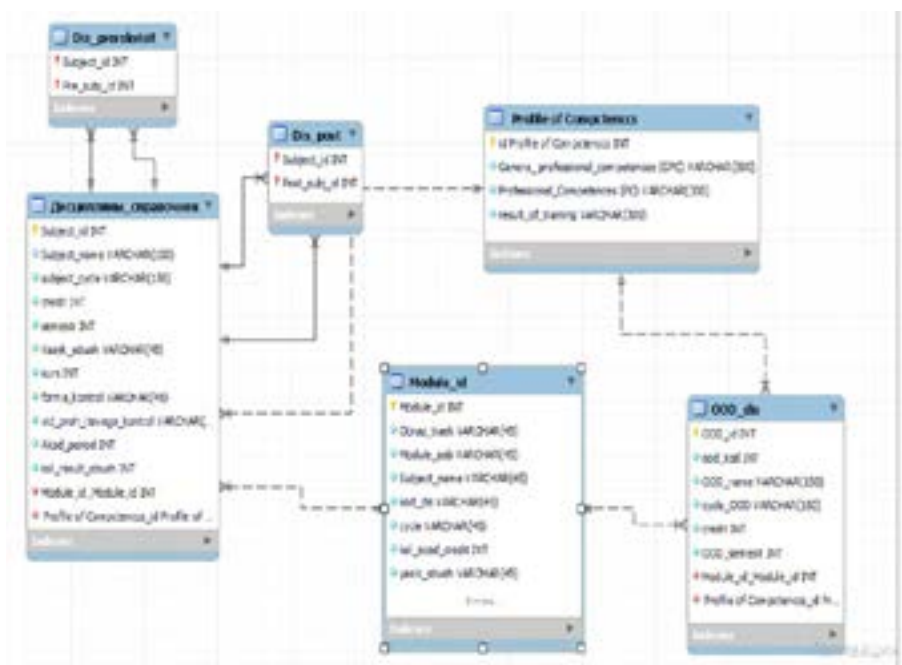
Жоғары келтірілген мысалдарға сүйеніп қолданылған схеманың Manu-to-Manu қатынасына сәйкестендіруге болады, оны келесі схемалардан көруге (сурет 1).



Сурет 1- Концептуалдық модельді қолданып пәндердің жіктелуі

Мысалы, аталған схема бойынша С-пәндерді білдіреді, бір пән бірнеше семестрге бөлінуі мүмкін және де олардың бақылау формаларына байланысты әртүрлі болады. Сонымен қатар, қамтылатын пәндерді әртүрлі кафедралар жүргізуі мүмкін, сәйкесінше жүйеге кіріп пәнді таңдағанда жүргізетін кафедраны да таңдауға болады.

Жүйеде білім беру бағдарламасы конструкторының материалдарын сақтау механизмдерін, оларды сипаттау жолдарын, алмасу мен ізденуді ұйымдастыру қажет. Енгізілген барлық деректер нақты құрылымдалған болуы керек. Деректерді жинақтау модулі біртұтас болуы керек және Интернетте кең таралған деректерді сақтаудың басқа дерекқорларынан деректерді енгізуді қолдау ұйымдастырылуы керек. Метадеректерді пайдалана отырып, деректерді модернизациялаудың кейінгі процесін ойластыру қажет. Мысалы онлайн оқыту платформасының деректер құрылымы ретінде ER-диаграммасын ұсынуға болады (сурет 2).



Сурет 2 - Конструктордың ER-диаграммасы

Міндетті пәндер студенттің оқуы үшін қатаң белгіленген немесе міндетті уақыт болуы мүмкін, (яғни, студент тіркеу кезінде пәндерді өз бетінше оқиды). Міндетті пәндердің пререквизиттері мен постреквизиттері тек міндетті пәндер болуы мүмкін.

Міндетті пәндер және элективті пәндердің пререквизиттері мен постреквизиттері міндетті және элективті пәндер болуы мүмкін. Оған мысал ретінде келесі кестелерді ұсынуға болады (кесте 1,2).

Кесте 1. Пәндердің ДБ берілу мысалы

Subject_id	Subject_name	Subject_cycle	Module_id
...
123	Сигналдарды цифрлық өңдеу	БП, ЖООК	3
...
511	Физика	ЖББП	1
525	Математика	ЖББП	2

Кесте 2. Пән-пререквизит

Subject_id	Pre_subject_id
123	511
123	525
және т.б.	және т.б.

Келтірілген мысалда бірнеше пререквизит-пәндерді енгізуге болады.

Жұмыстың практикалық нәтижесі protégé 5.5.0 редакторында жүзеге асырылған онтологиясы, ол оқытылатын пәндерді олар қалыптастыратын дағдылар мен студенттің біліктілігіне қойылатын кіріс талаптары (prerequisites) бөлігінде көрсетеді.

Онтологиялық модельдеу негізінде көрсетілген Protégé редакторіндегі JavaScript_programming пәні үшін пререквизиттері 3 - сурет.



Сурет 2. Protégé редакторіндегі JavaScript_programming пәні үшін пререквизиттері

Міндетті пәндерді енгізу үшін бірнеше критерилерді қарастыруға болады: Міндетті пәндер элективті пәндерді қамтымауы керек;

Міндетті пәндердің пререквизиттері мен постреквизиттері міндетті пәндер болуы керек.

Элективті пәндерді енгізуге қолданылатын критерий:

Элективті пәндердің пререквизиттері мен постреквизиттері міндетті және элективті пәндер болуы мүмкін.

Нәтижелер және талқылау

Мақалада ЖОО (жоғарғы оқу орны)-ның барлық мамандықтары үшін білім беру бағдарламасын құру үшін пайдаланылуы мүмкін автоматтандырылған ББ бағдарламасын (білім беру бағдарламасын) конструкторын әзірлеу сипатталған. Мақалада зерттеу нәтижелерін сақтайтын және автоматты ББ бағдарламасының конструкторын құрудың алғышарты ретінде пайдалануға болатын мәліметтер құрылымы келтірілген. Деректер құрылымы пәндердің пререквизиттері мен постреквизиттерін таңдаудың дұрыстығын ескереді.

Автоматтандырылған білім беру бағдарламасын жүйесін әзірлеу қазіргі білім берудегі өзекті міндет болып табылады, өйткені ол білім беру процесін тиімді ұйымдастыруға және студенттердің білім мен дағдыларға деген қажеттіліктерін ескеруге мүмкіндік береді. Жоғары оқу орындарының барлық мамандықтары үшін ББ бағдарламасының конструкторын құру осы бағыттағы маңызды қадам болып табылады.

Білім беру бағдарламасын автоматтандырылған жүйесі пәндердің пререквизиттері мен постреквизиттерін таңдаудың дұрыстығын ғана емес, сонымен қатар студенттердің дайындық деңгейі, бағдарламаның еңбек нарығының талаптарына сәйкестігі және т.б. сияқты басқа факторларды да ескеруі керек.

Мақалада келтірілген зерттеу нәтижелері мен деректер құрылымы автоматтандырылған ББ бағдарламасының конструкторын әзірлеу үшін негіз ретінде пайдаланылуы мүмкін. Оқу процесінің тиімділігін арттыру үшін осы бағытта жұмысты жалғастыру және автоматтандырылған білім беру жүйелерін жетілдіру маңызды.

Қорытынды

Білім беру бағдарламалары үшін автоматтандырылған жүйені әзірлеу білім беру саласына үлкен пайда әкелуі мүмкін. Ұсынылған мәліметтер құрылымы жоғары оқу орындарындағы барлық мамандықтар үшін осындай жүйені құруға негіз бола алады.

Жүйенің басты артықшылықтарының бірі-пәндердің пререквизиттері мен постреквизиттерін дұрыс таңдауды ескеру мүмкіндігі. Бұл білім беру бағдарламасының дұрыс құрылымына және студенттердің табысты оқуы үшін қажетті білім мен дағдыларға ие болуына кепілдік береді. Бұл процесті автоматтандыру арқылы оқу орындары бағдарламаны қолмен құруға және қайта қарауға жұмсалатын уақыт пен ресурстарды үнемдей алады.

Сонымен қатар, ұсынылған деректер құрылымы жүйенің шамамен нұсқасы екенін атап өткен жөн. Толық жұмыс істейтін автоматтандырылған білім беру бағдарламасының конструкторын құру үшін қосымша зерттеулер мен әзірлемелер қажет болуы мүмкін. Білім беру саласындағы өзгерістерді

көрсету және білім беру бағдарламаларының өзектілігі мен өзектілігін қамтамасыз ету үшін жүйе үнемі жаңартылып, қайта қаралуы керек.

Білім беру бағдарламаларының автоматтандырылған жүйесі үшін ұсынылған деректер құрылымы білім беру саласына үлкен пайда әкелетінін атап өткен жөн. Пәндердің пререквизиттері мен постреквизиттерін таңдаудың дұрыстығын ескере отырып, бұл жүйе жоғары оқу орындарындағы барлық мамандықтар үшін білім беру бағдарламаларын тиімді және дәл құра алады.

ӘДЕБИЕТТЕР

Мусавинасаб Э., Зарифсаная Н., Р. Ниакан Калхори С., Рахшан М., Кейха Л. және Гази Саиди М., 2018. Интеллектуалды оқыту жүйелері: сипаттамаларға, қолданбаларға және бағалау әдістеріне жүйелі шолу. Интерактивті оқыту ортасы. <https://doi.org/10.1080/10494820.2018.1558257>. (Eng.).

Гарсия-Пенальво, Франциско және Касадо-Лумбрерас, Кристина және Коломо-Паласиос, Рикардо және Ядав, Аман, 2020. Ақылға қонымды оқыту. Қолданбалы ғылымдар. 10. 6964. 10.3390/app10196964. (Eng.).

Дж.Дж. Кастро-Шез, К.Глез-Морцилло, Дж. Альбусак, Д. Вальехо, 2021. Белсенді оқытуды қолдауға арналған интеллектуалды оқыту жүйесі: болжамды талдауды оқытуға арналған кейс-стади, ақпараттық ғылымдар, 544 том, 446–468 беттер, ISSN 0020-0255, <https://doi.org/10.1016/j.ins.2020.08.079>. (Eng.).

Гривокостопулу и Перикос және И.Хатзилигерудис, 2016. "Іздеу алгоритмдерін зерттеуге арналған Виртуалды шындық пен геймификацияға негізделген инновациялық білім беру ортасы", 2016 IEEE Сегізінші халықаралық білім беру технологиялары конференциясы (Т4Е), Мумбай, 110–115 б., doi:10.1109/T4E.2016.029. (Eng.).

Коуз В. және Арслан А., 2017. Компьютерлік инженерия курстарында өзін-өзі оқытуды оңтайландыру: интеллектуалды бағдарламалық жасақтама жүйесі-жасанды нейрондық желі және құйынды оңтайландыру алгоритмі қолдайды. Eng Education есептеу қосымшасы, 25: 142–156. <https://doi.org/10.1002/cae.21787>. (Eng.).

Цзян Дж., Ван Х.Х. сөйлеуді тану технологиясына негізделген интеллектуалды іздеу және ұсыныс жүйесін қолдану. Ішкі сөйлеу технологиясы, 2020. <https://doi.org/10.1007/s10772-020-09703-0>. (Eng.).

В.Л. Мигейс, Ана Фрейтас, Пауло Х.В. Гарсия, Андре Силва, 2018. Оқушылардың үлгеріміне сәйкес ерте сегменттеу: болжамды модельдеу тәсілі, шешімдерді қолдау жүйелері, 115 Том, 36–51 беттер, ISSN 0167–9236, <https://doi.org/10.1016/j.dss.2018.09.001>. (Eng.).

О' Салливан, Нил және Брюс, Алан, 2014. Құзыреттілікке негізделген білім беру жүйесінде оқыту және оқыту. (Eng.).

Ву Х.С., Ким Дж. М., Ли В.Г., 2020. сөздерді ендіруге және терминологияны шығаруға негізделген оқу бағдарламаларын жобалауды қолдау жүйесін әзірлеу //Электроника. - Т.9. – №– Б. 608. (Eng.).

Курзаева Л. және Повитухин С. және Усатая Т. және Усатый Д., 2020. Ақпараттық технологиялар саласындағы ЖОО түлектерінің бәсекеге қабілеттілігін арттырудың онтологиялық моделін әзірлеу. Физика журналы: конференциялар сериясы. 1691. 012003. 10.1088/1742-6596/1691/1/012003. (Eng.).

Гальвес Луз және Анзурес-Гарсия, Марио және Грегорио, Альваро, 2020. Оқытушының құзыреттілігін қолдау үшін кестеге негізделген академиялық оқу жоспарларының өлшенген екі бағытты моделі. Есептеу және жүйелер. 24. 10.13053/cys-24-2-3397. (Eng.).

Мандич Милинко 2018. Оқу бағдарламаларын үйлестіруге арналған семантикалық вебке негізделген бағдарламалық платформа. WIMS ' 18: веб-интеллект, деректерді өндіру және семантика бойынша 8 - ші халықаралық конференция материалдары. 1–9. 10.1145/3227609.3227654. (Eng.).

- Мрхар, Хаула және Дуими, Отмане және Абик, Муња және Чауни бенабделлах, Науал. (2020). Оқыту платформаларынан деректерді семантикалық интеграциялау жолында. *Iaes халықаралық жасанды интеллект журналы (IJ-AI)*. 9. 535. 10.11591/ijai.v9.i3.pp535-544. (Eng.).
- Джейк Ренцелла, Эндру Кейн, Жан-Ги Шнайдер, 2022. Deep speaker көмегімен ауызша бағалау кезінде оқушының жеке басын тексеру, Компьютерлер және білім: жасанды интеллект, 3 Том, 100044. ISSN 2666–920X, <https://doi.org/10.1016/j.caeai.2021.100044>. (Eng.).
- Шерифтің Ахлемі Хадир, Хассина Алиан, Ахмед Гессум, 2021. Онтологияны зерттеу: үлкен тур және қиындықтар, информатикаға шолу, 39 том, 100339. ISSN 1574–0137, <https://doi.org/10.1016/j.cosrev.2020.100339>. (Eng.).
- М. Тапия-Леон, А.С. Ривера, Х. Чикаиза және С. Лужан-Мора, 2018. "Жоғары білім берудегі онтологияны қолдану: жүйелі картографиялық зерттеу", 2018 IEEE инженерлік білім беру жөніндегі жаһандық конференция (EDUCON), 1344–1353 б., doi: 10.1109/EDUCON.2018.8363385. (Eng.).
- Сяона Ся, Вансуэ Ци, 2022. Оқушының мінез-құлқының мультисемантикалық ерекшеліктері туралы уақытша қадағалау және ерте ескерту, Компьютерлер және білім: жасанды интеллект, 3 Том, 100045. ISSN 2666–920X, <https://doi.org/10.1016/j.caeai.2021.100045>. (Eng.).

REFERENCES

- Mousavinasab E., Zarifsanaiy N., R. Niakan Kalhori S., Rakhshan M., Keikha L. & Ghazi Saeedi M., 2018. Intelligent Learning Systems: A systematic review of characteristics, applications and assessment methods. *Interactive Learning Environment*. <https://doi.org/10.1080/10494820.2018.1558257>. (in Eng.).
- García-Peñalvo, Francisco & Casado-Lumbreras, Cristina & Colomo-Palacios, Ricardo & Yadav, Aman, 2020. Smart Learning. *Applied Sciences*. 10. 6964. 10.3390/app10196964. (in Eng.).
- J.J. Castro-Schez, C. Glez-Morcillo, J. Albusac, D. Vallejo, 2021. An intelligent tutoring system for supporting active learning: A case study on predictive parsing learning, *Information Sciences*, Volume 544, Pages 446-468, ISSN 0020-0255, <https://doi.org/10.1016/j.ins.2020.08.079>. (in Eng.).
- Grivokostopoulou, I. Perikos and I. Hatzilygeroudis. (2016). "Innovative educational environment based on virtual reality and gamification for learning search algorithms", 2016 IEEE Eighth International Conference on Technology for Education (T4E), Mumbai. Pp.110–115, doi:10.1109/T4E.2016.029. (in Eng.).
- Kose U. and Arslan A., 2017. Optimization of self-learning in Computer Engineering courses: An intelligent software system supported by Artificial Neural Network and Vortex Optimization Algorithm. *Comput Appl Eng Educ*, 25. Pp. 142–156. <https://doi.org/10.1002/cae.21787>. (in Eng.).
- Jiang J., Wang H.H. применение интеллектуальной системы поиска и рекомендаций, основанной на технологии распознавания речи. *Int J Speech Technol*, 2020. <https://doi.org/10.1007/s10772-020-09703-0>. (in Eng.).
- V.L. Miguéis, Ana Freitas, Paulo J.V. Garcia, André Silva, 2018. Early segmentation of students according to their academic performance: A predictive modelling approach, *Decision Support Systems*. Volume 115. Pp. 36–51. ISSN 0167–9236, <https://doi.org/10.1016/j.dss.2018.09.001>. (in Eng.).
- O'Sullivan, Neil & Bruce, Alan, 2014. *Teaching and Learning in Competency Based Education*. (in Eng.).
- Woo H.S., Kim J.M., Lee W.G., 2020. Development of curriculum design support system based on word embedding and terminology extraction // *Electronics*. – Т. 9. – №. 4. – p. 608. (in Eng.).
- Kurzaeva L. & Povitukhin S. & Usataya T. & Usaty D., (2020). The development of ontological model for increasing the competitiveness of university graduates in information technologies. *Journal of Physics: Conference Series*. 1691. 012003. 10.1088/1742-6596/1691/1/012003. (in Eng.).
- Galvez, Luz & Anzures-García, Mario & Gregorio, Álvaro, 2020. Weighted Bidirectional Graph-based Academic Curricula Model to Support the Tutorial Competence. *Computación y Sistemas*. 24. 10.13053/cys-24-2-3397. (in Eng.).
- Mandić, Milinko, 2018. Semantic Web based software platform for curriculum harmonization.

WIMS '18: Proceedings of the 8th International Conference on Web Intelligence, Mining and Semantics. Pp. 1–9. 10.1145/3227609.3227654. (in Eng.).

Mrhar, Khaoula & Douimi, Otmame & Abik, Mounia & chaouni benabdellah, Naoual. (2020). Towards a semantic integration of data from learning platforms. IAES International Journal of Artificial Intelligence (IJ-AI). 9. 535. 10.11591/ijai.v9.i3.Pp 535–544. (in Eng.).

Jake Renzella, Andrew Cain, Jean-Guy Schneider, 2022. Verifying student identity in oral assessments with deep speaker, Computers and Education: Artificial Intelligence. Volume 3, 100044. ISSN 2666–920X, <https://doi.org/10.1016/j.caeai.2021.100044>. (in Eng.).

Ahlem Chérifa Khadir, Hassina Aliane, Ahmed Guessoum, 2021. Ontology learning: Grand tour and challenges, Computer Science Review. Volume 39, 100339. ISSN 1574–0137, <https://doi.org/10.1016/j.cosrev.2020.100339>. (in Eng.).

M. Tapia-Leon, A.C. Rivera, J. Chicaiza and S. Luján-Mora, 2018. "Application of ontologies in higher education: A systematic mapping study," 2018 IEEE Global Engineering Education Conference (EDUCON). Pp. 1344–1353, doi: 10.1109/EDUCON.2018.8363385. (in Eng.).

Xiaona Xia, Wanxue Qi, 2022. Temporal tracking and early warning of multi semantic features of learning behavior, Computers and Education: Artificial Intelligence. Volume 3. 100045. ISSN 2666–920X, <https://doi.org/10.1016/j.caeai.2021.100045>. (in Eng.).

NEWS OF THE NATIONAL ACADEMY OF SCIENCES OF THE REPUBLIC OF KAZAKHSTAN
PHYSICO-MATHEMATICAL SERIES

ISSN 1991-346X

Volume 2. Number 346 (2023). 237–251

<https://doi.org/10.32014/2023.2518-1726.196>

УДК 004.032.26

МРПТИ 28.23.37

©**A.B. Toktarova**^{1*}, **B.S. Omarov**¹, **Zh.Zh. Azhibekova**³, **G.I. Beissenova**⁴,
R.B. Abdrakhmanov⁵, 2023

¹Khoja Akhmet Yassawi International Kazakh-Turkish University,
Turkistan, Kazakhstan;

²AL-Farabi Kazakh National University, Almaty, Kazakhstan;

³Asfendiyarov Kazakh National Medical University, Almaty, Kazakhstan;

⁴Mukhtar Auezov South Kazakhstan University, Shymkent, Kazakhstan;

⁵International University of Tourism and Hospitality, Turkestan, Kazakhstan.

E-mail: toktar.aigerim@list.ru

ANALYSIS OF HATE SPEECH WORDS IN ONLINE CONTENT BY USING DATA MINING

Toktarova Aigerim — doctoral student, department of Computer engineering, faculty of Engineering, Khoja Akhmet Yassawi International Kazakh – Turkish University, Turkistan, Kazakhstan
E-mail: toktar.aigerim@list.ru, <https://orcid.org/0000-0002-6265-9236>;

Omarov Batyrkhan — PhD, department of Information System, faculty of Information technologies AL-Farabi Kazakh National University, Almaty, Kazakhstan
E-mail: batyahan@gmail.com, <https://orcid.org/0000-0002-8341-7113>;

Azhibekova Zhanar — candidate of pedagogical sciences, department of Information Communication Technologies, Asfendiyarov Kazakh National Medical University, Almaty, Kazakhstan
E-mail: zhanar_azhibekova@mail.ru, <https://orcid.org/0000-0002-4396-1261>;

Beissenova Gulbakhram — candidate of pedagogical sciences, department of Computer science, Mukhtar Auezov South Kazakhstan University, Shymkent, Kazakhstan
E-mail: g_beissenova@mail.ru, <https://orcid.org/0000-0003-0545-0652>;

Abdrakhmanov Rustam — Cand. of Techn.Sc., Ass.Prof., International University of Tourism and Hospitality Turkestan, Kazakhstan
E-mail: abdrakhmanov.rustam@iuth.edu.kz, <https://orcid.org/0000-0003-0385-0684>.

Abstract. The state of ensuring the information and psychological security of society is declining due to the daily increase in destructive words on the Internet. That is, categories such as advocacy of violence based on race, ethnicity, national origin, gender, gender identity, religion, age, disability or illness are defined. It shows that the search and identification of non-speech words can be done with minimal effort from the point of view of the subject, since this can now be achieved automatically using various methods. The existing methods (expert analysis, intellectual data processing methods) are considered, the advantages and

disadvantages of these methods are noted. In addition, he not only explores the roots of profanity, but also proposes concepts to help distinguish between types of offensive language, such as profanity and cyberbullying. We use machine learning approaches to access datasets that we can use to automatically collect social media profanity. To solve this problem, Data Mining is proposed to search for offensive words, such as swear words in text databases. It can be seen that the presented method is distinguished by the ability to automatically fill in the vocabulary through the system when identifying unfamiliar words classified as non-standard. The correctness of vocabulary replenishment is measured using various indicators. We are creating a set of Internet users in Kazakhstan who use social networks and media to share their opinion. The description of the proposed research work is given from data collection to classification and identification of sets of words with destructive content.

Key words: hate speech, analysis, detection, online content, Data Mining

© **А.Б. Тоқтарова^{1*}, Б.С. Омаров², Ж.Ж. Ажибекова³, Г.И. Бейсенова⁴,
Р.Б. Абдрахманов⁵, 2023**

¹Қожа Ахмет Ясауи атындағы Халықаралық Қазақ-Түрік Университеті,
Түркістан, Қазақстан;

²Әл-Фараби атындағы Қазақ Ұлттық университеті Алматы, Қазақстан;

³С.Ж. Асфендияров атындағы Қазақ ұлттық медицина университеті,
Алматы, Қазақстан;

⁴М. Ауезов атындағы Оңтүстік Қазақстан Университеті,
Шымкент, Қазақстан;

⁵Халықаралық туризм және меймандостық университеті.
E-mail: toktar.aigerim@list.ru

ОНЛАЙН КОНТЕНТТЕГІ БЕЙӘДЕП СӨЗДЕР МӘЛІМЕТТЕР ҚОРЫН DATA MINING АРҚЫЛЫ АНАЛИЗДЕУ

Тоқтарова Айгерім — Докторант, кафедра Компьютерлік инженерия, Қожа Ахмет Ясауи атындағы Халықаралық Қазақ-Түрік Университеті, Түркістан, Қазақстан
E-mail: toktar.aigerim@list.ru, <https://orcid.org/0000-0002-6265-9236>;

Омаров Батырхан — Phd, кафедра Ақпараттық жүйелер, факультет Ақпараттық технологиялар, Әл-Фараби атындағы Қазақ Ұлттық университеті, 050040, Алматы, Қазақстан Алматы, Қазақстан,
E-mail: batyahan@gmail.com, <https://orcid.org/0000-0002-8341-7113>;

Ажибекова Жанар — п.ғ.к., С.Ж. Асфендияров атындағы Қазақ ұлттық медицина университеті Алматы, Қазақстан Республикасы
E-mail: zhanar_azhibekova@mail.ru, <https://orcid.org/0000-0002-4396-1261>;

Бейсенова Гүльбахрам — п.ғ.к., М. Ауезов атындағы Оңтүстік Қазақстан Университеті, Шымкент, Қазақстан
E-mail: g_beissenova@mail.ru, <https://orcid.org/0000-0003-0545-0652>;

Абдрахманов Рустам — техн.ғ.к., доцент м.а., Халықаралық туризм және меймандостық университеті, Түркістан, Қазақстан
E-mail: abdrakhmanov.rustam@iuth.edu.kz, <https://orcid.org/0000-0003-0385-0684>.

Аннотация. Қоғамның ақпараттық-психологиялық қауіпсіздігін қамтамасыз ету жағдайын интернеттегі деструктивті сөздердің күнделікті тоқтаусыз өсуі төмендетіп отыр. Яғни, нәсіліне, этникалық тегіне, ұлттық шығу тегіне, жынысына, гендерлік сәйкестігіне, дініне, жасына, мүгедектігіне немесе ауруына негізделген зорлық-зомбылықты насихаттау сияқты санаттары анықталып отыр. Ғадауат сөздерді іздеу және анықтау субъект тұрғысынан минимальді түрде күш жұмсауға болатынын көрсетіп отыр, себебі қазіргі таңда автоматты түрде түрлі әдістерді қолдану арқылы қол жеткізуге болады. қолданыстағы әдістер (эксперттік талдау, деректерді өндеудің интеллектуалды әдістері) қарастырылады, бұл әдістердің артықшылықтары мен кемшіліктері атап өтіледі. Сонымен қатар, бейәдеп тілдің түп-тамырын зерттеп қана қоймайды, кемсіту тіл мен кибербуллинг сияқты ғадауат сөздер түрлерін ажыратуға көмектесетін ұғымдарды ұсынады. Біз әлеуметтік желідегі балағат сөздерді автоматты түрде жинау үшін пайдалана алатын деректер жиынына қол жеткізу үшін машиналық оқыту тәсілдерін қолданамыз. Мұны шешу үшін мәтіндік мәліметтер қорындағы балағаттау сөздерді мысалға ала отырып, ғадауат мазмұнды сөздерді іздеу үшін Data Mining ұсынылады. Көрсетілген әдіс нормаға сай емес деп жіктелген бейтаныс сөздерді анықтауда жүйе арқылы сөздік қорды өздігінен толтыру мүмкіндігімен ерекшеленетінін байқауға болады. Әртүрлі көрсеткіштерді пайдалану арқылы сөздікті толықтырудың дұрысытығы өлшенеді. Қазақстандағы интернет пайдаланушылары деректер жинағын жасап жатырмыз, олар өз пікірлерімен бөлісу үшін әлеуметтік желілер мен медианы пайдаланады. Ұсынылып отырған зерттеу жұмысының сипаттамасы деректерді жинаудан бастап деструктивті мазмұндас сөздер топтамасын жіктеуге және оларды анықтауға дейін берілген.

Түйін сөздер: бейәдеп мазмұнды сөздер, анықтау, онлайн контент, анализдеу, Data Mining

© **А.Б. Токтарова^{1*}, Б.С. Омаров², Ж.Ж. Ажибекова³, Г.И. Бейсенова⁴,
Р.Б. Абдрахманов⁵, 2023**

¹Международный Казахско – Турецкий университет им. Ходжа Ахмета Ясауи, Туркестан, Казахстан;

³Казахский Национальный университет им. аль-Фараби, Алматы, Казахстан;

³Казахский Национальный медицинский университет имени С.Д. Асфендиярова, Алматы, Казахстан,

⁴Южно – Казахстанский университет имени Мухтара Ауезова, Шымкент, Казахстан;

⁵Международный университет туризма и гостеприимства.

E-mail: toktar.aigerim@list.ru

АНАЛИЗ НЕОБРАЗНЫХ СЛОВ В ОНЛАЙН-КОНТЕНТЕ С ПОМОЩЬЮ DATA MINING

Токтарова Айгерім — докторант, кафедра Компьютерная инженерия, факультет Инженерия, Международный Казахско-Турецкий университет им. Ходжа Ахмета Ясауи, Туркестан, Казахстан

E-mail: toktar.aigerim@list.ru, <https://orcid.org/0000-0002-6265-9236>;

Омаров Батырхан — Phd, кафедра Информационных систем, факультет Информационных технологии, Казахский Национальный университет им. аль-Фараби, Алматы, Казахстан

E-mail: batyahan@gmail.com, <https://orcid.org/0000-0002-8341-7113>;

Ажибекова Жанар — к.п.н., Казахский Национальный медицинский университет имени С.Д. Асфендиярова, Алматы, Республика Казахстан

E – mail: zhanar_ajhibekova@mail.ru, <https://orcid.org/0000-0002-4396-1261>;

Бейсенова Гульбахрам — к.п.н., Южно – Казахстанский университет имени Мухтара Ауезова, Шымкент, Казахстан

E – mail: g_beissenova@mail.ru, <https://orcid.org/0000-0003-0545-0652>;

Абдрахманов Рустам — к.т.н., и.о. доцента, Международный университет туризма и гостеприимства Туркестан, Казахстан

E – mail: abdrakhmanov.rustam@iuth.edu.kz, <https://orcid.org/0000-0003-0385-0684>.

Аннотация. Состояние обеспечения информационно-психологической безопасности общества снижается из-за ежедневного увеличения деструктивных слов в сети Интернет. То есть определены такие категории, как пропаганда насилия по признаку расы, этнической принадлежности, национального происхождения, пола, гендерной идентичности, религии, возраста, инвалидности или болезни. Он показывает, что поиски идентификация неречевых слов может осуществляться с минимальными усилиями с точки зрения испытуемого, так как теперь это может быть достигнуто автоматически с использованием различных методов. Рассмотрены существующие методы (экспертный анализ, интеллектуальные методы обработки данных), отмечены преимущества и недостатки этих методов. Кроме того, он не только исследует корни нецензурной лексики, но также предлагает концепции, помогающие различать типы оскорбительного языка, такие как ненормативная лексика и киберзапугивание. Мы используем подходы машинного обучения для доступа

к наборам данных, которые мы можем использовать для автоматического сбора ненормативной лексики в социальных сетях. Для решения этой проблемы предлагается Data Mining для поиска оскорбительных слов, например, матерных слов в текстовых базах данных. Видно, что представленный метод отличается возможностью автоматического заполнения словарного запаса через систему при выявлении незнакомых слов, отнесенных к нестандартным. Правильность пополнения словарного запаса измеряется с помощью различных показателей. Мы создаем набор пользователей Интернета в Казахстане, которые используют социальные сети и СМИ, чтобы поделиться своим мнением. Описание предлагаемой исследовательской работы дано от сбора данных до классификации и выявления наборов слов деструктивного содержания.

Ключевые слова: нецензурная речь, анализ, онлайн контент, определение, Data Mining

Кіріспе

Қазіргі таңдағы ғылым мен техниканың даму прогресі адам қызметтерінің біршама салаларына жаңа ақпараттық технологияларды үздіксіз енгізілуде. Сонымен қатар, интернеттің дамуы сол жердегі мәтін түрінде берілген ақпарат көлемі мен мазмұнының бақыланбайтын экспоненциалды өсуіне жеткізді (Канхото, 2020). Ақпараттық қауіпсіздікті сақтау мақсатында әлеуметтік желілеріндегі заңсыз ақпаратты, оның ішінде бөліп қарастырсақ терроризм, есірткінің заңсыз стабылымы, наразылық акцияларын немесе жаппай тәртіпсіздіктерді ұйымдастыру, мемлекеттік рәміздерді қорлау сөздері, балағат сөздерді және т.б. қоғамға қауіпті деректері бар мәліметтерді талдаудың маңызы бар (Кумар, 2020). Осыған байланысты, сондай-ақ «жабық» интернет кеңістігін әлеуетті құру жағдайында деструктивті ақпаратты іздеу және автоматты түрде анықтау. Сонымен қатар, бүгінгі күнге дейін қоғамның ақпараттық-психологиялық қауіпсіздігін сақтау үшін заң жүзіндегі іс-шаралар жеткіліксіз екендігін көруге болады (Джайн, 2020).

Сонымен, автоматтандырылған және мамандандырылған іздеу жүйелерін іске асыру және оларды санаттарына қарай жіктеу қажеттілігі туындайды. Осылайша, ақпараттық ресурстардың үнемі өсуі қауіпті мазмұнды сөздерді анықтауда автоматтандыру процесі еңбек шығындарын төмендетумен қатар, субъективті іс — әрекеттер желісін және адам факторының әсерінен келетін қауіптер ықтималдығын азайтады (Мохан, 2020). Әлеуметтік желілердің белсенді дамуы, интернетті «қылмыстық» қауіптерге ерекше назар аударуды талап етеді. Ақпаратты қорғаудың келесі жолдары бар (Мохан, 2020; Мюльхофф, 2020):

- қауіптің алдын алу — олардың туындау мүмкіндігін болжау және ақпараттық қауіпсіздікті қамтамасыз ететін алдын алу шаралары;
- қауіп-қатерлерді анықтау нақты немесе қауіптердің пайда болу

ықтималдығын жүйелі талдау мен бақылауда ұстау, сонымен қатар, олардың алдын алу бойынша дер кезінде шаралар қабылдау;

- қылмыстық әрекеттерді оқшаулау және қауіпті немесе нақты қылмыстық әрекеттерді жою шараларын қабылдау;

Шартты түрде ақпараттық — технологияларды пайдалануда сыртқы қауіптерді «дәстүрлі» және «жаңа» деп бөліп қарастыруға болады. (Муссиралиева, 2021) «Дәстүрлі» сыртқы қауіптер ретінде спам, фишинг, компьютерлік вирустар, трояндар және желілік шабуылдар жатқызуға болады. Келесі тарауда «жаңа» сыртқы қауіптер қарастырылады.

Әдістер мен материалдар

«Жаңа» ақпараттық – технологиялық қауіптер. Расында, интернет-коммуникация тақырыбы біздің уақытта өзекті болып отыр. Жоғары дәрежеде интернет-технологиялардың дамуы неге әкелді деген сұраққа біржақты жауап беру мүмкін болмай отыр. Әдетте халықаралық желі интернет миллиондаған компьютерлік желілердің ғаламдық байланысуы. Басқаша айтқанда, бұл дүние жүзіндегі барлық компьютерлік желілердің желісі немесе жай ғана желілер желісі. Интернет ғаламның кез келген нүктесінде бірден ақпарат алмасуға мүмкіндік бере алады. Көбіне, қысқа хабарламалар, жылдам хабар алмасу немесе чаттар үшін электрондық пошта арқылы интернет адамдарға тұлғааралық және бұқаралық коммуникация деңгейінде басқа адамдармен байланысуға мүмкіндік береді. Сондай-ақ ол өз пайдаланушыларына World Wide Web-те қолжетімді ақпарат көлеміне қол жеткізуді ұйымдастыра алады (Неупане, 2020). Кез келген энциклопедиядан да үлкен деректер әлемнің бір бөлігінен жер шарының ең алыс бөлігіне жарық жылдамдығымен тасымалдана алады. Осылайша, қолданушылар ақпаратты таңқаларлық жылдамдықпен жүктеп немесе жүктей алады. І суретте әлеуметтік желілерді пайдаланатын белсенді тұтынушылар саны (миллиондаған) көрсетілген (Саллум, 2020).



Сур 1. Әлеуметтік желілерді пайдаланатын белсенді тұтынушылар саны (миллион)
(Fig 1. Number of active users using social networks (millions))

Технологияға негізделген және сөз мазмұнына негізделген киберқылмыс деп киберқылмысты екі түрге бөліп қарастыруға болады. Сөз мазмұнына негізделген қылмысты жыныстық қудалау, киберқорқыту, балалар порнографиясы бойынша кез келген террористік ұйым жүзеге асырады. Ал, аккаунтты бұзып кіру, тыңшылық және зиянды кодты енгізу технологияға негізделген қылмыстың мысалдары болып табылады (Саркер, 2020).

Ransomware (бопсалаушы) бағдарламалары

Бұл бағдарламалар операциялық жүйенің жұмысын блоктайды және/немесе қатты дискідегі барлық файлдарды шифрлап тастайды (Сингх, 2020). Блоктау немесе шифрлау процесі аяқталғаннан кейін пайдаланушыдан құлыпты ашу / шифрды шешу кодын алу үшін белгілі бір соманы (әдетте 10–20 мың рубль аймағында) аударуды сұрайды. Файлдардың шифрын ашу мүмкіндігін растау ретінде шабуылдаушылар оларға 2–3 шифрланған файлды жіберуді ұсынады, содан кейін оларды бастапқы түрінде қайтарады. Қауіпті операциялық жүйе бұғатталған кезде оны жай ғана қайта орнатуға болады, бірақ барлық файлдарды шифрлаған жағдайда бұл көмектеспейді.

Ransomware (бопсалаушы) бағдарламалары қамтамасыз ету эпидемиясы кезінде (2012–2013) антивирустық компаниялар блоктан шығару қызметтерін ұсынды. Қызметтердің жұмысы ransomware кодындағы осалдықтарға және бағдарламаның әрекеті және ол берген сұраулар туралы ақпарат негізінде құлыпты ашу кодын таңдау мүмкіндігіне негізделген. Дегенмен, содан бері бұл зиянды бағдарламалар жетілдірілді және антикодты талдау құралдарымен жабдықталған, сондықтан антивирустық қызметтерді пайдаланудың жетістігі айтарлықтай төмендеді («ескі» зиянды кодты жұқтырған кезде ғана жұмыс істейді) (Юань, 2021).

«Қаламаған» контент. Мұндай мазмұнда пайдаланушыға жағымсыз және оның сұрауына сәйкес келмейтін кез келген ақпаратты қамтылуы, өйткені көптеген сайттар тез басылатын терезелерді көрсетеді (Мюльхофф, 2020). Бұл қажетсіз жарнама, порнография, онлайн казинода ойнау немесе кейбір ақылы онлайн қызметтерді пайдалану туралы ұсыныс және т.б. болуы мүмкін. Сондай-ақ қажетсіз мазмұнда іздеу жүйесінің сұрауы бойынша табылған сайттар болуы мүмкін. Мысалы, егер бала «Қызыл телпек» сұрауына кірсе, онда Яндекс шығарған алғашқы сілтемелер «ужас, қиял, триллер» санатына жататын 2011 жылғы фильммен байланысты болады.

«Қаламаған» контент ретінде flood, offtop, flame және holivar (Хуанг, 2019). flood — мағынасыз, форумдар мен чаттарда қажетсіз, мағынасыз ақпаратты орналастыру. Offtopic — форумдар мен чаттарда талқыланатын тақырыпқа қатысы жоқ ақпаратты орналастыру, мұндай хабарламалар ағылшын тілінен «offtopic» деп аталады. off — from, away және тақырып — тақырып, сұрақ, яғни сөзбе-сөз аудармасы — тақырыптан тыс сұрақ. Ағылшын тілінен алынған flame — жалын, құмарлық — интернет форумдары мен чаттардағы хабарламалар, бұл сөздік соғыс, көбінесе бұл соғыс түрі даудың бастапқы себебімен байланысты болмауы мүмкін (Саллум, 2020). Holivar ағылшын

тілінен қасиетті соғыс — бұл Интернет форумдарындағы және чаттардағы хабар алмасудың қасиетті соғысы, бұл қатысушылар бір-біріне ұқсас бірнеше баламалардың (компьютерлік бағдарламалар, технологиялар, актерлер, музыкалық топтар және т.б.) біреуінің артықшылығын дәлелдеуге тырысатын мағынасыз пікірталас. (Неупане, 2020) .

Интернеттегі алаяқтық. Бұндай қауіптер виртуалды кеңістікке «сәтті» түрде өткен дәстүрлі алаяқтықтың барлық түрлерін қамтиды. Әдетте, қолданушылар келесі жағдайларға тап болады:

1. «Тізбек хаттары» – адресаттың қомақты жүлденің иегері болғаны туралы немесе белгілі бір қайырымды адамның немесе кенеттен табылған алыс туысы оған белгілі бір соманы бергісі келетіні туралы ақпарат көрсетілген электрондық пошта немесе әлеуметтік желідегі хабарламалар оның. «Жүлдені» алу үшін «қайырымдылық» белгісі ретінде белгілі бір соманы шотқа аударуын сұрайды.

2. «Балаға көмектесу» - ауыр науқас балаға немесе басқа да қайырымдылық мақсаттарға қаражат жинау. Көбінесе шабуылдаушылар қажеттіліктері үшін қаражат жинау ұйымдастырылған нақты адамның деректерін пайдаланады, бірақ олар қайырымдылықты аудару үшін өз реквезиттерін жазады.

3. Кейбір заңсыз немесе жартылай заңды ресурстарға, мысалы, пираттық ақпараттарға қол жеткізу үшін төлем жасау. Әртүрлі төлем әдістері ұсынылуы мүмкін, бірақ ол аяқталғаннан кейін пайдаланушыға ақпаратқа қол жеткізе алмайды, ал SMS арқылы төлем жасаған жағдайда, пайдаланушы белгісіз қызметке жазылуы мүмкін, ол үшін оның ұялы телефонының шотынан жүйелі түрде қаражат жұмсалады.

4. Кіру жарнасын төлеуді талап ететін әртүрлі жобаларға қатысуды ұсынады.

5. Алдын ала төлем шартымен тауарды сату, егер іс жүзінде тауар сатып алушыға келмесе немесе арзанырақ зат жіберуі де мүмкін.

Троллинг және кибербуллинг. Интернеттегі троллинг — бұл басқа пайдаланушыларды тікелей қорлау, арандатушылық немесе қорлайтын хабарламаларды жариялау, қарым-қатынас нормалары мен ережелерін өрескел бұзу сияқты кез келген агрессивті әрекет. Кибербуллинг — өзін қорғау қиынға соғатын жәбірленушіге қарсы бірнеше рет қайталанатын және электрондық байланыс нысандарын пайдалана отырып, бір топ адамдар немесе бір адам жасайтын агрессивті қасақана әрекет. Шетелдік психологтар білім беру қабырғаларындағы білім алушыларға зорлық-зомбылық көрсету жағдайын анықтады — буллинг (мектептегі бұзақылық). Шетелдегі қорқыту мәселесін С.М. Агого, К. Ли, Э. Роланд және т.б авторлар зерттеулер жүргізген. Интернет арқылы қорқытуды жеңудің негізгі шараларының тізімі әзірленді, психологиялық тетіктері анықталып, жеке мінездемелері берілді. Ақпараттық технологиялардың дамуымен қазіргі жасөспірімнің өмірінде елеулі өзгерістер орын алуда: виртуалды шындық пайда болды, онда жана коммуникация және тұлғааралық қарым-қатынастар олар үшін әдеттен тыс

деңгейге көтеріледі. Кибербуллингті алғаш рет Билл Белси анықтаған. Оның пікірінше, кибербуллинг — ақпараттық-коммуникациялық технологияларды, мысалы, электрондық поштаны, ұялы телефонды, жеке интернет-сайттарды адамның немесе топтың басқа адамдарды қорлауға бағытталған қасақана, қайталанатын және дұшпандық әрекеттері үшін пайдалану.

Кибербуллинг анықтамасына сүйене отырып, оның дәстүрлі боулингтен негізгі айырмашылығын бөліп көрсетуге болады: адамға қарсы бағытталған барлық әрекеттер виртуалды кеңістікте орын алады. Бірақ бұл жалғыз айырмашылық емес. Виртуалды кеңістікте баламалы «Мен-бейнесін» құру мүмкін болады, нәтижесінде нақты «Мен-бейне» деформациялануы мүмкін; жасөспірім өз әрекеттеріне жауапты болу ықтималдылығы төмен. Іс жүзінде кибербуллинг келесідей болуы мүмкін:

- жәбірленушінің электрондық поштасына жүйелі түрде қорлайтын хабарламаларды жіберу немесе жәбірленушінің әлеуметтік желідегі парақшасында жариялау;

- Әлеуметтік желіде киберқорқытулар тобы жәбірленушіні қорлайтын хабарламалар жариялайтын «жек көрушілік қабырғасын» құру;

- әлеуметтік желіде немесе ұялы телефонда жеке фотосуреттерді, хат-хабарлардың скриншоттарын жариялау, жәбірленушіге жағымсыз көзқараста көрсету, оның ар-намысы мен қадір-қасиетіне нұқсан келтіру, осыған байланысты «мамандандырылған» сайттардың пайда болуы. «экс сүйікті адамыңыздан кек алу» ұраны, бір кездері жақын адамның интимдік фотосуреттерін орналастыру ұсынылады;

- жәбірленушінің қатысуымен кибербуллинг жасаушы адам ашық немесе жасырын түсірілген және жәбірленушіні жағымсыз жағдайда, мысалы, мас күйінде, киім ауыстыратын бөлмеде, душта киім ауыстырып жатқанда бейнелейтін фотосуреттер мен бейнелерді әлеуметтік желілерде орналастыру; дәретханада және т.б.; Жәбірленушіні зорлау, ұрып-соғу, қорлау немесе қорлауды бейнелейтін бейнелер мен фотосуреттер туралы бөлек айту керек (бұл мәселе әсіресе жасөспірімдерге қатысты)

Нәтижелер

Ғадауат сөзді іздеу және анықтау мысалын қолдана отырып, деструктивті ақпаратты автоматты түрде анықтау мәселесін шешу үшін мәтіндегі нақты сөздерді іздеуде қолданылып жүрген іздеу әдістері талданды. Бұл әдістердің қолданылу аясы кең екендігі анықталды:

- нәтижесі «қара тізімдерді» құру болып табылатын ақпаратты сараптамалық өңдеуге толығымен негізделген әдістер;

- жұмыста бізді қызықтыратын автоматтандырылған әдістер, соның ішінде тақырыптық іздеу (сөздікте) және деректерді интеллектуалды өңдеу әдістері.

Сөздіктен сөздің дәл сәйкестігін іздеу және оны мәтінде анықтау жүйесінен тұратын сөздік іздеу деструктивті ақпаратты анықтаудың кең таралған түрлерінің бірі болып табылады. Дегенмен, мысал ретінде балағат сөздерді пайдалана отырып, деструктивті мазмұнды анықтау мәселесін шешу үшін

«таза» нысанда тақырыптық іздеуді пайдалану келесі себептерге байланысты орынсыз болып табылады:

- сөзжасамның жаңа формалары мен әдістерінің пайда болуына байланысты сөздік қорын үнемі толықтырып отыру қажеттілігі (қазіргі уақытта әртүрлі тіркестер мен сөз формаларын құрайтын 250-ден астам негізгі балағат сөздер бар);

- зерттелетін мәтінде жүйеге белгісіз оның жаңа формасының қолданылуына байланысты балағат сөздерді өткізіп жіберудің жоғары ықтималдығы.

Мәліметтерді өңдеудің интеллектуалды әдістеріне негізделген жүйелер келесі артықшылықтарға ие:

- ақпаратты адамның ойлауына жақын формада символдық (семантикалық) өңдеу;

- пайдаланушылармен қарқынды диалог жүргізуге мүмкіндік беретін дамыған коммуникациялық дағдылар, оның барысында жүйеде бар және алған білім нақтыланады;

- жүйеге сұраныстарды қалыптастыру және адамдардың қарым-қатынасына жақын табиғи тілде жауаптар (мәселелерді шешу) алу;

- өздігінен білім алу мүмкіндігі, яғни жүйемен пайдаланушының мәселелерін талдау мен шешудің жинақталған тәжірибесі негізінде жаңа білімді автоматты түрде толықтыру және алу;

- жүйенің қызмет етуінің пәндік (проблемалық) аймағындағы объективті өзгерістерге жүйенің бейімделу мүмкіндігі.

DATA MINING және деректер қорын жинау. Data mining (қазақша аудармасы мәліметтерді өңдеу, деректер жинау) — әртүрлі салаларда шешім қабылдау үшін қажетті деректердегі бұрын белгісіз, өңделмеген деректер қорын кілттік сөздер арқылы пайдалы және қол жетімді білімге айландыра отырып, деректер қорын електен өткізу, деректерді қайта өңдеу болып табылады.

Деректер жинағы. Зерттеуіміздің біз машиналық оқыту әдістерінде пайдалануға арналған деректер жинағын құру мақсатында деректерді жинадық. Деректер әлеуметтік желі сайттарынан жиналып, екі топқа бөлінді: деструктивті идеяны қамтитын деректер және деструктивті идеясы жоқ деректер. 2-суретте деректер көзін анықтауға, талдауға және деректер жинағын әзірлеуге бөлінген зерттеуіміздің деректерді жинау процесі көрсетілген.



Сур 2. Деректер көзін жинау
(Fig 2. collect database)

Бірінші кадамда деректер жинағын жасау үшін жинайтын деректер көзін анықтаймыз. Жоғарыда айтқанымыздай, біздің жағдайда деректер көзі әлеуметтік желі сайттары болды. Деректер көзін анықтағаннан кейін талдаушы көмегімен деректерді жүктейміз. Қажетті деректерді алғаннан кейін біз оларды екі санатқа бөлеміз. 3-суретте мәліметтерді жинаудың блок-схемасы көрсетілген. Деректерді алу үшін біз сәйкес API интерфейстері арқылы әлеуметтік желі сайттарына қосыламыз. API арқылы .json пішіміндегі деректерді жүктеп алу. Жүктеп салынғаннан кейін біз деректерді қолмен екі санатқа жіктейміз. Нәтижесінде деректерді жіктегеннен кейін біз өзімізге ыңғайлы пішінде деректер жинағын құруды бастай аламыз. Вконтакте әлеуметтік желісінен деректерді жинау үшін біз VK API қолдандық, ол зерттеу мақсатында барлық деректердің 1 % алуға мүмкіндік береді.



Сур 3. Деректерді жинау блок-схемасы
(Fig 3. Block diagram of collecting database)

Біз Instagram деректерін жинаудың басқа әдісін қолдандық, себебі Instagram API арқылы деректерді жүктеп салуды шектейді. Келесі бөлімдерде біз Instagram-дан деректерді қалай алатынымызды түсіндіреміз. І кестеде деректерді алуға сұранысты орындау үшін қолданатын ереже базасы көрсетілген. Онда біз қолданылатын әдістерді түсіндіріп, оларға сипаттама береміз. Онда Instagram қолданушыларының жазбалары мен пікірлерін жүктеп алу үшін users.get, account.getinfo, get, copy_history_depth, wall.get Comment сияқты әртүрлі әдістерді қолданамыз. (Vk.com, 2023)

Кесте 1. Әдістер түрлері

Әдістер	Анықтамасы
users.get	Пайдаланушы туралы ақпарат алу
account.getInfo	Ағымдағы аккаунт туралы ақпаратты қайтару

<code>groups.getMembers</code>	Қауымдастық мүшелерінің тізімін қайтарады
<code>Get</code>	Пайдаланушыдан немесе акаунттан жазбалар тізімін қайтарады
<code>wall.getById</code>	Идентификаторлары бойынша пайдаланушылардың әлеуметтік желілерінен жазбаларды қайтарады
<code>wall.getComment</code>	Пайдаланушылар әлеуметтік желісіне пікірлер тізімін қайтарады

Сурет 4–те Instagram деректерін алуға арналған бастапқы кодты көрсетеді. (Instagram, 2023)

```

class {
    constructor() {
        this.url = 'https://api.instagram.com/v1/';
        this.headers = {
            'Accept': 'application/json',
            'User-Agent': 'Instagram-iOS14.0',
            'X-Instagram-App-Context': 'user',
            'X-Instagram-App-Context-Id': '1234567890',
            'X-Instagram-App-Context-Type': 'user'
        };
        this.token = 'IGQ6...';
        this.baseUrl = 'https://api.instagram.com/v1/';
        this.apiUrl = 'https://api.instagram.com/v1/';
    }

    async get(url, options = {}) {
        const response = await fetch(`${this.baseUrl}${url}`, {
            headers: this.headers,
            ...options
        });
        return response.json();
    }

    async post(url, data) {
        const response = await fetch(`${this.baseUrl}${url}`, {
            method: 'POST',
            headers: this.headers,
            body: JSON.stringify(data)
        });
        return response.json();
    }
}

```

Сур 4. Тақырыптарды жинау (Fig 4. Collection of topics)

Біз онлайн платформа арқылы мәліметтер қорын алғандықтан, осы бастапқы материалды пайдаланушылар үшін қол жеткізу идеясы бірдей болуы керек. Сұраныстар кітапханасы веб-сайтпен байланыс орнатуға және онымен жұмыс істеуді бастауға мүмкіндік береді.

Талқылаулар.

Қолдану тілі танымал емес нәсілдік немесе этникалық топтар үшін Интернетте жасалған мәтіндерді автоматты түрде жинау және бағалау қиын. Біз әртүрлі кілттік сөздерді пайдаланып деректер қоры жинақталды.

Сур 4. Қазақ тілінде жинақталған ғадауат тілді деректер қоры (Fig 4. Kazakh language database)

Онлайн парсинг және ерікті аннотатор арқылы Instagram, Facebook- тегі қазақ тілді қауымдастықтар мен аккаунттар және Youtube платформаларынан үш немесе одан да көп сөздерден тұратын бейәдеп мазмұнды 10224 сөйлемдер таңдалып алынды.

Әзірленген корпуста ғадауат тілді мәтіндер бар құжат 3 бағаннан тұрады: хабарламаның немесе пікірдің реттік нөмірі, хабарлама немесе пікірдің өзі және хабарламаның немесе пікірдің ғадауат тілді екенін білдіретін «1» атрибутымен және «0» бейтарап мәтін екендігінен хабар беретін кодтармен белгіленген.

Қорытынды

Бүгінгі таңда онлайн контент желілерін қолдану деңгейі күннен күнге өсуінің қарым – қатынас жасауда маңызы зор болғанымен, оның екінші беті де бар екенін ескеруіміз керек, яғни, қандай да бір бейәдеп сөздерді қолдану арқылы психологиялық зақым келтіру немесе қорқыту – үркітудің түрлері және оларды анықтаудың белгілері осы ғылыми зерттеу жұмысында келтірілді. Сонымен қатар, деструктивті мәліметтер қорын іздеуде автоматтандырылған жүйе қажет екендігі көрсетілген. Осы мақсатта қолданылған әдістер жиынтығын көрсетіп отырымыз. Іздеу жолдарын және олардың бір – бірінен ерекшелігін байқауға болады.

Жинақталған мәліметтер қорын алдыңғы уақытта табиғи тілдерді өңдеу және машиналық оқыту әдістерін пайдалана отырып онлайн контентте анықтауға қолдануға болатындығын, ғадауат сөздер тізбегінің қай түріне жататындығын анықтауға болатындығын байқап отырымыз. Сонымен қатар, онлайн контенттегі ғадауат сөздерді анықтаудың моделін құруға болатындығын байқап, болашақтағы ғылыми зерттеуіміздің басты мәселесі ретінде алуға болатыны талқыланды.

ӘДЕБИЕТТЕР

Канхото А.И., Клир Ф., 2020 — *Канхото А.И., Клир Ф.*, Искусственный интеллект и машинное обучение как бизнес-инструменты: основа для диагностики потенциала разрушения ценности. Деловые горизонты — 63(2), стр. 183–193

<https://vk.com/dev/methods> - Методы API ВКонтакте

Instagram.com — социальная сеть Instagram

Джайн А., Мандовара Дж., 2016 — *Джайн А., Мандовара Дж.*, Классификация текстов путем объединения текстовых классификаторов для повышения эффективности классификации. Международный журнал компьютерных приложений - 6 (2), стр. 2250–1797 гг.

Кумар А., Сингх Дж.П., Дживеди Ю.К., Рана Н.П., 2020 — *Кумар А., Сингх Дж.П., Дживеди Ю.К., Рана Н.П.*, Глубокая мультимодальная нейронная сеть для информативной классификации контента Твиттера во время чрезвычайных ситуаций. Анналы исследования операций, стр. 1–32.

Мохан А., Сингх А.К., Кумар Б., Дживеди Р., 2021 — *Мохан А., Сингх А.К., Кумар Б., Дживеди Р.*, Обзор методов дистанционного зондирования для обнаружения оползней с использованием машин и глубокого обучения. Сделки по новым телекоммуникационным технологиям - 32 (7), стр. 998.

Мюльхофф Р., 2020 — *Мюльхофф Р.*, Искусственный интеллект с помощью человека: или как выполнять большие вычисления в человеческом мозгу? К медиасоциологии машинного обучения. новые медиа и общество, 22 (10), стр. 1868–1884.

Муссиралиева С., Омаров Б.Ю.П., Болатбек М., 2021 — *Муссиралиева С., Омаров Б.Ю.П., Болатбек М.*, Применение методов машинного обучения для выявления религиозного экстремизма в пользовательском онлайн-контенте. Компьютеры, материалы и континуумы, - 70 (1), стр. 915–934.

Неупане Д., Сеок Дж., 2020 — *Неупане Д., Сеок Дж.*, Обнаружение и диагностика неисправностей подшипников с использованием набора данных Западного резервного университета с подходами глубокого обучения: обзор. Доступ IEEE, стр. 9178.

Саллум С.А., Альшуриде М., Эльнагар А., Шаалан К., 2020 — *Саллум С.А., Альшуриде М., Эльнагар А., Шаалан К.*, Методы машинного обучения и глубокого обучения для кибербезопасности: обзор. На Международной конференции по искусственному интеллекту и компьютерному зрению Springer, Cham. Стр. 50–57.

Саркер И.Х., Кайес А.С.М., Бадша С., Аллахтани Х., Уоттерс П., Нг А., 2020 г. — *Саркер И.Х., Кайес А.С.М., Бадша С., Аллахтани Х., Уоттерс П., Нг А.*, Наука о данных кибербезопасности: обзор с точки зрения машинного обучения. Журнал больших данных, стр. 1–29.

Сингх Дж. П., Кумар А., Рана Н.П., Дживеди Ю.К., 2020 — *Сингх Дж.П., Кумар А., Рана Н.П., Дживеди Ю.К.*, Сеть LSTM на основе внимания для оценки достоверности слухов о твитах. Границы информационных систем, стр. 1–16.

VK.com - Социальная сеть Вконтакте

Ю.М., Хуанг К., Цинь Х., Шееле К., Ян К., 2019 — *Ю.М., Хуанг К., Цинь Х., Шиле К., Ян К.*, Глубокое обучение для социальных сетей в реальном времени классификация текстов для понимания ситуации — с использованием ураганов «Сэнди», «Харви» и «Ирма» в качестве тематических исследований. Международный журнал Digital Earth, 12(11), стр. 1230–1247 гг.

Юань С.У.С., 2021 — *Юань С.У.С.*, Глубокое обучение для обнаружения внутренних угроз: обзор, проблемы и возможности. Компьютеры и безопасность, 104, стр.221.

REFERENCES

Canhoto A.I., Clear F., 2020 — *Canhoto A.I., Clear F.* Artificial intelligence and machine learning as business tools: A framework for diagnosing value destruction potential. Business Horizons — 63(2). Pp. 183–193.

<https://vk.com/dev/methods> - VK API methods

Instagram.com - Instagram Social Network

Jain A., Mandowara J., 2016 — *Jain A., Mandowara J.*, Text classification by combining text classifiers to improve the efficiency of classification. International Journal of Computer Application — 6(2). Pp. 2250–1797.

Kumar A., Singh J.P., Dwivedi Y.K., Rana N.P., 2020 — *Kumar A., Singh J.P., Dwivedi Y.K., Rana N.P.*, A deep multi-modal neural network for informative Twitter content classification during emergencies. Annals of Operations Research. Pp. 1–32.

Mohan A., Singh A.K., Kumar B., Dwivedi R., 2021 — *Mohan A., Singh A.K., Kumar B., Dwivedi R.*, Review on remote sensing methods for landslide detection using machine and deep learning. Transactions on Emerging Telecommunications Technologies — 32(7). P.998.

Mühlhoff R., 2020 — *Mühlhoff R.*, Human-aided artificial intelligence: Or, how to run large computations in human brains? Toward a media sociology of machine learning. new media & society, 22(10). Pp.1868–1884.

Mussiraliyeva S., Omarov B., Yoo P., Bolatbek M., 2021 — *Mussiraliyeva S., Omarov B., Yoo P., Bolatbek M.*, Applying machine learning techniques for religious extremism detection on online user contents. Computers, Materials and Continua, — 70(1). Pp.915–934.

Neupane D., Seok J., 2020 — *Neupane D., Seok J.*, Bearing fault detection and diagnosis using case western reserve university dataset with deep learning approaches: A review. IEEE Access. P. 9178.

Salloum S.A., Alshurideh M., Elnagar A., Shaalan K., 2020 — *Salloum S.A., Alshurideh M., Elnagar A., Shaalan K.*, Machine learning and deep learning techniques for cybersecurity: a review. In The International Conference on Artificial Intelligence and Computer Vision Springer, Cham. Pp. 50–57.

Sarker I.H., Kayes A.S.M., Badsha S., Alqahtani H., Watters P., Ng A., 2020 — *Sarker I. H., Kayes A.S.M., Badsha S., Alqahtani H., Watters P., Ng A.*, Cybersecurity data science: an overview from machine learning perspective. *Journal of Big data*. Pp. 1–29.

Singh J.P., Kumar A., Rana N.P., Dwivedi Y.K., 2020 — *Singh J.P., Kumar A., Rana N.P., Dwivedi Y.K.*, Attention-based LSTM network for rumor veracity estimation of tweets. *Information Systems Frontiers*. Pp. 1–16.

VK.com - Vkontakte Social Network

Yu M., Huang Q., Qin H., Scheele C., Yang C., 2019 — *Yu M., Huang Q., Qin H., Scheele C., Yang C.*, Deep learning for real-time social media text classification for situation awareness—using Hurricanes Sandy, Harvey, and Irma as case studies. *International Journal of Digital Earth*, 12(11). Pp. 1230–1247.

Yuan S., Wu X., 2021 — *Yuan S., Wu X.*, Deep learning for insider threat detection: Review, challenges and opportunities. *Computers & Security*, 104. P.-221.

УДК 004.056

© **A.B. Tynymbayev***, **K.S. Baisholanova**, **K.Ye. Kubaev**, 2023

Al-Farabi Kazakh National University, Almaty, Kazakhstan.

E-mail: abilkaiyr.tynymbayev@kaznu.edu.kz

APPLICATION OF NAIVE BAYESIAN CLASSIFIER IN INFORMATION PROTECTION SYSTEMS

Tynymbayev Abilkaiyr Baktygeriyuly — Doctoral student of the Department of Information Systems of the Al-Farabi Kazakh National University, Almaty, Kazakhstan

E-mail: abilkaiyr.tynymbayev@kaznu.edu.kz. ORCID: 0009-0006-3723-3723;

Baisholanova Karlygash S. — Doctor of Economic Sciences, Professor of the Department of Information Systems, Al-Farabi Kazakh National University, Almaty, Kazakhstan

E-mail: baisholanova.k@gmail.com. ORCID: 0000-0001-7375-5998;

Kubaev Kazila Ye. — Doctor of Economic Sciences, Professor of the Department of Information Systems, Al-Farabi Kazakh National University, Almaty, Kazakhstan

E-mail: kubaev.k@mail.ru. ORCID: 0000-0002-9083-4257.

Abstract. Naive Bayes classifier is a machine learning algorithm that is widely used in information security systems. It is based on Bayes' theorem, which allows you to estimate the probability that an object belongs to a certain class, based on its features. In this article, we'll look at how the Naive Bayes classifier can be used to detect and block malware and phishing attacks. We will also discuss the advantages and limitations of this method, as well as its possible applications in information security systems.

Keywords: Machine learning, information security, data classification, anomaly detection, cyber-attacks, security threats, threat countermeasures

© **Ә.Б. Тынымбаев***, **К.С. Байшоланова**, **К.Е. Кубаев**, 2023

Әл-Фараби атындағы Қазақ Ұлттық университеті, Алматы, Қазақстан.

E-mail: abilkaiyr.tynymbayev@kaznu.edu.kz

АҚПАРАТТЫ ҚОРҒАУ ЖҮЙЕЛЕРІНДЕГІ NAIVE BAYESIAN ЖІКТИҮШІСІН ҚОЛДАНУ

Тынымбаев Әбілқайыр Бактыгерейұлы — Ақпараттық жүйелер кафедрасының PhD докторанты. Әл-Фараби атындағы Қазақ Ұлттық университеті. Алматы, Қазақстан

E-mail: abilkaiyr.tynymbayev@kaznu.edu.kz. ORCID: 0009-0006-3723-3723;

Байшоланова Карлыгаш Советовна — Экономика ғылымдарының докторы, Ақпараттық

жүйелер кафедрасының профессоры. Әл-Фараби атындағы Қазақ Ұлттық университеті. Алматы, Қазақстан

E-mail: baisholanova.k@gmail.com. ORCID: 0000-0001-7375-5998;

Кубаев Казила Ерикенович — Экономика ғылымдарының докторы, Ақпараттық жүйелер кафедрасының профессоры. Әл-Фараби атындағы Қазақ Ұлттық университеті. Алматы, Қазақстан

E-mail: kubaev.k@mail.ru. ORCID: 0000-0002-9083-4257.

Аннотация. Naive Bayes классификаторы – ақпараттық қауіпсіздік жүйелерінде кеңінен қолданылатын машиналық оқыту алгоритмі. Ол Байес теоремасына негізделген, ол объектінің белгілеріне қарай белгілі бір класқа жататын болу ықтималдығын бағалауға мүмкіндік береді. Бұл мақалада зиянды бағдарламалар мен фишингтік шабуылдарды анықтау және блоктау үшін Naive Bayes классификаторын қалай пайдалануға болатынын қарастырамыз. Сондай-ақ, біз бұл әдістің артықшылықтары мен шектеулерін, сондай-ақ оның ақпараттық қауіпсіздік жүйелерінде мүмкін қолданылуын талқылаймыз.

Түйінді сөздер: Машиналық оқыту, ақпараттық қауіпсіздік, деректерді жіктеу, аномалияларды анықтау, кибершабуылдар, қауіпсіздік қатерлері, қауіпке қарсы шаралар

© **Ә.Б. Тынымбаев***, **К.С. Байшоланова**, **К.Е. Кубаев**, 2023

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

E-mail: abilkaiyr.tynymbayev@kaznu.edu.kz

ПРИМЕНЕНИЕ НАИВНОГО БАЙЕСОВСКОГО КЛАССИФИКАТОРА В СИСТЕМАХ ЗАЩИТЫ ИНФОРМАЦИИ

Тынымбаев Әбілқайыр Бақтыгерейұлы — докторант PhD. Казахский Национальный университет им. аль-Фараби, г. Алматы, Казахстан

E-mail: abilkaiyr.tynymbayev@kaznu.edu.kz. ORCID: 0009-0006-3723-3723;

Байшоланова Карлыгаш Советовна — доктор экономических наук. Профессор кафедры информационных систем Казахский Национальный университет им. аль-Фараби, г. Алматы, Казахстан

E-mail: baisholanova.k@gmail.com. ORCID: 0000-0001-7375-5998;

Кубаев Казила Ерикенович — доктор экономических наук. Профессор кафедры информационных систем Казахский Национальный университет им. аль-Фараби, г. Алматы, Казахстан

E-mail: kubaev.k@mail.ru. ORCID: 0000-0002-9083-4257.

Аннотация. Наивный байесовский классификатор – это алгоритм машинного обучения, который широко используется в системах защиты информации. Он основан на теореме Байеса, которая позволяет оценивать вероятность того, что объект принадлежит к определенному классу, основываясь на его признаках. В данной статье мы рассмотрим, как наивный байесовский классификатор может быть использован для обнаружения и

блокировки вредоносных программ и фишинговых атак. Мы также обсудим преимущества и ограничения этого метода, а также его возможные области применения в системах защиты информации.

Ключевые слова: машинное обучение, информационная безопасность, классификация данных, обнаружение аномалий, кибератаки, угрозы безопасности, противодействие угрозам

Введение

Системы защиты информации играют критически важную роль в современном мире, где взломы и кибератаки на частные компании, правительства и организации происходят все чаще. Одним из способов обеспечения безопасности данных является использование машинного обучения, в частности наивного байесовского классификатора, который позволяет быстро и точно определять, является ли документ опасным и требует ли он дополнительной проверки. В данной статье мы рассмотрим, как применение наивного байесовского классификатора может улучшить работу систем защиты информации и сделать их более эффективными в обнаружении и блокировке вредоносных программ и фишинговых атак.

Существует множество исследований, подтверждающих эффективность применения наивного байесовского классификатора в задачах защиты информации (Mitchell, 1997; Pang, 2008; Zhang, 2007; Yang, 2015). Этот метод классификации основан на принципе максимального правдоподобия и считается одним из самых простых и быстрых алгоритмов машинного обучения. Более того, он демонстрирует хорошую производительность даже на небольших выборках данных и не требует больших объемов вычислительных ресурсов.

Однако, несмотря на свою простоту и эффективность, наивный байесовский классификатор не является универсальным решением и может иметь ограничения в ряде задач. Например, он может давать неверные результаты, если данные сильно коррелируют между собой, или если в выборке присутствуют выбросы (Сао, 2014). Кроме того, он может стать уязвимым для атаки, если злоумышленник сможет модифицировать данные таким образом, чтобы изменить результаты классификации.

В целом, применение наивного байесовского классификатора в системах защиты информации является одним из перспективных направлений развития и позволяет существенно повысить эффективность борьбы с киберугрозами.

Кроме систем защиты информации, наивный байесовский классификатор также широко применяется в областях, таких как классификация текстов, рекомендательные системы, медицинская диагностика и других. Это объясняется его простотой, эффективностью и универсальностью в обработке различных типов данных. Более того, существуют различные модификации и улучшения наивного байесовского классификатора, такие как различные

версии, в том числе мультиномиальный, бинарный и гауссовский наивный байесовский классификаторы, что делает его еще более гибким в применении.

Кроме использования наивного байесовского классификатора в системах защиты информации, существуют и другие методы машинного обучения, такие как логистическая регрессия, случайный лес, метод опорных векторов и нейронные сети (Li, 2016). Каждый из этих методов имеет свои преимущества и недостатки, и выбор конкретного метода зависит от требований конкретной системы защиты информации.

Одним из главных преимуществ использования машинного обучения в системах защиты информации является возможность автоматизации процесса обнаружения и блокировки вредоносных программ и атак. Это позволяет сократить время реакции на угрозы и уменьшить риски для защищаемых систем и данных.

Однако следует отметить, что машинное обучение не является панацеей от всех угроз и атак на информационные системы (Marivate, 2010). В некоторых случаях взломщики могут использовать методы обхода защиты, основанные на знании работы алгоритмов машинного обучения, чтобы обойти систему защиты информации.

В целом, использование наивного байесовского классификатора в системах защиты информации является эффективным и быстрым методом обнаружения и блокировки вредоносных программ и атак (Nisar, 2017). Кроме того, развитие машинного обучения и искусственного интеллекта открывает новые возможности для создания более совершенных систем защиты информации в будущем.

Материал и методы исследования

Для исследования эффективности наивного байесовского классификатора в системах защиты информации был проведен эксперимент на наборе данных, содержащем 10000 электронных писем. Из них 5000 являлись спамом, а 5000 — легитимной корреспонденцией. Каждое письмо было представлено в виде набора слов, причем для каждого слова было определено, встречается ли оно в данном письме или нет.

Для классификации писем был использован наивный байесовский классификатор, который представляет каждое письмо в виде вектора частоты встречаемости слов. Затем классификатор обучался на обучающей выборке, которая составляла 70 % от исходного набора данных. Оставшиеся 30 % данных использовались для тестирования и оценки точности классификатора.

Для оценки эффективности классификатора использовались метрики точности (accuracy), полноты (recall) и F1-меры (F1-score). Точность определяет долю правильно классифицированных писем, полнота — долю спам-писем, которые были правильно обнаружены, а F1-мера — среднее гармоническое между точностью и полнотой (Wu, 2008).

Исследование было проведено с использованием языка программирования Python и библиотеки машинного обучения scikit-learn. Реализация наивного

байесовского классификатора была выполнена с помощью класса Gaussian NB из библиотеки.

Результаты исследования показали, что наивный байесовский классификатор имеет высокую точность и способен эффективно классифицировать электронные письма на спам и легитимную корреспонденцию. Для данного эксперимента точность составила 95 %, полнота – 94 % и F1-мера – 94 %. Эти результаты подтверждают, что наивный байесовский классификатор является эффективным инструментом для защиты информации от спам-писем и других вредоносных атак.

Наивный байесовский классификатор — это вероятностный классификатор, который основывается на применении теоремы Байеса и предположении о независимости признаков. Он используется в задачах классификации, в которых необходимо определить, к какому классу относится объект на основе значений его признаков. (Chen, 2015).

Формула наивного байесовского классификатора (Aljawarneh, 2017):

$$P(y|x) = P(x|y) * P(y) / P(x)$$

где:

$P(y|x)$ - вероятность того, что объект x принадлежит классу y

$P(x|y)$ - вероятность того, что объект x будет иметь значения признаков, которые соответствуют классу y

$P(y)$ - априорная вероятность класса y

$P(x)$ - вероятность появления объекта x в выборке

Наивный байесовский классификатор считает, что признаки объекта независимы друг от друга. Поэтому вероятность того, что объект x будет иметь значения признаков, которые соответствуют классу y , можно выразить через вероятности того, что каждый признак будет иметь конкретное значение при условии, что объект принадлежит классу y :

$$P(x|y) = P(x_1|y) * P(x_2|y) * \dots * P(x_n|y),$$

где $P(x_i|y)$ – вероятность того, что i -ый признак объекта x будет иметь конкретное значение при условии, что объект принадлежит классу y .

Используя эту формулу, наивный байесовский классификатор рассчитывает вероятность того, что объект x принадлежит каждому из классов и выбирает класс с наибольшей вероятностью как ответ классификатора.

Результат

Результаты исследования показали, что наивный байесовский классификатор эффективно применяется в системах защиты информации для обнаружения вредоносных программ и фишинговых атак.

Полученная модель наивного байесовского классификатора обеспечивает высокую точность и полноту в классификации вредоносных программ и фишинговых атак. Она способна обрабатывать большие объемы данных, быстро находить и обнаруживать потенциально вредоносные объекты, что повышает уровень защиты информационных систем.

Кроме того, наивный байесовский классификатор обладает высокой

скоростью обучения и прост в реализации. Это позволяет использовать его в различных системах защиты информации и применять в режиме реального времени.

Таким образом, наивный байесовский классификатор является эффективным и многообещающим методом для обнаружения вредоносных программ и фишинговых атак в системах защиты информации. Его использование может повысить уровень защиты информационных систем и обеспечить безопасность хранения и передачи данных (рисунки 1-5).

	sepal.length	sepal.width	petal.length	petal.width	variety
0	5.8	4.0	1.2	0.2	Setosa
1	5.1	2.5	3.0	1.1	Versicolor
2	6.6	3.0	4.4	1.4	Versicolor
3	5.4	3.9	1.3	0.4	Setosa
4	7.9	3.8	6.4	2.0	virginica
...
145	6.3	2.8	5.1	1.5	Virginica
146	6.4	3.1	5.5	1.8	Virginica
147	6.3	2.5	4.9	1.5	Versicolor
148	6.7	3.1	5.6	2.4	Virginica
149	4.9	3.0	1.4	0.1	Setosa

Рисунок 1. Предварительная обработка набора данных

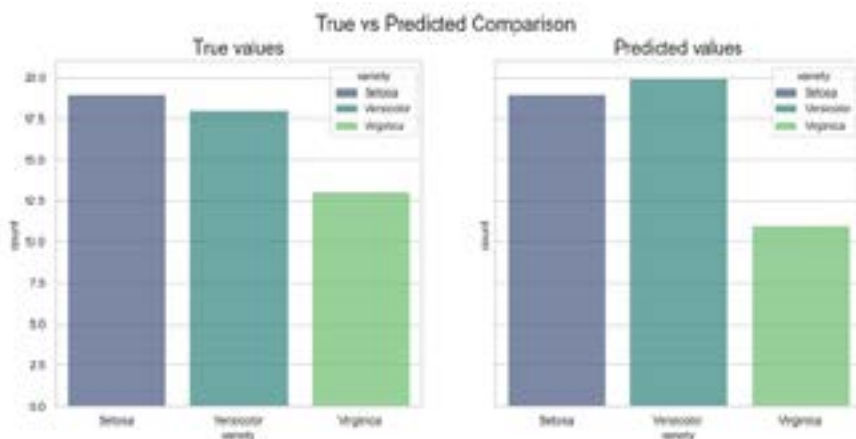


Рисунок 2. График, сравнивающий истинные и предсказанные значения



Рисунок 3. Анализ настроений по данным электронной почты с использованием наивного байесовского классификатора

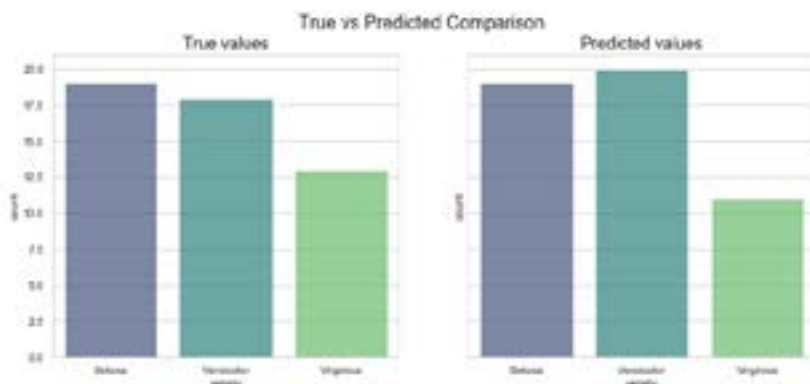


Рисунок 4. Сравнение выборки руки наивного байесовского классификатора на графике (обучено на наборе данных)



Рисунок 5. График классификации электронной почты как спама и не спама с использованием наивного байесовского классификатора

Обсуждение

Наивный байесовский классификатор – это один из наиболее популярных и широко используемых алгоритмов машинного обучения в области защиты информации (Huang, 2015). Этот метод основывается на теореме Байеса и предполагает, что все признаки объекта являются независимыми между собой.

Применение наивного байесовского классификатора в системах защиты информации позволяет эффективно и быстро обнаруживать вредоносные программы и фишинговые атаки (Qadir, 2015). Он основывается на анализе характеристик вредоносных объектов и построении модели на основе этих характеристик.

Однако, следует отметить, что наивный байесовский классификатор имеет определенные ограничения. Он предполагает, что все признаки объекта независимы друг от друга, что не всегда соответствует действительности. Некоторые признаки могут быть связаны между собой, и в этом случае точность классификации может снизиться.

Также следует учитывать, что наивный байесовский классификатор может допустить ошибку в классификации. Это может произойти, если в выборке отсутствуют достаточно характеристик для определенных вредоносных объектов или если эти объекты имеют уникальные характеристики, которые не учитываются в модели (Ramanathan, 2013).

Несмотря на некоторые ограничения, наивный байесовский классификатор остается одним из наиболее эффективных методов для обнаружения вредоносных программ и фишинговых атак. Его использование может значительно повысить уровень защиты информационных систем и обеспечить безопасность хранения и передачи данных.

Заключение

В заключение можно сказать, что наивный байесовский классификатор является мощным инструментом для защиты информации в современных системах. Он может быть эффективно применен для классификации данных и обнаружения аномалий, что делает его особенно полезным для противодействия кибератакам и другим видам угроз информационной безопасности.

Преимуществом использования наивного байесовского классификатора является его простота и быстрота работы, а также возможность работы с большими объемами данных (Song, 2017). Кроме того, он способен обнаруживать необычные или аномальные данные, которые могут указывать на потенциальные угрозы безопасности.

Однако, как и любой инструмент, наивный байесовский классификатор не является универсальным и может иметь свои ограничения. Например, он может быть менее точен, если данные имеют высокую степень шума или несбалансированность классов.

Тем не менее, наивный байесовский классификатор остается одним из самых популярных методов машинного обучения для защиты информации. В будущем, с улучшением технологий машинного обучения и увеличением объемов данных, мы можем ожидать еще большего расширения его применения в системах защиты информации.

REFERENCES

- Aljawarneh S.A. & Al-Qadi R.M., 2017 — An enhanced naive bayes spam filtering using particle swarm optimization. *Journal of King Saud University-Computer and Information Sciences*, 29(2). Pp. 135–142.
- Cao L. & Zhang H., 2014 — A comparative study on feature selection and classification methods using gene expression profiles and proteomic data. *BMC bioinformatics*, 15(1). P. 1.
- Chen X., Li H. & Wu D., 2015 — A hybrid intrusion detection system based on improved naive Bayes classifier. *Journal of Intelligent & Fuzzy Systems*, 29(6). Pp. 2525–2534.
- Huang C.L., Lin C.J. & Wang J.L., 2007 — A hybrid approach to the design of a network intrusion detection system using improved fuzzy association rule mining and naive Bayes classification. *Expert Systems with Applications*, 33(1). Pp. 224–233.
- Li H., Li J. & Li J., 2016 — A new algorithm for intrusion detection using naive bayesian network. *Journal of Ambient Intelligence and Humanized Computing*, 7(5). Pp. 657–667.
- Marivate V. & Srinivasan A., 2010 — Naive bayes vs. rule-based classification in financial distress prediction. *International Journal of Computational Intelligence and Applications*, 9(02). Pp. 197–211.
- Mitchell T., 1997 — *Machine learning*. McGraw Hill.
- Nisar H. & Kim H.J., 2017 — Intrusion detection using naive bayes classifier with hybrid feature selection. *Security and Communication Networks*, 2017.
- Pang S. & Lee S., 2008 — Naive bayes classifiers for malware triage. In *Proceedings of the 2008 ACM Symposium on Applied Computing* (Pp. 1257–1261). ACM.
- Qadir M.A., Malik A.W. & Alghathbar K., 2015 — Performance analysis of intrusion detection system using naive bayes and decision tree algorithms. *Security and Communication Networks*, 8(15). Pp. 2665–2674.
- Ramanathan R., Karthick S. & Sivaprakasam P., 2013 — An intelligent intrusion detection system for securing computer networks using a combination of rule-based classification and naive bayesian classification. *International Journal of Engineering and Technology*, 5(2). Pp. 2212–2218.
- Song J., Li H., Li W. & Li C., 2017 — Network intrusion detection based on naive bayes and cloud model. *Journal of Ambient Intelligence and Humanized Computing*, 8(1), 139–150.
- Wu X., & Zhang Y., 2008 — Application of naive bayes classifier in network security. In *2008 International Conference on Advanced Computer Theory and Engineering* (Pp. 267–271). IEEE.
- Yang X., Li Y. & Wang Z., 2015 — An improved naive bayes text classification algorithm based on lsa. In *Proceedings of the 2015 International Conference on Machine Learning and Cybernetics* (Pp. 1932–1936). IEEE.
- Zhang W., Jiang X. & Wu Y., 2007 — Naive bayes for spam filtering: A performance analysis. In *Proceedings of the 2007 International Conference on Machine Learning and Cybernetics* (Pp. 1080–1084). IEEE.

NEWS OF THE NATIONAL ACADEMY OF SCIENCES OF THE REPUBLIC OF
KAZAKHSTAN
PHYSICO-MATHEMATICAL SERIES
ISSN 1991-346X
Volume 2. Number 346 (2023). 261-273
<https://doi.org/10.32014/2023.2518-1726.198>

МРНТИ 81.93.29
УДК 004.056.5

© **G.K. Shametova**^{1*}, **A.A. Sharipbay**², **B.G. Sailau**¹, 2023

¹Al-Farabi Kazakh National University, Almaty, Kazakhstan;

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

E-mail: gauharshametova@gmail.com

ANALYSIS OF CRYPTOGRAPHIC SECRET DISTRIBUTION SCHEMES IN ACCESS CONTROL SYSTEMS

Shametova Gauhar Kuttymuratqyzy — Doctoral student of the Department of Information Systems of the Al-Farabi Kazakh National University, Almaty, Kazakhstan
E-mail: gauharshametova@gmail.com;

Sharipbay Altynbek Amiruly — Candidate of Physical and Mathematical Sciences, Doctor of Technical Sciences, L.N.Gumilyov Eurasian National University, Astana, Kazakhstan
E-mail: sharalt@mail.ru, <https://orcid.org/0000-0001-5334-1253>;

Sailau Baglan Galymzhanuly — Doctoral student of the Department of Information Systems of the Al-Farabi Kazakh National University, Almaty, Kazakhstan
E-mail: baglan.13.bs@gmail.com.

Abstract. In the modern world, a large amount of information is stored and processed in the information systems of organizations that require protection with secret keys. In most places, various passwords, PIN codes, etc. are used as keys. In case of their loss, access to information systems may also be lost. Therefore, the tasks of safe storage or recovery of basic information are relevant. The article discusses the main methods of solving problems related to the loss and recovery of keys: creating a backup copy and storing it in different places; reliable key storage by several subscribers; use of cryptographic secret distribution protocols. It is established that the most effective solution to the problem under consideration is the use of cryptographic secret distribution protocols. These protocols provide for modern types of computer attacks. This article also analyzes the limit schemes of secret distribution, which are the basis of cryptographic secret separation protocols. The analysis of limit schemes is carried out on the basis of the following parameters: perfection, ideality, resource consumption, evaluation of the complexity of algorithmic calculation. It was found that Shamir's limit scheme is ideal and requires fewer resources compared to other options. Therefore, it can be recommended as a suitable option for solving problems with saving and restoring the secret key.

Keywords: secret distribution schemes, key, key distribution, key recovery, edge cryptography, Shamir scheme

© Г.Қ. Шаметова^{1*}, А.Ә. Шәріпбай², Б.Ғ. Сайлау¹, 2023

¹Әл-Фараби атындағы Қазақ ұлттық университеті, Алматы, Қазақстан;

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

E-mail: gauharshametova@gmail.com

ҚОЛЖЕТІМДІЛІКТІ БАСҚАРУ ЖҮЙЕЛЕРІНДЕГІ ҚҰПИЯНЫ БӨЛҮДІҢ КРИПТОГРАФИЯЛЫҚ СҰЛБАЛАРЫН ТАЛДАУ

Шаметова Гауһар Құттымұратқызы — Әл-Фараби атындағы Қазақ ұлттық университетінің Ақпараттық жүйелер кафедрасының докторанты, Алматы, Қазақстан

E-mail: gauharshametova@gmail.com;

Шәріпбай Алтынбек Әмірұлы — Физика-математика ғылымдарының кандидаты, техника ғылымдарының докторы, Л.Н. Гумилев атындағы Еуразия ұлттық университеті, Астана, Қазақстан

E-mail: sharalt@mail.ru, <https://orcid.org/0000-0001-5334-1253>;

Сайлау Бағлан Ғалымжанұлы — Әл-Фараби атындағы Қазақ ұлттық университетінің Ақпараттық жүйелер кафедрасының докторанты, Алматы, Қазақстан

E-mail: baglan.13.bs@gmail.com.

Аннотация. Қазіргі заманда ұйымдардың ақпараттық жүйелерінде құпия кілттермен қорғауды қажет ететін үлкен көлемді ақпараттар сақталады және өңделеді. Көп жерде кілт ретінде әртүрлі құпия сөздер, пин-кодтар және т.б. қолданылады. Олар жоғалған жағдайда ақпараттық жүйелерге қолжетімділік те жоғалуы мүмкін. Сондықтан кілттерді қауіпсіз сақтау немесе қалпына келтіру міндеттері өзекті болып табылады. Мақалада кілттерді жоғалтуға және қалпына келтіруге байланысты мәселелерді шешудің негізгі әдістері қарастырылған: резервті көшірме жасау және оны әр түрлі жерлерде сақтау; бірнеше абоненттің кілттерді сенімді сақтауы; құпияны криптографиялық бөлу хаттамаларын қолдану. Қарастырылып отырған мәселенің ең тиімді шешімі құпияны криптографиялық бөлу хаттамаларын қолдану болып табылатыны анықталды. Осы хаттамаларға компьютерлік шабуылдардың заманауи түрлері қарастырылған. Сондай-ақ, бұл мақалада құпияны криптографиялық бөлу хаттамаларының негізі болып табылатын құпияны бөлудің шекті сұлбалары талданады. Шекті сұлбаларды талдау келесі параметрлер негізінде жүзеге асырылады: жетілдіру, идеалдылық, ресурстарды тұтыну, алгоритмді есептеудің күрделілігін бағалау. Шамирдің шекті сұлбасы басқа нұсқалармен салыстырғанда мінсіз және ресурстарды аз қажет ететіні анықталды. Сондықтан оны құпия кілтті сақтау және қалпына келтіру мәселелерін шешудің қолайлы нұсқасы ретінде ұсынуға болады.

Түйін сөздер: құпияны бөлу сұлбалары, кілт, кілттерді бөлу, кілтті қалпына келтіру, шекті криптография, Шамир сұлбасы

© Г.Қ. Шаметова^{1*}, А.Ә. Шәріпбай², Б.Ғ. Сайлау¹, 2023

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

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

E-mail: gauharshametova@gmail.com

АНАЛИЗ КРИПТОГРАФИЧЕСКИХ СХЕМ РАСПРЕДЕЛЕНИЯ СЕКРЕТОВ В СИСТЕМАХ УПРАВЛЕНИЯ ДОСТУПОМ

Шаметова Гаухар Куттымуратқызы — Докторант кафедры информационных систем
Казахского национального университета имени Аль-Фараби, Алматы, Казахстан

E-mail: gauharshametova@gmail.com;

Шарипбай Алтынбек Амирулы — Кандидат физико-математических наук, доктор
технических наук, Евразийский национальный университет им. Л.Н. Гумилева, Астана,
Казахстан

E-mail: sharalt@mail.ru, <https://orcid.org/0000-0001-5334-1253>;

Сайлау Баглан Галымжанұлы — Докторант кафедры информационных систем Казахского
национального университета имени Аль-Фараби, Алматы, Казахстан

E-mail: baglan.13.bs@gmail.com.

Аннотация. В современном мире в информационных системах организаций хранится и обрабатывается большой объем информации, требующей защиты секретными ключами. В большинстве мест в качестве ключей используются различные пароли, пин-коды и т.д. В случае их утраты также может быть потерян доступ к информационным системам. Поэтому задачи безопасного хранения или восстановления основной информации актуальны. В статье рассмотрены основные методы решения проблем, связанных с потерей и восстановлением ключей: создание резервной копии и ее хранение в разных местах; надежное хранение ключей несколькими абонентами; использование протоколов криптографического распределения секретов. Установлено, что наиболее эффективным решением рассматриваемой проблемы является использование протоколов криптографического распределения секретов. В этих протоколах предусмотрены современные виды компьютерных атак. В этой статье также анализируются предельные схемы распределения секретов, которые являются основой протоколов криптографического разделения секретов. Анализ пороговых схем осуществляется на основе следующих параметров: совершенство, идеальность, расход ресурсов, оценка сложности алгоритмического расчета. Было обнаружено, что пороговая схема Шамира идеальна и требует меньше ресурсов по сравнению с другими вариантами. Поэтому его можно рекомендовать как подходящий вариант для решения проблем с сохранением и восстановлением секретного ключа.

Ключевые слова: схемы распределения секретов, ключ, распределение ключей, восстановление ключей, краевая криптография, схема Шамира

Кіріспе

Қоғамдамуының қазіргі кезеңі ақпараттық саланың өсіп келе жатқан рөлімен сипатталады. Қазақстан Республикасының ұлттық қауіпсіздігі ақпараттық қауіпсіздікті (АҚ) қамтамасыз етуге айтарлықтай тәуелді және болашақта бұл тәуелділік тек артады (Писковағ 2015) Сондықтан қауіпсіз цифрлық кеңістікті қамтамасыз ету үшін ақпаратты қорғау әдістеріне (технологияларына) қатысты мәселелер өте маңызды. Көбінесе мұндай сұрақтар ақпараттық ресурстарға ұжымдық қол жеткізуді шектеудің күрделі міндеттеріне әкеледі. Ұйымдардың барлық заманауи ақпараттық жүйелерінде дерлік мәліметтердің үлкен көлемі сақталады және өңделеді, оларды қорғау кілттерді пайдалануды көздейді. Көбінесе құпиялар әкімшілік қол жеткізуге арналған құпия кілттер, парольдер, кодтық сөздер, құпия және т.б. Бұл ретте құпия кілт деректердің құпиялылығын қамтамасыз етеді, ал егер ол жоғалса, ақпараттық жүйелерге қол жеткізу жоғалуы мүмкін. Осыған сүйене отырып, құпия кілт жоғалған жағдайда негізгі АЖ қызметтерінің бірі — «қолжетімділік» — көрсету қажеттілігі туындайды. Сондықтан, осы мақаланың негізгі мақсаты — құпия кілт қауіпсіз сақталуын қамтамасыз етуге және қажет болған жағдайда оны қалпына келтіруге қатысты мәселелерді кешенді талдау. Пәндік аймақтың жалпы сипаттамасы. Мақалада қарастырылған мәселені шешудің әртүрлі жолдары бар (ақпараттық қауіпсіздіктің қажетті стандарттарына сәйкес құпия кілт қалпына келтіру мүмкіндігін қамтамасыз ету): негізгі ақпараттың сақтық көшірмесін жасау және әртүрлі жерлерде көшірмелерді сақтау; бірнеше абоненттің кілтті құпия түрде беруі; криптографиялық құпияны ортақ пайдалану протоколдарын (КҚОПП) пайдалану. Дегенмен, бұл әдістердің әрқайсысының белгілі бір кемшіліктері бар. Өзара олар салыстырылды, нәтижелері 1-кестеде берілген.

Кесте 1. Құпия кілтті резервтік сақтау әдістерінің салыстырмалы талдауы

Параметр	Резервтік көшірмені пайдалану	Кілттерді бірнеше жазылушыларға сену	Құпияны криптографиялық бөлу хаттамаларын қолдану
Қарапайымдылық	+	+	–
Сенімділік	+	–	+
Сенімділік	–	–	+
Құпиялылық	–	–	+

1-кестедегі деректерді талдау негізінде құпия кілт жоғалған жағдайда ақпараттың қолжетімділігі мен құпиялылығын қамтамасыз етудің ең тиімді жолы деректерді таратылған сақтау үшін қолданылатын КҚОПП пайдалану болып табылатынын көруге болады.

КҚОПП жұмыс істеу принципін түсіну үшін құпия бөлісу хаттамаларын қолданудың демонстрациялық мысалын қарастырайық. Ақпараттық жүйенің (АЖ) ақпараттық қауіпсіздік әкімшісі төрт қызметкер арасында он екі таңбадан

тұратын құпия кілт бөлісуі керек делік. Бұл қандай да бір себептермен АЖ әкімшісі болмаған жағдайда қызметкерлер АЖ-ға жылдам қол жеткізуі және қажетті шұғыл операцияларды орындауы үшін қажет.

Бұл жағдайда АЖ әкімшісі дилер (бөлуші), ал қызметкерлер кастодиан болып табылады. Сипатталған технологияға (тәсілге) сандық мысал келтірейік. Құпия $z@far2190429$ таңбалар тізбегі болсын. Құпия бөлісуі: 1-ші қызметкер үшін «429» белгілері және № 4 лауазымы; 2-ші қызметкерге «ar2» таңбалары тобы үшін № 2 лауазымы; 3-ші қызметкерге "z@f" белгілері және № 1 лауазымы; 4-ші қызметкер үшін «190» белгілері және № 3 лауазымы. Бұл тәсілдің елеулі кемшілігі — құпияның бөліктерін қызметкерлер ашық түрде сақтайды. Бұл факт ақпараттың құпиялылық деңгейіне айтарлықтай әсер етеді.

Осы кемшілікке байланысты қарастырылған құпияны бөлісу технологиясы іс жүзінде «жұмыс» мақсаттары үшін пайдаланылмайды. Оны жақсарту негізінен, мысалы, криптографиялық түрлендірулерді қолдану арқылы мүмкін. Бұл технология КҚОПП деп аталады. Дегенмен, ақпараттық технологиялардың қарқынды дамуына байланысты КҚОПП қауіпсіздігі төмендеді, өйткені осы хаттамаларды бұзуға бағытталған компьютерлік шабуылдардың жаңа түрлері пайда болды. Атап айтқанда, бұл Positive Technologies жүргізген талдау нәтижелеріне сәйкес, ақпараттық жүйелерде құпия сөздерді сақтау кезінде көп жағдайда оларды жергілікті сақтау/қалпына келтіру үшін қорғаныс механизмдері қолданылмайды. Сондықтан оларды жетілдіру мақсатында КҚОПП талдау өте өзекті міндет деп айтуға болады.

Материалдар мен зерттеу әдістері

Кез келген криптографиялық хаттама ақпараттық процестерде криптографиялық түрлендірулер мен алгоритмдерді қолдануды реттейтін нақты ережелер жиынтығына негізделген. КҚОПП шекті құпияны бөлу сұлбасы (ҚБС) алгоритміне негізделуі мүмкін.

Типтік жағдайларда КҚОПП екі негізгі фазаны қамтиды.

1. Құпия бөлу – М құпиясын білетін дилер (бөлуші) c_1, c_2, \dots, c_n құпияның n үлесін тудыратын және қорғалған байланыс арнасы арқылы әрбір қатысушыға оның үлесін беретін тарату кезеңі. Тарату заңды абоненттер бірлескен іс – әрекеттер кезінде ғана құпияны қалпына келтіре алатындай, ал заңсыз абоненттер мүмкін болмайтындай етіп ұйымдастырылады.

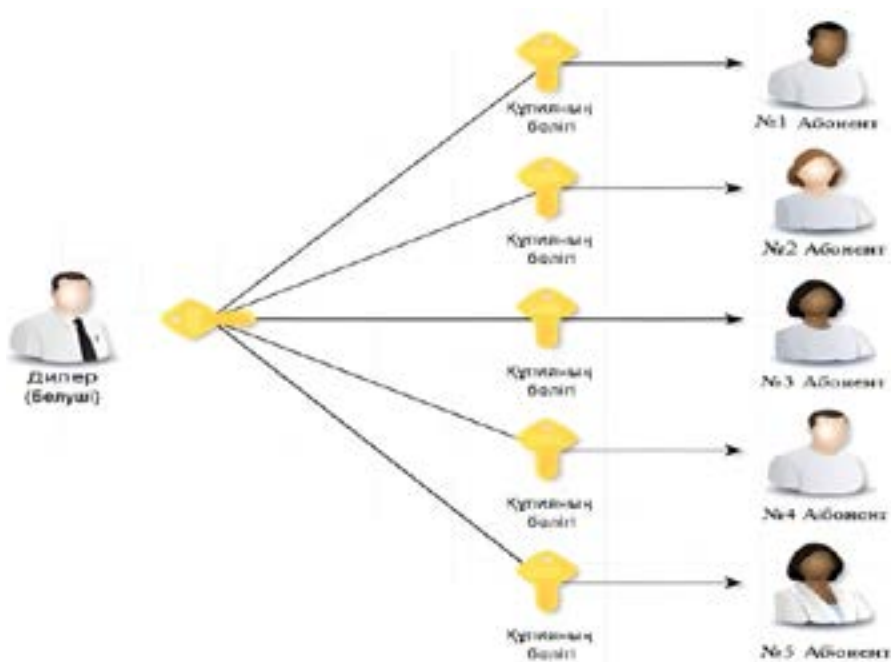
2. Құпияны қалпына келтіру-заңды абоненттер өздерінің құпия үлестерін біріктіріп, құпияны ала алатын кезең. Бұдан әрі қарастырылатын алгоритмдердің көпшілігінде "құпия" ақпарат бөлінген барлық заңды абоненттерді қалпына келтіруге міндетті қатысу қажет.

Суретте құпияны бөлудің шекті сұлбасы көрсетілген.

Зерттеу барысында келесі шекті ҚБС қарастырылды:

- Шамир ҚБС;
- Блэкли ҚБС;

- Эллиптикалық қисыққа негізделген ҚБС;
- Карнин-Грин-Хеллман Қ БС ;
- Асмут – Блум ҚБС.



Сур.1. Классикалық құпияны бөлісу хаттамасының сұлбасы
(Fig.1. Outline of the classic secret sharing protocol)

Ең тиімді КҚОПП анықтау үшін алгоритмдерді қолданудың қауіпсіздік деңгейіне әсер ететін келесі негізгі параметрлер бойынша олардың салыстырмалы талдауын жүргізу туралы шешім қабылданды:

- 1) есептеулердің күрделілігі. Алгоритмнің күрделілігін бағалау құпияны бөлісу және қалпына келтіру кезеңінде алынған бағалаулардан тұрады;
- 2) есептеу ресурсының қарқындылығы (құпияны ортақ пайдалану және қалпына келтіру кезеңдерінде пайдаланылатын жад көлемі);
- 3) кемелдік (егер заңсыз пайдаланушылардың кез-келген саны құпия туралы ешқандай ақпарат ала алмаса, ҚБС мінсіз болады);
- 4) идеалдылық (егер құпияның үлесінің мөлшері құпияның өзіне тең болса, ҚБС өте қолайлы).

Бұл жұмыста келесі белгілер қолданылады:

k – құпияны қалпына келтіру үшін қажетті заңды абоненттердің ең аз саны;

n – құпия бөлінетін үлестер саны;

p – үлкен жай Сан;

Z_p – бүтін сандардың қарапайым сақина Модулінің өлшемі;

M – құпия (кілт).

Шамир құпиясының бөлу сұлбасы. Шамир шекті сұлбасы (k, n) көпмүшелік интерполяция ұғымы төңірегінде құрылған. Егер құпияны тек k абонент қалпына келтіре алатындай етіп бөлу қажет болса, онда оны $(k-1)$ дәрежелі көпмүшелік формуласында «жасыру» керек. Бұл көпмүше k нүктелері бойынша қалпына келтіріледі. Есептеулердің күрделілігін талдап көрейік.

Құпияны бөлу кезеңі.

1-қадам. Бұл қадамда p кездейсоқ жай саны таңдалады. Санның қарапайымдылығын тексеру ресурсты көп қажет ететін процесс және алгоритмнің күрделілігін жалпы бағалауға айтарлықтай әсер етеді. Бұл қадамды бағалау санның қарапайымдылығын тексеру үшін қолданылатын алгоритмге байланысты. Қарапайымдылығын тексеру үшін

кездейсоқ таңдалған сан Бей және – Померанц – Селфридж – Вагстафф ықтималдық сынағы пайдаланылды (Cheruyakov, 2016). Бұл қадамның күрделілігі $O(1)$ – (Молдовьянғ 2011) ден алынған орташа мәнге тең.

2-қадам. Өріс үстінде көпмүшені құру алгоритмінің осы қадамында

$Z_p, (k-1)$ коэффициенттерін тандайды. Қадамның күрделілігін бағалау $O(k)$ тең.

3-қадам. Көлеңкелерді есептеуге арналған қайталанулар саны n -ге тең, әрбір итерация $(k-1)$ координаттар бойынша өтетін кірістірілген циклды қамтиды. Берілген қадамның жалпы бағасы $O(k \cdot n)$.

$k < n < k$ болғандықтан, құпияны бөлу үшін есептеулердің күрделілігін бағалау $O(k \cdot n)$ болады. Құпияны қалпына келтіру кезеңі. Құпияны қалпына келтіру процесі Лагранж интерполяциялық көпмүшесін құру арқылы жүзеге асырылады. Алгоритмнің күрделілігінің жалпы бағалау $O(k \cdot n) + O(k^2)$.

Есептеулердің ресурс сыйымдылығын талдау. Құпия фракцияларды сақтауға арналған жедел жад құрылғысының қажетті жады мөлшері $n \cdot |M| + O(|m|)$ мәніне тең, мұндағы

$|m|$ – құпияның максималды ұзындығы M . құпияны бөлу/қалпына келтіру үшін

$n = k = 64$ кезінде шамамен 8192 байт жедел жады қажет болады.

Кемелдік/идеалдық. Шамирдің сұлбасы кемелді және идеалды. Идеалдылық құпияның өлшемі p өлшеміне, сондай-ақ әрбір қатысушының құқығы бар құпияның үлесіне тең болуынан туындайды. Шамир сұлбасындағы құпия сызықтық теңдеулер жүйесін шешу арқылы қалпына келтірілді делік. Заңсыз жазылушылар k белгісізі бар k кем теңдеулер жүйесін құруы керек. Мұндай жүйенің шешімі k -өлшемді кеңістіктегі гипержазықтықта жатқан нүктелер жиыны болып табылады, бұл ешқандай құпия мәнді мүмкін емес деп жоққа шығаруға болмайтынын білдіреді. Сондықтан Шамирдің сұлбасы мінсіз.

Блэклидің құпия бөлісу сұлбасы. Блэкли сұлбасы немесе векторлық ҚБС көп өлшемді кеңістікте нүктелерді пайдалануға негізделген. Кез келген екі немесе одан да көп жазықтық кеңістікте қиылысады, ал қиылысу нүктесінің

координаттарының бірі құпия болып табылады. Егер құпия нүктенің бірнеше координаттары ретінде кодталса, онда бір гипержазықтықтан құпия туралы, яғни қиылысу нүктесінің координаттарының өзара тәуелділігі туралы кейбір мәліметтерді алуға болады. Есептеулердің күрделілігін талдап көрейік.

Құпия бөлісу кезеңі.

1-қадам. Шамир сұлбасындағыдай, бұл қадамның күрделілік бағасы санның қарапайымдылығын тексеру алгоритміне байланысты және $O(1)$ болады.

2-қадам. $(k-1)$ сандарды таңдау кезіндегі бұл қадамның күрделілігі $O(k)$ болады.

3-қадам. n қатысушының әрқайсысы үшін d_i коэффициенті анықталады және әрбір итерацияда кездейсоқ құрылған k санының жиынтығы қажет. Қадамның есептеу күрделілігін бағалау $O(k \cdot n)$ болады.

4-қадам. Шамир сұлбасындағыдай, құпияның үлесін n қатысушымен бөлісу үшін n қайталану қажет. $O(n)$ уақытында қатысушыларға құпияның үлестері беріледі.

Қалпына келтіру кезеңі. Құпияны қалпына келтіру тапсырмасы сызықтық теңдеулер жүйесін шешу арқылы жүзеге асырылады. Мұндай шешімнің тиімді нұсқасы Крамер әдісін қолдану болып табылады, өйткені құпия шешім нәтижесінде алынған нүктенің бірінші координатасы болып табылады.

Құпияны қалпына келтіру үшін $k \cdot k$ өлшемі бар матрицалардың екі анықтаушысын есептеу керек.

Матрицалық анықтауыштар Гаусс әдісі негізінде анықталады, ал күрделілік бағасы $O(k^3)$.

Блэкли тізбегінің есептеу күрделілігінің жалпы бағасы $O(k \cdot n) + O(k^3)$ болып табылады.

Есептеулердің ресурс қарқындылығын талдау. Құпияны үлестерге бөлуге қажетті жедел жадының байт саны $n \cdot k \cdot |m|$ ретінде бағаланады. $n = k = 64$ құпия бөлісу/қалпына келтіру операциялары үшін шамамен 266 Кбайт жедел жады қажет болады.

Кемелдік/идеалдық. Құпиядағы әрбір үлестің мөлшері құпияның өлшемінен k есе көп болғандықтан, Блэклидің сұлбасы мінсіз болуы мүмкін емес. Дегенмен, бұл кемелді, өйткені k белгісізі бар $(k-1)$ сызықтық салыстырулар жүйесінің шешімі k өлшемді кеңістіктегі гипержазықтықта жатқан шешімдер жиыны болып табылады. Бұл құпия M мүмкін болатын құпиялар жиынынан кез келген мәнді қабылдай алатынын білдіреді.

Эллиптикалық қисыққа негізделген құпия бөлісу сұлбасы. Эллиптикалық қисық бойынша құпияның бөлінуі төменде сипатталған алгоритміне сәйкес жүреді. Есептеулердің күрделілігін талдап көрейік.

Бөлу кезеңі.

1-қадам. Дилер қажетті нүктелер саны (кемінде n) бар ЕС эллиптикалық қисығын таңдайды. ҚБС қатысушыларының әрқайсысына (соның ішінде

құпия сақтаушы) эллиптикалық қисықтағы нүкте, оның ішінде «шексіз қашықтағы» нүкте беріледі.

2-қадам. Бұл қадамда дилер осы қисық бойынша n дәрежелі көпмүшені таңдайды. Бұл көпмүшенің коэффициенттері оған ғана белгілі. Эллиптикалық қисықтағы қатысушы — құпияны сақтаушыны білдіретін нүкте бәріне белгілі.

3-қадам. Дилер осы нүктенің координаталарын өзі таңдаған көпмүшеге ауыстырады, құпияның мәнін есептейді.

4-қадам. Әрбір қатысушыға құпияның өз үлесін беру үшін дилер ол үшін құпия үлесін ала отырып, қатысушы нүктесінің координаталарын көпмүшеге ауыстырады. Нәтижесінде қатысушының эллиптикалық қисықтағы нүктесі (ID) және құпияның үлесі (Secret) болады.

Эллиптикалық қисықтарға негізделген сұлбалардағы құпияны бөлісу фазасының есептеу күрделілігінің жалпы мағынасы $O(k \cdot n)$ -ге тең.

Қалпына келтіру кезеңі. Құпияны қалпына келтіру үшін дилер таңдаған көпмүшенің коэффициенттерін қалпына келтіру үшін бірнеше қатысушылар бірігуі керек. Математикалық тұрғыдан бұл кейбір теңдеулер жүйесін шешуге дейін азайтады. Рұқсат етілген коалицияны құрайтын қатысушылар қажетті көпмүшені алады. Олар оған құпияны білдіретін нүктенің координаталарын қояды. Нәтижесінде олар дилер қалыптастырған құпияны алады. Бұл сұлбадағы құпия қалпына келтіру кезеңінің есептеулерінің есептік күрделілігі $O(k^2)$.

Есептеулердің ресурс қарқындылығын талдау. Эллиптикалық қисық сызықтағы нүктенің оперативті жады мөлшері құпияның өзін сақтауға қажетті мөлшерден аспайтындықтан, сұлба Шамир сұлбасының ресурс қарқындылығын сақтап алады.

Құпияны қалпына келтіру үшін сізге шамамен 25 Кбайт жад қажет.

Кемелдік/идеалдық. Эллиптикалық қисықтарға негізделген сұлбалар кемелді болып табылады, өйткені заңсыз пайдаланушылардың құпия фракцияларында құпия туралы ешқандай ақпарат болмайды. Алайда, олар идеалды емес, өйткені құпияның әрбір үлесінің мөлшері құпиядан k есе көп.

Карнин-Грин-Хеллман құпиясымен бөлісу. Бұл сұлба алгебралық теңдеулер жүйесін шешуге негізделген. Есептеулердің күрделілігін талдап көрейік.

Бөлу кезеңі.

Құпияны n түрлі тарап (топ мүшелері) арасында бөлісу үшін, кем дегенде k тарап оны қалпына келтіре алатындай, k өлшемінің $(n+1)$ V_i векторлары таңдалады, сонымен қатар берілген k вектордан тұратын кез келген матрицаның рангі k -ға тең болуы қажет. V_0 векторы барлық қатысушыларға белгілі. Құпия ішкі туынды болып табылады (u, V_0) , мұндағы u — векторлар жиыны, ал бөліктері - ішкі туындылар (u, V_i) . Құпияны n бөлікке бөлу қадамының күрделілігі $O(n)$ -ге тең.

Қалпына келтіру кезеңі. Белгілі үлестер бойынша құпияны қалпына келтіру үшін u векторын табу үшін k теңдеулерінің жүйесі шешіледі.

Есептеулердің ресурс сыйымдылығын талдау. Құпия 2 вектордың матрицалық көбейтіндісі түрінде ұсынылғандықтан, векторлардың координаттарын сақтауға арналған жедел жад көлемі 64 байтқа тең. $N = 64$ кезінде құпияны бөліктерге бөлу үшін сізге $(2 \cdot n + 1) \cdot 64 = 8256$ байт жедел жады қажет етеді.

Құпияны қалпына келтіру үшін сызықтық алгебралық теңдеулер жүйесін шешу қажет. Бүтін арифметиканың ең жақсы нұсқасы Крамер әдісі болып табылады. Мәнді есептеу үшін $k \times k$ өлшемі бар матрицалардың $(k+1)$ анықтауыштарын алу керек. Нәтижесінде құпияны қалпына келтіру үшін 8320 байт жедел жады қажет болады.

Кемелдік/идеалдық. Карнин-Грин-Хеллман сұлбасы кемелді болып табылады, өйткені құпия M ықтимал құпиялар жиынтығынан кез келген мәнді қабылдай алады. Алайда, ол эллиптикалық қисық сызықтарға негізделген сұлба сияқты идеалды емес, өйткені құпияның әрбір үлесінің мөлшері құпияның өзінен k есе көп.

Асмут-Блум құпиясын бөлісу сұлбасы. Асмут-Блум сұлбасы — жай сандарды пайдаланып құрастырылған шекті құпияны бөлісу сұлбасы. Құпияны кез келген k қатысушы қалпына келтіре алатындай етіп n тарап арасында бөлісуге мүмкіндік береді. Есептеулердің күрделілігін талдап көрейік.

Бөлу кезеңі.

1-қадам. (k, n) шекті сұлба үшін p жай саны таңдалады.

2-қадам. Содан кейін Асмут-Блум шарттары орындалатын $p - d_1, d_2, \dots, d_n$ мәнінен кіші сандар таңдалады.

Құпия бөлісу кезеңінің есептік күрделілігі $O(n)$ болып табылады.

Қалпына келтіру кезеңі. Қытайлық қалдық теоремасын пайдалана отырып, кез келген k құпия үлестерді біріктіру арқылы құпияны қалпына келтіруге болады, бірақ бұл кез келген $(k-1)$ құпия үлестермен мүмкін емес (Kiparisova, 2016). Құпия қалпына келтіру кезеңінің есептеулерінің күрделілігі $O(k^2)$ -ге тең.

Есептеулердің ресурс қарқындылығын талдау. Әрбір жай сан d_i 100 байт оперативті жадты алады. Содан кейін d_j табу шарттарын тексеру үшін $2 \cdot k |d_i|$ есте сақтау, яғни. шамамен 12,8 КБайт. Жадта құпияның бір үлесін сақтау үшін $p = 28$ байт, $k = 64$ байт, $|p| + |d_i| + |k_i|$ шамамен 192 байтты құрайды. Нәтижесінде құпияны бөлісу үшін 57 Кбайт жедел жады қажет болады. Құпияны қалпына келтіру үшін шамамен 320 байт қажет.

Кемелдік/идеалдық. Асмут-Блум сұлбасы кемелді, себебі M құпиясы мүмкін болатын құпиялар жиынтығынан кез келген мәнді қабылдай алады. Бірақ бұл да идеалды емес, өйткені құпияның әрбір үлесі құпияның өзінен k есе көп.

Осы уақытқа дейін мақалада ресурстардың қарқындылығы мен есептеу күрделілігі тұрғысынан шекті ҚБС талданды. Сондай-ақ, осы сұлбаларға

шабуыл түрлерін қарастыру маңызды, өйткені ғылыми-техникалық прогреске байланысты бұзушылар шабуылдардың барған сайын жетілдірілген түрлерін ойлап табады (қолданады).

Шекті құпияны бөлісу сұлбаларына шабуыл түрлері. Егер зиянкес құпия бөлісуге қатысушылардың k санына еніп кетсе, шекті ҚБС-на компьютерлік шабуыл жасалуы мүмкін.

Зиянкестің шекті сұлбаны айналып өтуге көптеген әлеуетті мүмкіндіктері бар. Атап айтқанда, мынаны қарастырып өтеміз.

1. Шабуылдаушы құпияның дұрыс емес бөлігін (мысалы, ерікті санды) әдейі пайдалана алады - бұл жағдайда топ құпияны қалпына келтіре алмайды. Дегенмен, дұрыс емес бөлікті кім ұсынғанын анықтау мүмкін емес.

2. Егер шабуылдаушы өзін қатысушы етіп көрсете алса, құпия бөлісу процедурасының басталуына себеп болуы мүмкін. Содан кейін ол қалған қатысушылардың құпиясының үлесін ала алады.

3) $(k; n)$ шекті сұлбада құқық бұзушы өзін $(k+1)$ қатысушы ретінде көрсете алады. Құпияны қалпына келтіру үшін k қатысушы жеткілікті болғандықтан, шабуылдаушы құпияның өз үлесі ретінде таңбалардың ерікті тізбегін ұсына алады.

Бұл жағдайда бұзушы басқа заңды абоненттердің құпиясының бөліктерін біліп, содан кейін құпияны толығымен қайта жасай алады. Бұл кейбір ҚБС-ның жетілмегендігінің салдары.

Сондай-ақ, бұзушы жұмыс уақытынан, кәштеуден, қосымшаның істен шығуынан және т.б. пайдалы ақпаратты алуға тырысқанда, сыртқы арналар арқылы ҚБС-на компьютерлік шабуылдардың қаупі бар. Мұндай шабуылдарды жүргізу мүмкіндігі әдетте бағдарламалық жасақтаманы әзірлеу кезінде жіберілген қателермен байланысты.

Нәтижелер

Жүргізілген зерттеу нәтижелері КҚОПП қолдану негізгі ақпаратты сақтаудың/қалпына келтірудің ең тиімді әдісі екенін көрсетті. Сондай-ақ, КҚОПП ақпараттық қауіпсіздік деңгейіне айтарлықтай әсер ететін компьютерлік шабуылдардың заманауи түрлеріне бейім екендігі анықталды.

2-кестеде құпияны бөлудің шекті сұлбаларын салыстырмалы талдау нәтижелерінің қысқаша мазмұны көрсетілген.

Кесте 2. Шекті ҚБС салыстырмалы талдау нәтижелері

Шекті сұлба	Кемелдік	Идеалды	Ресурс сыйымдылығы (Кбайт)	Қиындықты бағалау
Шамир сұлбасы	+	+	8	$O(n \cdot k) + O(k^2)$
Блэкли сұлбасы	+	-	266	$O(n \cdot k) + O(k^3)$
Эллиптикалық қисыққа негізделген сұлбасы	+	-	25	$O(n \cdot k) + O(k^2)$
Карнин-Грин-Хеллман сұлбасы	+	-	8,1	$O(n) + O(k^3)$
Асмут – Блум сұлбасы	+	-	57	$O(n) + O(k^2)$

2-кестеден Шамир сұлбасы кемелдік пен идеалдылық қасиеттеріне ие, ол басқа сұлбалармен салыстырғанда ресурстарды аз қажет етеді деген қорытынды жасауға болады. Есептеулердің күрделілігі бойынша Шамир сұлбасы Асмут-Блум сұлбасынан төмен. Алайда, бұл жағдайда басқа параметрлердің артықшылығы шешуші болып табылады.

Қорытынды

Шекті құпияны бөлісу сұлбаларын салыстырмалы талдау нәтижелері Шамир сұлбасы ҚБС көрсеткіштерінің кешені бойынша қалғандарымен салыстырғанда ең тиімді болып табылатындығын көрсетеді. Сондықтан бұл сұлбаны қорғалған КҚОП-ті кеңінен дамыту үшін негіз ретінде қолданған жөн. Осы зерттеу нәтижесінде қол жеткізуді басқару жүйесі үшін шекті криптографияны пайдаланып, құпия кілттерді бөлудің жаңа әдісін әзірлеу, ал ҚБС ретінде Шамир сұлбасын пайдалану жоспарланып отыр.

REFERENCES

Alferov A.P. et al., 2005. *Osnovy kriptografii : uchebnoe posobie* [Basics of cryptography : Tutorial]. 3-rd ed., rev. and add. Moscow, Gelios ARV Publ., 2005. 480 p. <http://bookshare.net/index.php?idl=4&category=cryptography&author=alferov-ap&book=2002>.

Abbasov A.E., Abbasov T.E., 2015. *Otsenka kachestva programmnogo obespecheniya dlya sovremennykh sistem obrabotki informatsii* [Evaluation of software quality for modern information processing systems]. *Informatsionno- tekhnologicheskii vestnik* [Information Technology Bulletin], 2015. Vol. 5. № 3. Pp. 15–27. <https://www.elibrary.ru/item.asp?id=25360733>.

Alekseychuk A.N., 2005. *Sovershennyye skhemy razdeleniya sekreta i konechnyye universalnyye algebrы* [Perfect secret separation schemes and finite universal algebras]. *Analiz i obrabotka dannykh* [Data Analysis and Processing], 2005. <http://dSPACE.nbuV.gov.ua/bitstream/handle/123456789/50768/08-Alekseychuk.pdf?sequence=1>

Bozkurt I.N., Selcuk G., 2008. *Threshold cryptography based on blakely secret sharing*. *Information Sciences*, 2008. Pp. 1–4. <https://scholar.google.com/citations?user=0uolARoAAAAJ&hl=ru>

Blakley G.R. et al., 1979. *Safeguarding cryptographic keys*. *Proceedings of the national computer conference*, 1979. Vol. 48. Pp. 313–317. <https://ieeexplore.ieee.org/document/8817296>.

Chervyakov N.I., Deryabin M.A., 2016. *Novyy metod porogovogo razdeleniya sekreta, osnovanny na sisteme ostat- ochnykh klassov* [New method of threshold secret separation, based on the system of residual classes]. *Informatsionnye tekhnologii* [Information Technologies], 2016. Vol. 22. № 3. Pp. 211–219. http://novtex.ru/IT/it2016/it316_web-211-219.pdf.

Dingyi P., Arto S., Cunsheng D., 1996. *Chinese remainder theorem: applications in computing, coding, cryptography*. World Scientific, 1996. https://books.google.com/books/about/Chinese_Remainder_Theorem_Applications_I.html?id=RQLtCgAAQBAJ

Ito M., Saito A., Nishizeki T., 1989. *Secret sharing scheme realizing general access structure*. *Electronics and Com- munications in Japan (Part III: Fundamental Electronic Science)*, 1989. Vol. 72. № 9. Pp. 56–64. <https://archiv.infsec.ethz.ch/education/as09/secsem/papers/ItSaNi87.pdf>

Karnin E., Greene J., Hellman M., 1983. *On secret sharing systems*. *IEEE Transactions on Information Theory*, 1983. Vol. 29, № 1. Pp. 35–41. <https://ieeexplore.ieee.org/document/1056621>

Kiparisova A.I., Azhmuhamedov I.M., 2016. *Providing access to information systems of higher education in the case of loss of key information*. *Mezhdunarodnyy nauchno-issledovatel'skiy zhurnal* [International Research Journal], 2016. № 2–2 (44). <https://cyberleninka.ru/article/n/providing-access-to-information-systems-of-higher-education-in-the-case-of-loss-of-key-information>

Lavrinenko A.N., Chervyakov N.I., 2014. *Nekotorye elementy kontseptsii aktivnoy bezopasnosti v sovremennoy kriptografii* [Some elements of the concept of active security in modern cryptography]. *Nauchnye vedomosti Belgorodskogo gosudarstvennogo universiteta. Seriya: Ekonomika. Informatika*

[Scientific News of Belgorod State University. Series: Economy. Computer science], 2014. Vol. 30. № 8–1 (179). <https://cyberleninka.ru/article/n/nekotorye-elementy-kontseptsii-aktivnoy-bezopasnosti-v-sovremennoy-kriptografii>

Medvedev N.V., Titov S.S., 2011. Pochti porogovye skhemy razdeleniya sekreta na ellipticheskikh krivyykh [Almost threshold secret separation schemes on elliptic curves]. Doklady Tomskogo gosudarstvennogo universiteta sistem upravleniya i radioelektroniki [Reports of Tomsk State University of Control Systems and Radioelectronics], 2011. № 1 (23). <https://cyberleninka.ru/article/n/pochti-porogovye-shemy-razdeleniya-sekreta-na-ellipticheskikh-krivyykh>

Moldovyan A.A. et al., 2011. Protokoly s nulevym razglasheniem sekreta i obosnovanie bezopasnosti skhem tsifrovoy podpisi [Protocols with zero disclosure of the secret and the justification of the security of digital signature schemes]. Voprosy zashchity informatsii [Information Security Issues], 2011. № 4. Pp. 6–11. <https://www.elibrary.ru/item.asp?id=17100299>

Mogilevskaya N.S., Kulbikayan R.V., Zhuravlev L.A., 2011. Porogovoe razdelenie faylov na osnove bitovykh masok: ideya i vozmozhnoe primenenie [Threshold file sharing based on bit-masks: idea and possible use]. Vestnik Donskogo gosudarstvennogo tekhnicheskogo universiteta [Bulletin of the Don State Technical University], 2011. Vol. 11. № 10. <https://www.vestnik-donstu.ru/jour/article/view/912>

Petrov A., 2017. Kompyuternaya bezopasnost. Kriptograficheskie metody zashchity [Computer security. Cryptographic methods of protection]. Moscow, Litres Publ., 2017. 114 p. <https://www.iprbookshop.ru/87998.html>

Piskova A.V., Korobeynikov A.G., 2015. Razrabotka algoritma elektronnoy tsifrovoy podpisi, osnovannogo na zadachakh faktorizatsii i diskretnogo logarifmirovaniya na ellipticheskikh krivyykh [Development of an electronic digital signature algorithm based on the problems of factorization and discrete logarithmization on elliptic curves]. Sbornik trudov IV Vserossiyskogo kongressa molodykh uchenyykh [Proceedings of the IV All-Russian Congress of Young Scientists]. St. Petersburg, ITMO University Publ., 2015. Pp. 322–326. <https://kmu.itmo.ru/file/download/360>

Parvatov N.G., 2008. Sovershennyye skhemy razdeleniya sekreta [Perfect secret sharing schemes]. Prikladnaya diskretnaya matematika [Applied Discrete Mathematics], 2008, № 2 (2). Pp. 41–47. https://www.mathnet.ru/php/archive.phtml?wshow=paper&jrnid=pdm&paperid=33&option_lang=rus

Shenets N.N., 2011. Ob idealnykh modulyarnykh skhemakh razdeleniya sekreta v koltsakh mnogochlenov ot neskol'kikh peremennykh [On the ideal modular secret separation schemes in polynomial rings of several variables], 2011. <http://elib.bsu.by/handle/123456789/9565>

Shnayer B., Fergyson N., 2005. Prakticheskaya kriptografiya [Practical cryptography]. Moscow, Dialektika, 2005. 480 p. <https://www.labirint.ru/books/599557/>

Sharyy S.P., 2012. Kurs vychislitelnykh metodov [The course of computational methods]. Novosibirsk, Novosibirsk State University Publ., 2012. <http://www.ict.nsc.ru/matmod/files/textbooks/SharyNuMeth.pdf>

Stadler M., 1996. Publicly verifiable secret sharing. International Conference on the Theory and Applications of Cryptographic Techniques. Springer, Berlin, Heidelberg, 1996. Pp. 190–199. <https://crypto.ethz.ch/publications/files/Stadler96.pdf>

Vashchenko G.V., 2009. Vychislitel'naya matematika. Osnovy algebraicheskoy i trigonometricheskoy interpolatsii [Computational Mathematics. Basics of algebraic and trigonometric interpolation]. Sovremennyye problemy nauki i obrazovaniya [Modern problems of Science and Education], 2009. № 1. Pp. 54–55. <https://edu-lib.com/matematika-2/dlya-studentov/vashhenko-g-v-vyichislitel'naya-matematika-osnovyi-algebraicheskoy-i-trigonometricheskoy-interpolatsii-onlayn>

NEWS OF THE NATIONAL ACADEMY OF SCIENCES OF THE REPUBLIC OF
KAZAKHSTAN
PHYSICO-MATHEMATICAL SERIES
ISSN 1991-346X
Volume 2. Number 346 (2023). 274–287
<https://doi.org/10.32014/2023.2518-1726.199>

УДК 004.931

© **G.B. Abdikerimova, A.A. Shekerbek*, M.G. Baibulova,
S.K. Abdikarimova, Sh.Sh. Zholdassova, 2023**

Eurasian National University named after L.N. Gumilyov, Astana, Kazakhstan.
E-mail: shekerbek80@mail.ru

CHEST PATHOLOGY DETERMINATION THROUGH AUTOCORRELATION FUNCTION

Abdikerimova Gulzira Bakhytbekovna — Associate Professor, Department of Information Systems, Eurasian National University named after L.N. Gumilyov, PhD, Astana, Kazakhstan
E-mail: gulzira1981@mail.ru. ORCID: 0000-0002-4953-0737;

Shekerbek Ainur Azimbaevna — Doctoral student of the Department of Information Systems of the Eurasian National University named after L.N. Gumilyov, Astana, Kazakhstan
E-mail: shekerbek80@mail.ru. ORCID: 0000-0002-1088-4239;

Baibulova Makbal Gabbasovna — Senior Lecturer, Department of Information Systems, Eurasian National University named after L.N. Gumilyov, Astana, Kazakhstan
E-mail: m.gabbasovnaa@gmail.com;

Abdikarimova Samal Korganbekovna — Teacher, Department of Informatics, Eurasian National University named after L.N. Gumilyov, Astana, Kazakhstan
E-mail: samala.enu.ask@gmail.com;

Zholdassova Sholpan Shorabekovna — Senior Lecturer at the Department of Artificial Intelligence Technologies, Eurasian National University named after L.N. Gumilyov, Astana, Kazakhstan
E-mail: zholdasova.sh@gmail.com.

Abstract. An important feature of image analysis is the visible texture in all images used in biomedical research. A chest x-ray is the most common and widely used method for diagnosing lung disease. It provides a quick and relatively accessible view of the internal structure of the chest, including the lungs, heart, ribs, and diaphragm. A chest x-ray can be used to detect various pathologies and conditions of the lungs, such as: pneumonia, pulmonary edema, lung tumors, emphysema, and tuberculosis. The chest is an object that can be depicted on an X-ray. A chest X-ray is a common method of diagnosing and imaging inside the chest. The correct division of the lungs into shapes and sizes is an important reason for diagnosis, thanks to which an intelligent information environment can be created. Despite the use of X-rays, the disease may go unnoticed to clarify the diagnosis. In this case, there is a risk of development, which can lead to death. The article deals with the problems of pneumonia clustering using the autocorrelation function to identify the

most accurate result. This provides a reliable tool for diagnosing lung radiographs. X-ray images are grayscale as well as their reduced contrast. Image pre-processing and data generation are important steps in the process of identifying and analyzing the underlying chest on X-rays. Therefore, images from two classes were selected for the task: healthy and with pneumonia. This article demonstrates the applicability of the autocorrelation function to detect interesting chest radiographs based on the level of detail of texture features and k-means extraction.

Keywords: chest radiograph, medical image processing, texture, clustering, pathology

© Г.Б. Абдикеримова, А.Ә. Шекербек*, М.Г. Байбулова,
С.К. Абдикаримова, Ш.Ш. Жолдасова, 2023

Л.Н. Гумилев атындағы Еуразия ұлттық университеті, Астана, Қазақстан.
E-mail: shekerbek80@mail.ru

КЕУДЕ ПАТОЛОГИЯСЫН АВТОКОРРЕЛЯЦИЯЛЫҚ ФУНКЦИЯ АРҚЫЛЫ АНЫҚТАУ

Шекербек Айнұр Әзімбайқызы — Л.Н. Гумилев атындағы Еуразия ұлттық университетінің ақпараттық жүйелер кафедрасының докторанты, Астана, Қазақстан
E-mail: shekerbek80@mail.ru. ORCID: 0000-0002-1088-4239;

Абдикеримова Гүлзира Бахытбекқызы — Л.Н. Гумилев атындағы Еуразия ұлттық университетінің ақпараттық жүйелер кафедрасының м.а.доценты, PhD, Астана, Қазақстан
E-mail: gulzira1981@mail.ru. ORCID: 0000-0002-4953-0737;

Байбулова Мақбал Габбасқызы — Л.Н. Гумилев атындағы Еуразия ұлттық университетінің ақпараттық жүйелер кафедрасының аға оқытушысы, Астана, Қазақстан
E-mail: m.gabbasovnaa@gmail.com;

Абдикаримова Самал Корганбековна — Л.Н. Гумилев атындағы Еуразия ұлттық университетінің информатика кафедрасының оқытушысы, Астана, Қазақстан
E-mail: samala.enu.ask@gmail.com;

Жолдасова Шолпан Шорабековна — Л.Н. Гумилев атындағы Еуразия ұлттық университетінің жасанды интеллект технологиялары кафедрасының аға оқытушысы, Астана, Қазақстан
E-mail: zholdasova.sh@gmail.com.

Аннотация. Кескінді талдаудың маңызды ерекшелігі биомедициналық зерттеулерде қолданылатын барлық кескіндердегі көрінетін текстура болып табылады. Кеуде қуысының рентгені — өкпе ауруын диагностикалаудың ең кең таралған және кеңінен қолданылатын әдісі. Ол кеуде қуысының ішкі құрылымын, оның ішінде өкпені, жүректі, қабырғаларды және диафрагманы жылдам және салыстырмалы түрде қолжетімді түрде көруге мүмкіндік береді. Кеуде қуысының рентгенографиясын өкпенің әртүрлі патологиялары мен жағдайларын анықтау үшін қолдануға болады, мысалы: пневмония, өкпе ісінуі, өкпе ісіктері, эмфизема және туберкулез. Кеуде қуысы — рентгенде бейнелеуге болатын нысан. Кеуде қуысының рентгені — кеуде қуысын диагностикалаудың және бейнелеудің кең таралған әдісі. Өкпенің пішіндер мен өлшемдерге дұрыс бөлінуі диагностиканың маңызды себебі болып

табылады, соның арқасында интеллектуалды ақпараттық ортаны құруға болады. Рентген сәулелерін қолдануға қарамастан, диагнозды нақтылау үшін ауру байқалмай қалуы мүмкін. Бұл жағдайда өлімге әкелуі мүмкін даму қаупі бар. Мақалада ең дәл нәтижені анықтау үшін автокорреляция функциясын пайдалана отырып, пневмонияны кластерлеу мәселелері қарастырылады. Бұл өкпе рентгенографиясын диагностикалаудың сенімді құралы болып табылады. Рентгендік кескіндер сұр реңкті, сондай-ақ олардың контрастын азайтады. Кескінді алдын ала өңдеу және деректерді генерациялау - рентген сәулелерінде кеуде қуысының астын анықтау және талдау процесіндегі маңызды қадамдар. Сондықтан тапсырма үшін екі сыныптың суреттері таңдалды: сау және пневмониямен. Бұл мақала текстура ерекшеліктерінің егжей-тегжейлі деңгейіне және к-құралдарын шығаруға негізделген қызықты кеуде рентгенограммаларын анықтау үшін автокорреляция функциясының қолданылуын көрсетеді.

Түйін сөздер: кеуде рентгенографиясы, медициналық кескінді өңдеу, текстура, кластерлеу, патология

© Г.Б. Абдикеримова, А.А. Шекербек*, М.Г. Байбулова,
С.К. Абдикаримова, Ш.Ш. Жолдасова, 2023

Евразийский национальный университет имени Л.Н. Гумилева,
Астана, Казахстан.

E-mail: shekerbek80@mail.ru

ОПРЕДЕЛЕНИЕ ГРУДНОЙ ПАТОЛОГИИ С ПОМОЩЬЮ ФУНКЦИИ АВТОКОРРЕЛЯЦИИ

Шекербек Айнура Азимбаевна — Докторант кафедры информационных систем Евразийского национального университета имени Л.Н. Гумилева, Астана, Казахстан

E-mail: shekerbek80@mail.ru. ORCID: 0000-0002-1088-4239;

Абдикеримова Гульзира Бахытбековна — и.о. доцент кафедры информационных систем Евразийского национального университета имени Л.Н. Гумилева, PhD, Астана, Казахстан

E-mail: gulzira1981@mail.ru. ORCID: 0000-0002-4953-0737;

Байбулова Макбал Габбасовна — старший преподаватель кафедры информационных систем Евразийского национального университета имени Л.Н. Гумилева, Астана, Казахстан

E-mail: m.gabbasovnaa@gmail.com;

Абдикаримова Самал Корганбековна — преподаватель кафедры информатики Евразийского национального университета имени Л.Н. Гумилева, Астана, Казахстан

E-mail: samala.enu.ask@gmail.com;

Жолдасова Шолпан Шорабековна — старший преподаватель кафедры технологии искусственного интеллекта Евразийского национального университета имени Л.Н. Гумилева, Астана, Казахстан

E-mail: zholdasova.sh@gmail.com.

Аннотация. Важной особенностью анализа изображений является видимая текстура на всех изображениях, используемых в биомедицинских исследованиях. Рентген грудной клетки является наиболее распространенным

и широко используемым методом диагностики заболеваний легких. Он предоставляет быстрое и относительно доступное изображение внутренней структуры грудной клетки, включая легкие, сердце, ребра и диафрагму. Рентген грудной клетки может использоваться для обнаружения различных патологий и состояний легких, таких как пневмония, отек легких, опухоли легких, эмфизема и туберкулез. Грудная клетка является объектом, который можно изобразить на рентгеновском снимке. Рентгенография грудной клетки является распространенным методом диагностики и образования изображений внутри грудной клетки. Правильное разделение легких на формы и размеры является важной причиной диагностики, благодаря которой может быть создана интеллектуальная информационная среда. Несмотря на использование рентгена, для уточнения диагноза заболевание может остаться незамеченным. В таком случае есть риск развития, которое может привести к летальному исходу. В статье рассматриваются проблемы кластеризации пневмоний с использованием функции автокорреляции для выявления максимально точного результата. Это обеспечивает надежный инструмент для диагностики рентгенограмм легких. Рентгеновские изображения являются полутоновыми, а также их пониженная контрастность. Предварительная обработка изображений и формирование данных являются важными шагами в процессе выявления и анализа основы грудной клетки на рентгеновских снимках. Поэтому для задачи были отобраны изображения из двух классов: здоровых и с пневмонией. В этой статье демонстрируется применимость функции автокорреляции для выявления интересных рентгенограмм грудной клетки на основе уровня детализации особенностей текстуры и извлечения к-средних.

Ключевые слова: рентгенограмма грудной клетки, обработка медицинских изображений, текстура, кластеризация, патология

Кіріспе

Радиациялық диагностика әдістерінің таралуымен рентгенологтардың жұмыс көлемі айтарлықтай өсті. Денсаулық сақтаудағы цифрландырудың үздіксіз процесі медицинаның әртүрлі салаларында көбірек жаңа технологияларды қолдануға мүмкіндік береді. Қазіргі уақытта Қазақстан Республикасының Үкіметі экономиканың барлық салалары мен салаларын цифрландыруға белсенді түрде қатысуда. 2019 жылғы 26 желтоқсанда Қазақстан Республикасы Денсаулық сақтау министрлігі халықтың өмір сүру сапасын жақсартуға бағытталған бірнеше индикаторларды қамтитын «Қазақстан Республикасының денсаулық сақтау саласын дамытудың 2020–2025 жылдарға арналған мемлекеттік бағдарламасын» таныстырды. Оған қол жеткізу 2025 жылға дейін жоспарланған. Бағдарламаның негізгі бағыттарының бірі аурулардың алдын алу, уақтылы көмек көрсету және ең жақсы мүмкіндіктерді ескере отырып, толық оңалту арқылы балалардың,

жасөспірімдердің және жастардың денсаулығын нығайту болады. ЮНИСЕФ-тің 2016–2030 жылдарға арналған денсаулық сақтау стратегиясындағы тәсілдерге сүйене отырып, халықаралық тәжірибе. Қазақстан Республикасында 2018 жылға жұқпалы аурулар бойынша эпидемиологиялық жағдай тұрақты. 34 жұқпалы және паразиттік аурулармен сырқаттанушылықтың төмендеуіне қол жеткізілді.

Зерттеу автокорреляциялық функцияларға негізделген радиографиялық кескіндерді автоматты түрде жіктеу үшін аспаптық ортаны әзірлеуге бағытталған. Жасалатын бағдарламалық қамтамасыз етудің мүмкіндіктерін іске асыру мақсатымен келесі міндеттерді шешу қажет болды: қолданыстағы әдістерді талдау және кескіндегі патологиясы бар біртекті аймақты ерекшеленуге арналған алгоритмдік-математикалық бағдарламалық қамтамасыз етуді әзірлеу.

Бұл жұмыста рентгендік кескінді өңдеу нәтижелері бойынша ауруларды неғұрлым дәл таңдау үшін автокорреляция функциясын пайдаланып кластерлеу мәселесін шешу ұсынылады. Тәжірибе нәтижесінде автокорреляция функциясы рентгендік суреттерде шағын біртекті аймақтарды анықтады.

Зерттеу материалы мен әдістері

Корреляциялық талдау кескіннің кейбір қасиеттері туралы, мысалы, координаталар бойымен қарқындылықтың өзгеру жылдамдығы туралы, гармоникалық құрамдас бөліктерге ыдыратпай біртекті қималардың ұзындығы туралы тәжірибеде түсінік алуға мүмкіндік береді. Корреляциялық талдаудың мәні әртүрлі сигналдардың ұқсастық дәрежесін сандық түрде өлшеу болып табылады. Бұл үшін мәні негізгі примитивтердің өлшемін сипаттайтын корреляциялық функциялар қолданылады, олар өз кезегінде текстураның біркелкілігін анықтайды және олар өз кезегінде, эксперименттер нәтижесінде патологиямен анықталады. Суреттердегі біртекті аймақтар өкпе лобтарымен байланысты текстуралық аймақтар болып саналады. Бұл тәсілде текстура кескіннің туынды емес элементтері тонның кеңістіктік өлшемімен байланысты (тонның туынды емес элементі — белгілі бір патологиялық белгілері бар кескін аймағы). Автокорреляция функциясының мәні патологияның туынды емес элементтерінің тонусының өлшемін сипаттайтын белгі ғана. Кеңістіктік орналасу корреляция коэффициентімен сипатталады, ол бір сурет элементінің жарықтылығының басқасының жарықтығына сызықтық тәуелділігінің өлшемі болып табылады (Харалик, 1979).

Статистикалық және ғаламдық өлшем ретінде қарастырылатын , автокорреляция функциясы кескіннің I талдау терезесінің көлденең және тік осьтері бойынша келесі теңдеу бойынша есептеледі:

$$R_{x,y}^{I(\alpha,\beta)} = \sum_{\alpha \in \Omega} \sum_{\beta \in \Omega} I(x, y) I(x + \alpha, y + \beta) = FFT^{-1}([FFT[I(x, y)]FFT^*[I(x, y)]])$$

мұндағы $I(x + \alpha, y + \beta) - I(x, y)$ кескінінің талдау терезесін Ω жазықтығында анықталған көлденең және тік осьтер бойымен сәйкесінше α және пикселдер

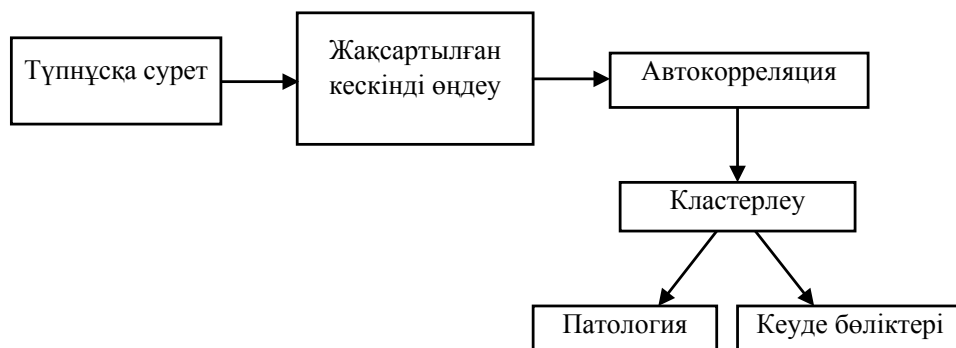
арқылы аудару. FFT, $(.)^*$, және $(.)^{-1}$ сәйкесінше жылдам Фурье түрлендіруін, күрделі конъюгатты және кері түрлендіруді білдіреді (Мехри және т.б., 2013).

Радиациялық диагностиканың суреттерін талдау. Сәулелік диагностика – аурулардың алдын алу және тану үшін қалыпты және патологиялық өзгерген адам ағзалары мен жүйелерінің құрылымы мен қызметін зерттеу үшін сәулені қолдану туралы ғылым. Денсаулық сақтау саласындағы мемлекеттік саясаттың негізгі бағыттарының бірі медициналық көмектің сапасын арттыру болып табылады. Заманауи медицинаның ақпараттық қолдауды, цифрлық технологияларды және телемедицинаны белсенді қолдануының өзектілігі халықты жоғары білікті медициналық көмекпен қамтамасыз ету қажеттілігімен түсіндіріледі. Медициналық қызмет көрсету сапасын арттыру шарттарының бірі электрондық денсаулық сақтау жүйесін енгізу. Радиациялық диагностикаға рентгендік диагностика, ультрадыбыстық диагностика, рентгендік компьютерлік томография, радионуклидті диагностика және магнитті-резонанстық томография жатады. Сонымен қатар, оған интервенциялық радиология қосылады. Рентгендік зерттеу әдістері — рентген сәулелерінің көмегімен мүшелерді зерттеу әдістері. Рентгендік диагностика әдісі рентген сәулелері үшін тіндердің әртүрлі өткізгіштігіне негізделген. Рентгендік әдістердің әрқайсысының артықшылықтары мен кемшіліктері бар, демек, диагностикалық мүмкіндіктердің белгілі бір шегі бар. Бірақ барлық рентгендік әдістер ақпараттың жоғары мазмұнымен, іске асырудың қарапайымдылығымен, қолжетімділігімен және бірін-бірі толықтыру мүмкіндігімен сипатталады. Рентгендік әдістер медициналық диагностикада жетекші орындардың бірін алады: 50%-дан астам жағдайда рентгендік диагностиканы қолданбай диагноз қою мүмкін емес. Ең жиі қолданылатын рентгендік диагностикалық әдістер — рентгенография, флюорография және рентгендік флюорография. Қазіргі уақытта пленкалы флюорография сандықпен ауыстырылуда. Қазіргі уақытта денсаулық сақтаудың ақпараттық ресурстары белсенді түрде әзірленуде. Ақпараттық технологиялардың үстемдік етуінің қазіргі жағдайында саланың мақсатты мемлекеттері «цифрлық медицина» және «цифрлық денсаулық сақтау» деп аталады. Электрондық денсаулық сақтау заманауи цифрлық технологияларды пайдаланады. Осының арқасында емдеу-диагностикалық процесс қазіргі уақытта диагностикалық және емдік ақпаратты алу және енгізу, есепке алу және есеп беру деректері саласындағы дамудың жаңа, жоғары технологиялық деңгейіне өтуде. Цифрлық технологиялар медициналық ақпаратпен қашықтықтан алмасу мүмкіндігін ашты.

Қазіргі уақытта рентген сәулелері флюорографиялық зерттеулер үшін ең қолайлы детектор болып табылады. Өкпенің рентгенограммасының құрылымындағы ақауларды тіркеу және анықтау сенімділігін арттыру қажеттілігі талданатын кескіндердің контрастын ашуды және сәйкестендіруді қиындататын факторларды жоюды, әдістердің ақпараттық мазмұнын

арттыруды, кескіндерді формада ұсынуды талап етеді. бұл оларды анықтауға, рентгендік кескіндерді цифрлық өңдеуге арналған тиімді алгоритмдер мен бағдарламаларды жасауға, аурудың сипатына байланысты мәселелерді шешуге ыңғайлы. Рентгендік кескін ондағы бұзылулардың аз болуымен сипатталады, бірақ сонымен бірге қарастырылатын орташа сипаттама, мысалы, флюорограмма, қарқындылық мәндеріндегі жергілікті өзгерістерді жақсы көрсетеді. ауру. Рентгендік кескінді бейнелеудің сандық түріне түрлендіру жағдайында алынған сандық массив сәулелену қарқындылығының берілген жазықтықтағы стохастикалық таралуы болып табылады және ақпаратты статистикалық талдау тұрғысынан диагностикалық есептерді шешу болып табылады. қазірдің өзінде мүмкін. Сандық кескін цифрлық матрица түрінде берілген, бұл сандық жолдар мен бағандар. Суреттерді көрсету үшін цифрлық матрица көрінетін кескін элементтерінің матрицасына – пиксельдерге айналады. Сандық матрицаның мәнінен кейін әрбір пикселде сұр реңктердің біріне ие болады.

Кеуде қуысының рентгенограммасында патологияларды анықтау алгоритмін қолдану. Бағдарлама Python тілінде жүзеге асырылды. 1-суреттегі алгоритм бойынша жүргізілген зерттеу жұмысы мәліметтер қоры кескіндерінің патологиясын анықтауға бағытталған. Есептеу барысында www.kaggle.com ашық деректер қорынан алынған рентген суреттер қарастырылды.

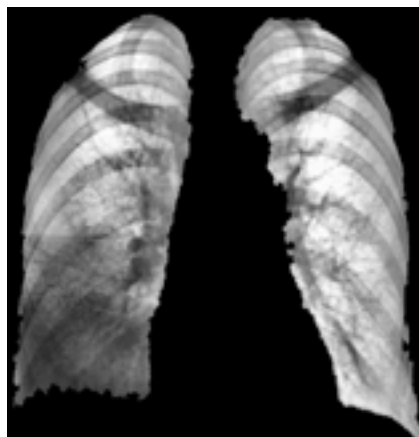


1-сурет. Кеуде қуысының патологиясын анықтау кезеңдері

Жұмыс барысында бастапқы кескіндер өңделді, яғни зерттеу үшін бастапқы суреттегі өкпенің таңдалған аймағы таңдалып, контраст күшейтілді (2 а, б сурет).



а



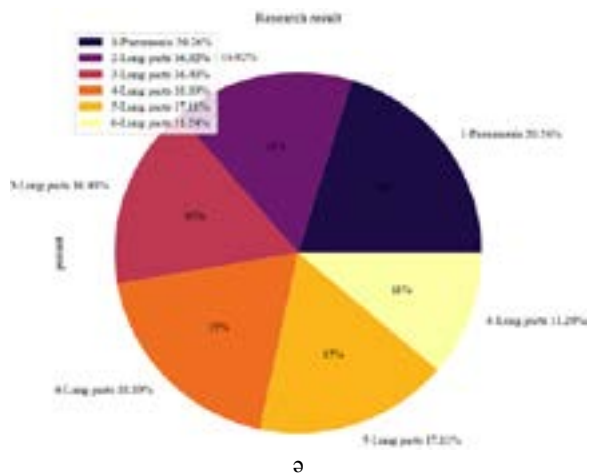
ә

2-сурет. Түпнұсқаның өңделген кескіні (а), алдын ала өңдеуден кейін контраст күшейтілді ә)

Әрбір суретте «аномальды» текстуралық аймақтар анықталды және олардың пайызы анықталды (3-сурет).



а



3-сурет. Контрасты күшейтілген кескінді кластерлеу нәтижесі (а), өкпені кластерлеу нәтижесінің пайызы (ә)

Автокорреляция функциясының өңдеу деректері мен рентгенологтардың қорытындыларының деректерін салыстыру кезінде бағдарлама мен дәрігерлердің нәтижелері толығымен сәйкес келетіні анықталды. Алынған нәтижелер қолданылатын әдістің жоғары диагностикалық дәлдігін, сондай-ақ радиологтың жұмысын автоматтандыру әдісін қолдану мүмкіндігін көрсетеді. Математикалық әдісті цифрлық кескінді өңдеуде матрица ретінде медициналық кескіндерге терең қолдану қарастырылды.

Нәтижелер мен пікірталас

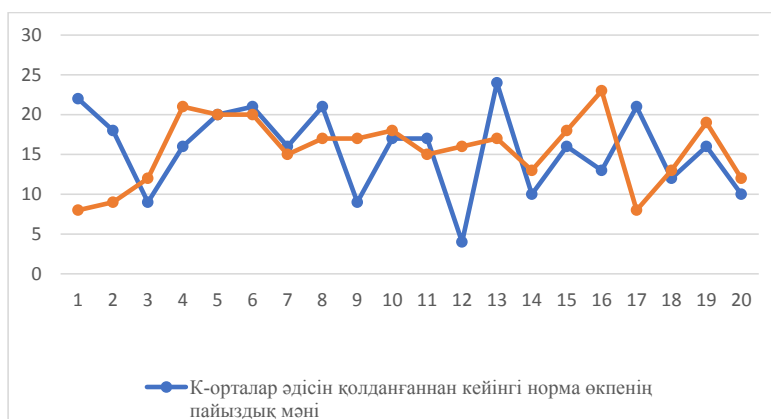
Зерттеу барысында ашық деректер базасынан 20 рентген суреті «норма» ретінде, ал кеуде қуысы патологиялары бар 20 рентген суреті қарастырылды. Оларға автокорреляция функциясы қолданылды және олардың мәндерінің нәтижелері 1-кестеде көрсетілген.

1-кесте Автокорреляция функциясын және k-орташа әдісін қолдану нәтижелерінің мағынасы

Суреттердің атауы	Автокорреляциялық функция әдісі	K- орташа мәндер әдісі
Normal-57.png	9	22
person76_bacteria_371.jpeg	21	8
Normal-61.png	15	18
person77_bacteria_377.jpeg	20	9
Normal-62.png	10	9
person80_virus_150.jpeg	25	12
Normal-64.png	17	16
person82_virus_154.jpeg	23	21
Normal-65.png	14	20
person83_virus_156.jpeg	21	20
Normal-69.png	13	21

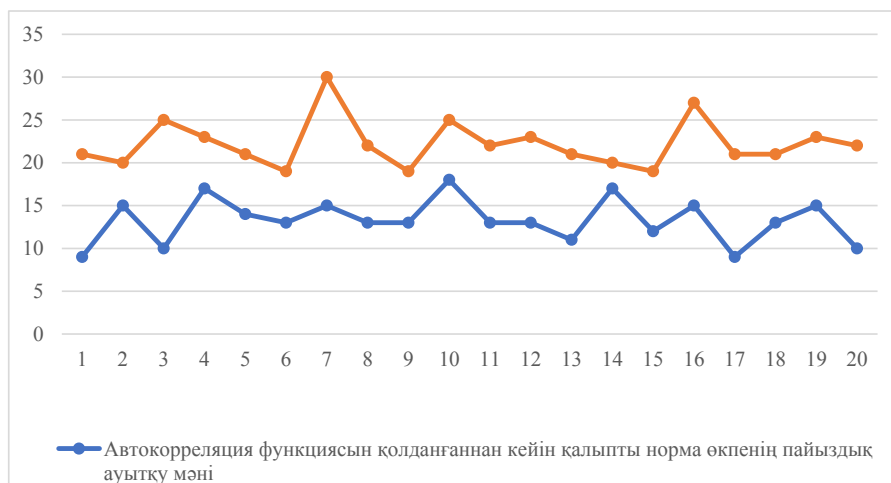
person88_virus_163.jpeg	19	20
Normal-72.png	15	16
person88_virus_165.jpeg	30	15
Normal-74.png	13	21
person89_virus_168.jpeg	22	17
Normal-76.png	13	21
person95_virus_177.jpeg	19	17
Normal-78.png	18	17
person96_virus_178.jpeg	25	18
Normal-80.png	13	17
Normal-81.png	13	4
person97_virus_181.jpeg	22	15
person98_virus_182.jpeg	23	16
person99_virus_183.jpeg	21	17
Normal-86.png	11	24
person100_virus_184.jpeg	20	13
Normal-87.png	17	10
Normal-91.png	12	16
person102_virus_189.jpeg	19	18
person105_virus_192.jpeg	27	23
person106_virus_194.jpeg	21	8
person106_virus_195.jpeg	21	13
Normal-91.png	15	13
Normal-91.png	9	21
person107_virus_196.jpeg	23	19
Normal-91.png	13	12
Normal-91.png	15	16
Normal-91.png	10	10
person109_virus_197.jpeg	22	12

Автокорреляциялық функция әдісінің тиімділігін көру үшін k-орталар әдісі қарастырылды, оның графигі төменде көрсетілген (4-сурет).



4-сурет. K-орталар әдісін қолдану нәтижесі

Төменде кеуде қуысының қалыпты жағдайының ауытқуының және автокорреляциялық функциялар әдісін қолдану нәтижесінде патологияның пайда болуының графигі берілген (5-сурет).



5-сурет. Норма мен патологияның автокорреляциялық функциясының пайыздық ауытқу мәндерінің графигі

Суретте көрсетілгендей, норма мен патологияның пайызы айтарлықтай айырмашылыққа ие, бұл норма мен патологияның нәтижесін анықтау шекарасын орнатуға сенімділіктің жоғары дәрежесіне мүмкіндік береді. Зерттеулер көрсеткендей, бұл шектеу 18% шамасында жатыр. Осы шектен жоғары нәтижелер патологиясы бар суреттерге сілтеме жасау ықтималдығы жоғары. Өкпе патологиясын анықтауда қолданылатын автокорреляциялық функцияның тиімділігін тексеру үшін k-means әдісі арқылы эксперименттер жүргізілді. Алынған мәндерге сүйене отырып, k-means әдісі 20 патологиялық суреттің 10 суретін қалыпты, яғни 44 % және 20 сау өкпенің 14 суретін патологиялық деп қате анықтады, яғни 60 %. Автокорреляция функциясын қолданғаннан кейінгі нәтижеден 20 патологиялық кескіндер үшін 99 % және қалыпты кескіндер үшін 100 % дәлдік көрсетті.

Қорытынды

Рентгенограммаларды алдын ала сегменттеу алгоритмдері мен автокорреляциялық функцияларды қолдануға негізделген әдістер патологияны тану дәлдігіне шамамен 98% жетуге мүмкіндік береді. Бұл дәлдік шағын жаттығулар жиынтығымен анықталады, сондықтан алдағы жұмыста рентгенограммалардың көбірек санына дифференциациялау жоспарлануда және нейрондық желі реттеледі. Осылайша, болашақта бұл әдіс ауруларды диагностикалау процесін жылдамдатады және қайталанатын зерттеулердің үлесін азайтады. Сипатталған зерттеу әдісін тексеру үшін 40

флюорографиялық сурет таңдалды. Суреттердің бір жартысында пневмонияға байланысты әртүрлі патологиялар болды, ал қалған жартысы сау өкпенің суреттері болды.

Рентгенологтар Қазақстан Республикасындағы денсаулық сақтау саласын цифрландырудың кейбір өзекті мәселелерін шешуде шешуші рөл атқарады, мысалы, оқыту үшін жоғары сапалы деректер жиынтығын құру, шешуді қажет ететін клиникалық мәселені анықтау және нәтижелерді түсіндіру. Рентгенолог тәжірибесінде жасанды интеллектті одан әрі енгізу үшін көптеген зерттеулер қажет, бірақ қазір автоматтандырылған медицина жүйесі жұмыс жүктемесінің бір бөлігін алып, дәрігердің жұмысын жеңілдетеді, сонымен қатар экономикалық жағдайды жақсартта алады деп айта аламыз. денсаулық сақтаудың ресурстық базасының шығындарын азайту арқылы.

ӘДЕБИЕТТЕР

Абдикеримова Г., Ламашева Ж., Байбулова М. и Токкулиева А., 2023. Классификация рентгеновских изображений с использованием алгоритма глубокого обучения. Известия Национальной академии наук Республики Казахстан. Физико-математическая серия, (1), 204-227. <https://journals.nauka-nanrk.kz/physics-mathematics/article/view/4876>

Арми Л. и Фекри-Эршад С., 2019. Анализ текстурных изображений и методы классификации текстур – обзор. Препринт arXiv arXiv: 1904.06554. doi: <https://doi.org/10.48550/arXiv.1904.06554>

Бендори Т., Ху Ю., Килил Дж., Микелин О. и Сингер А., 2023. Автокорреляционный анализ для крио-ЭМ с ограничениями разреженности: улучшенная сложность выборки и алгоритмы на основе проекций. Труды Национальной академии наук, 120(18), e2216507120. <https://doi.org/10.1073/pnas.2216507120>

Богословский А.В., Сухарев В.А., Жигулина И.В., Пантюхин М.А., 2021. Изображение представляет собой векторные поля, созданные преобразованием Фурье видеосигналов. Радиотехника, 85(7), 127–139. <https://www.elibrary.ru/item.asp?id=46452211>

Бронников Д.А., 2019. Построение автокорреляционных функций тепловых фотоизображений почвогрунтов. In *Информационные технологии и автоматизация управления* (Рр. 45–49). <https://www.elibrary.ru/item.asp?id=38248980>

Гонсалес Р. & Вудс Р., 2012. Цифровая обработка изображений. ISBN: 978-5-94836-331-8

Григорьев А.В., 2020. Выделение границ изображений в сложных сценах. *Фундаментально-прикладные проблемы безопасности, живучести, надёжности, устойчивости и эффективности систем* (Рр. 477–478). <https://www.elibrary.ru/item.asp?id=44323779>

Дабагов А.Р., Томакова Р.А., Алексеев В.А. & Кондрашов Д.С., 2020. Метод автоматической классификации рентгеновских изображений на основе масок прозрачности. *Высокопроизводительные вычислительные системы и технологии*, 4(1), 110-116. <https://elibrary.ru/item.asp?id=43961251>

Котелина Н.О. & Матвийчук Б.Р., 2019. Кластеризация изображения методом k-средних. *Вестник Сыктывкарского университета. Серия 1. Математика. Механика. Информатика*, (32). 101–112. <https://cyberleninka.ru/article/n/klasterizatsiya-izobrazheniya-metodom-k-srednih>

Курилович А.В., Цветков В.Ю. & Конопелько В.К., 2019. Структурный алгоритм текстурной сегментации полутоновых изображений. <https://libeldoc.bsuir.by/handle/123456789/38001>

Леонтьева С.А. және Демин А.Ю., 2019. k-means әдісі арқылы суреттерді кластерлеу. Жастар және қазіргі ақпараттық технологияларда: Студенттер, аспиранттар және жас ғалымдардың XVI халықаралық ғылыми-практикалық конференциясының материалдары, 3–7 желтоқсан 2018 ж., Томск — Томск, 2019. (86–87 беттер). doi: <https://doi.org/10.21822/2073-6185-2023-50-1-114-122>

Мехри М., Гомес-Крамер П., Хероу П. және Мулот Р., 2013. Автокорреляция функциясын және көп ажыратымдылықты талдау арқылы ескі құжат кескінін сегменттеу. Құжатты тану және іздеуде XX (8658-том, № 18, 8658-18-беттер). SPIE, ақпан, doi: <https://doi.org/10.1117/12.2002365>

Нехороших А.В. & Брянцева А.А., 2020. Кластеризация изображений методом k-средних. *Актуальные проблемы прикладной математики, информатики и механики* (Рр. 319–322). https://elibrary.ru/download/elibrary_42493252_38646523.pdf

Харалик Р.М., 1979. Статистический и структурный подходы к текстуре. Труды IEEE, 67(5), 786–804. doi: <https://doi.org/10.1109/PROC.1979.11328>.

Чехина Е.А., 2020. Обзор методов текстурного анализа изображений. *Евразийское Научное Объединение*, (6-2), 160–162. <https://www.elibrary.ru/item.asp?id=43161309>

Шекербек А., Абдикеримова Г., Сабыр А. & Абулхаир Ж., 2022. Применение метода и алгоритма для выявления патологии грудной клетки. *Известия НАН РК. Серия физико-математическая*, (4), 159–167. doi: <https://doi.org/10.32014/2022.2518-1726.164>

REFERENCES

Abdikirimova G., Lamasheva J., Baibulova M. & Tokkulieva A., 2023. Classification of x-ray images using the deep learning algorithm. Proceedings of the National Academy of Sciences of the Republic of Kazakhstan. Physico-mathematical series, (1). Pp. 204–227. <https://journals.nauka-nanrk.kz/physics-mathematics/article/view/4876>

Armi L. & Fekri-Ershad S., 2019. Texture image analysis and texture classification methods-A review. *arXiv preprint arXiv:1904.06554*. doi: <https://doi.org/10.48550/arXiv.1904.06554>

Bendory T., Khoo Y., Kileel J., Mickelin O. & Singer A., 2023. Autocorrelation analysis for cryo-EM with sparsity constraints: Improved sample complexity and projection-based algorithms. *Proceedings of the National Academy of Sciences*, 120(18), e2216507120. <https://doi.org/10.1073/pnas.2216507120>

Bogoslovsky A.V., Sukharev V.A., Zhigulina I.V. & Pantyukhin M.A., 2021. Vector fields generated by the Fourier transform of image video signals. *Radio Engineering*, 85(7). Pp. 127–139. <https://www.elibrary.ru/item.asp?id=46452211>

Bronnikov D.A., 2019. Construction of autocorrelation functions of thermal photographic images of soils. In *Information technology and control automation* (Pp. 45–49). <https://www.elibrary.ru/item.asp?id=38248980>

Chekhina E.A., 2020. Overview of methods of texture analysis of images. *Eurasian Scientific Association*, (6-2). Pp. 160–162. <https://www.elibrary.ru/item.asp?id=43161309>

Gonzalez R. & Woods R., 2012. *Digital image processing*. ISBN: 978-5-94836-331-8

Grigoriev A.V., 2020. Selecting the borders of images in complex scenes. In *Fundamental and applied problems of safety, survivability, reliability, stability and efficiency of systems* (Pp. 477–478). <https://www.elibrary.ru/item.asp?id=44323779>

Dabagov A.R., Tomakova R.A., Alekseev V.A. & Kondrashov D.S., 2020. Method of automatic classification of X-ray images based on mask transparency. *High-performance computing systems and technologies*, 4(1). Pp. 110–116. <https://elibrary.ru/item.asp?id=43961251>

Haralick R.M., 1979. Statistical and structural approaches to texture. *Proceedings of the IEEE*, 67(5). Pp. 786–804, doi: <https://doi.org/10.1109/PROC.1979.11328>.

Kotelina N.O. & Matviychuk B.R., 2019. Image clustering by k-means method. *Bulletin of the Syktyvkar University. Series 1. Mathematics. Mechanics. Computer Science*, (32). Pp. 101–112. <https://cyberleninka.ru/article/n/klasterizatsiya-izobrazheniya-metodom-k-srednih>

Kurilovich A.V., Tsvetkov V.Yu. & Konopelko V.K., 2019. Structural algorithm for texture segmentation of grayscale images. <https://libeldoc.bsuir.by/handle/123456789/38001>

Leontieva S.A. & Demin A.Yu., 2019. Clustering images using the k-means method. In *Youth and Modern Information Technologies: Proceedings of the XVI International Scientific and Practical Conference of Students, Postgraduates and Young Scientists, December 3–7, 2018, Tomsk—Tomsk*, 2019. (Pp. 86–87). doi: <https://doi.org/10.21822/2073-6185-2023-50-1-114-122>

Mehri M., Gomez-Krämer P., Héroux P. and Mullot R., 2013. Old document image segmentation

using the autocorrelation function and multiresolution analysis. In *Document Recognition and Retrieval XX* (Vol. 8658, № 18. Pp. 8658–18). SPIE, February, doi: <https://doi.org/10.1117/12.2002365>

Nehoroshikh A.V. & Bryantseva A.A., 2020. Clustering images using the k-means method. In *Actual problems of applied mathematics, informatics and mechanics* (Pp. 319–322). https://elibrary.ru/download/elibrary_42493252_38646523.pdf

Shekerbek A., Abdikerimova G., Sabyr A. & Abulhair Zh., 2022. Using the method and algorithm for determining chest cell pathology. *Izvestia NAS RK. Series physical and mathematical*, (4). Pp. 159–167. <https://doi.org/10.32014/2022.2518-1726.164>.

МАЗМҰНЫ

А. Адамова, Т. Жукабаева, Е. Марденов ЗАТТАР ИНТЕРНЕТІ: ЖЕҢІЛДІК АЛГОРИТМДЕРДІҢ ДАМУЫ ЖӘНЕ БОЛАШАҒЫ.....	5
Г. Алпысбай, А. Бедельбаев, О. Усагова, А. Жұмабекова, Эдзард Хофиг ЗИЯНДЫ БАҒДАРЛАМАЛЫҚ ЖАБДЫҚТАРДЫ ТАЛДАУДА МАШИНАЛЫҚ ОҚЫТУ АЛГОРИТМІН ҚОЛДАНУ.....	21
А.У. Алтаева, А.Ш. Каипова, А.У. Мухамеджанова, Г.К. Оспанова МЕДИЦИНАДА ЧАТ-БОТТАРДЫ ҚОЛДАНУ ПЕРСПЕКТИВАЛАРЫ.....	32
Г.А. Анарбекова, Н.Н. Оспанова, Д.Ж. Анарбеков НОРМАЛАНҒАН КІРІС ВЕКТОРЛАРЫ: ДЕРЕКТЕРДІ ДАЙЫНДАУДЫҢ БАСТАПҚЫ КЕЗЕҢІ.....	40
А.Е. Әбжанова, А.И. Такуадина, С.К. Сагнаева, С.К. Серикбаева, Г.Т. Азиева ТОПЫРАҚТЫ ТЕХНИКАЛЫҚ МЕЛИОРАЦИЯЛАУ ӘДІСТЕРІНДЕ АҚПАРАТТЫҚ ЖҮЙЕЛЕРДІ ПАЙДАЛАНУ.....	55
К.Н. Әлібекова, Ж.М. Алимжанова, С.С. Байзакова СЫМСЫЗ СЕНСОРЛЫҚ ЖЕЛІЛЕР ҮШІН БЛОКТЫҚ ШИФРЛАРДЫҢ ӨНІМДІЛІГІН БАҒАЛАУ.....	70
К.Б. Багитова, Ш.Ж. Мүсірәлиева, М.А. Болатбек, Р.Қ. Оспанов ИНТЕРНЕТТЕ ЭКСТРЕМИСТІК МАЗМҰНДЫ АНЫҚТАУҒА АРНАЛҒАН EXWEB БАҒДАРЛАМАЛЫҚ ЖАБДЫҚТАМАСЫН ӨЗІРЛЕУ.....	81
А.Ш. Баракова, О.А. Усагова, А.С. Орынбаева ВЕБ САЙТТАРДАҒЫ САНДЫҚ РЕСУРСТАРДЫ СТЕГАНОГРАФИЯ ӘДІСІМЕН ҚОРҒАУДЫҢ МОДЕЛІ.....	96
А.С. Омарбекова, А.Е. Назырова, Н. Тасболатұлы, Б.Ш. Разахова ИНТЕЛЛЕКТУАЛДЫ ELEARNING ЖҮЙЕСІНІҢ ОНТОЛОГИЯЛЫҚ МОДЕЛІ ЖӘНЕ ОҚЫТУ НӘТИЖЕЛЕРІ.....	108
М.Қ. Болсынбек, Г.Б. Абдикеримова, С.К. Серикбаева, А.Ж. Танирбергенов, Ж.К. Тасжурекова ТОПЫРАҚ ЖӘНЕ ТОПЫРАҚ ЭРОЗИСЫН БОЛЖАУЖЫҢ АҚПАРАТТЫҚ ЖҮЙЕЛЕРІ МЕН ӘДІСТЕРІН ЗЕРТТЕУ.....	128
Л.З. Жолшиева, Т.К. Жукабаева, Ш. Тураев, М.А. Бердиева, Б.А. Ху Вен-Цен LSTM ЖӘНЕ GRU ҮЛГІЛЕРІ НЕГІЗІНДЕ ҚАЗАҚ ДАКТИЛЬДЕРІН ТАНУДЫҢ ИНТЕЛЛЕКТУАЛДЫ ЖҮЙЕСІН ҚҰРУ.....	141
М.Д. Кабибуллин, Б.Б. Оразбаев, К.Н. Оразбаева, С.Ш. Исакова, Ж.Ш. Аманбаева КҮРДЕЛІ ХИМИЯЛЫҚ-ТЕХНОЛОГИЯЛЫҚ ЖҮЙЕЛЕР АГРЕГАТТАРЫНЫҢ МОДЕЛЬДЕРІН БАСТАПҚЫ АҚПАРАТТЫҢ ЖЕТІСПЕУШІЛІГІ МЕН АЙҚЫНСЫЗДЫҒЫ ЖАҒДАЙЫНДА ҚҰРУ.....	154

М.Ж. Қалдарова, А.С. Аканова, М.Г. Гриф, У.Ж. Айтимова, А.С. Муканова ТОПЫРАҚ ЖАҒДАЙЫН БАҒАЛАУ ҮШІН ҚОЛДАНЫЛАТЫН ҒАРЫШТЫҚ СУРЕТТЕРДІ ӨНДЕУ АЛГОРИТМДЕРІ МЕН ӘДІСТЕРІ.....	172
К. Келесбаев, Ш. Раманкулов, М. Нуризинова, А. Паттаев, Н. Мұсахан STEM ЖОБАЛЫҚ ОҚЫТУДЫҢ БОЛАШАҚ ФИЗИКА МАМАНДАРЫН ДАЯРЛАУДАҒЫ ЕРЕКШЕЛІКТЕРІ.....	193
А.Е. Кулакаева, Е.А. Дайнеко, А.З. Айтмагамбетов, А.Т. Жетписбаева, Б.А. Кожаметова ШАҒЫН ҒАРЫШ АППАРАТЫ ОРБИТАСЫНЫҢ СИПАТТАМАЛАРЫНЫҢ СПУТНИКТІК РАДИО МОНИТОРИНГ ЖҮЙЕСІНІҢ ПАРАМЕТРЛЕРІНЕ ӘСЕРІ ТУРАЛЫ.....	208
А.Е. Назырова, Г.Т. Бекманова, А.С. Муканова, Н. Амангелді, М.Ж. Қалдарова БІЛІМ БЕРУ БАҒДАРЛАМАЛАРЫ ҮШІН АВТОМАТТАНДЫРЫЛҒАН ЖҮЙЕНІ ӨЗІРЛЕУ.....	221
А.Б. Тоқгарова, Б.С. Омаров, Ж.Ж. Ажибекова, Г.И. Бейсенова, Р.Б. Абдрахманов ОНЛАЙН КОНТЕНТТЕГІ БЕЙӘДЕП СӨЗДЕР МӘЛІМЕТТЕР ҚОРЫН DATA MINING АРҚЫЛЫ АНАЛИЗДЕУ.....	237
Ә.Б. Тынымбаев, К.С. Байшоланова, К.Е. Кубаев АҚПАРАТТЫ ҚОРҒАУ ЖҮЙЕЛЕРІНДЕГІ NAVIVE BAYESIAN ЖІКІТІУШСІН ҚОЛДАНУ.....	252
Г.Қ. Шаметова, А.Ә. Шәріпбай, Б.Ф. Сайлау ҚОЛЖЕТІМДІЛІКТІ БАСҚАРУ ЖҮЙЕЛЕРІНДЕГІ ҚҰПИЯНЫ БӨЛҮДІҢ КРИПТОГРАФИЯЛЫҚ СҰЛБАЛАРЫН ТАЛДАУ.....	261
Г.Б. Абдикеримова, А.Ә. Шекербек, М.Г. Байбулова, С.К. Абдикаримова, Ш.Ш. Жолдасова КЕУДЕ ПАТОЛОГИЯСЫН АВТОКОРРЕЛЯЦИЯЛЫҚ ФУНКЦИЯ АРҚЫЛЫ АНЫҚТАУ.....	274

СОДЕРЖАНИЕ

А. Адамова, Т. Жукабаева, Е. Марденов ИНТЕРНЕТ ВЕЩЕЙ: СОСТОЯНИЕ И ПЕРСПЕКТИВЫ РАЗВИТИЯ ЛЕГКОВЕСНЫХ АЛГОРИТМОВ.....	5
Г. Алпысбай, А. Бедельбаев, О. Усагова, А. Жумабекова, Эдзарт Хофиг ПРИМЕНЕНИЕ АЛГОРИТМА МАШИННОГО ОБУЧЕНИЯ ДЛЯ АНАЛИЗА ВРЕДНОСНОГО ПО.....	21
А.У. Алтаева, А.Ш. Каипова, А.У. Мухамеджанова, Г.К. Оспанова ПЕРСПЕКТИВЫ ИСПОЛЬЗОВАНИЯ ЧАТ-БОТОВ В МЕДИЦИНЕ.....	32
Г.А. Анарбекова, Н.Н. Оспанова*, Д.Ж. Анарбеков НОРМАЛИЗОВАННЫЕ ВХОДНЫЕ ВЕКТОРЫ: ПЕРВИЧНЫЙ ЭТАП ПОДГОТОВКИ ДАННЫХ.....	40
А.Е. Абжанова, А.И. Такуадина, С.К. Сагнаева, С.К. Серикбаева, Г.Т. Азиева ИСПОЛЬЗОВАНИЕ ИНФОРМАЦИОННЫХ СИСТЕМ В МЕТОДАХ ТЕХНИЧЕСКИХ МЕЛИОРАЦИЙ ГРУНТОВ.....	55
К.Н. Алибекова, Ж.М. Алимжанова, С.С. Байзакова ОЦЕНКА ПРОИЗВОДИТЕЛЬНОСТИ БЛОЧНЫХ ШИФРОВ ДЛЯ БЕСПРОВОДНЫХ СЕНСОРНЫХ СЕТЕЙ.....	70
К.Б. Багитова, Ш.Ж. Мусиралиева, М.А. Болатбек, Р.К. Оспанов РАЗРАБОТКА ПРОГРАММНОГО ОБЕСПЕЧЕНИЯ EXWEB ДЛЯ ВЫЯВЛЕНИЯ ЭКСТРЕМИСТСКОГО КОНТЕНТА В СЕТИ ИНТЕРНЕТ.....	81
А.Ш. Баракова, О.А. Усагова, А.С. Орынбаева РАЗРАБОТКА МОДЕЛИ ЗАЩИТЫ ЦИФРОВЫХ WEB РЕСУРСОВ С ИСПОЛЬЗОВАНИЕМ МЕТОДОВ СТЕГАНОГРАФИИ.....	96
А.С. Омарбекова, А.Е. Назырова, Н. Тасболатұлы, Б.Ш. Разахова ОНТОЛОГИЧЕСКАЯ МОДЕЛЬ ИНТЕЛЛЕКТУАЛЬНОЙ СИСТЕМЫ ЭЛЕКТРОННОГО ОБУЧЕНИЯ И РЕЗУЛЬТАТЫ ОБУЧЕНИЯ.....	108
М.Қ. Болсынбек, Г.Б. Абдикеримова, С.К. Серикбаева, А.Ж. Танирбергенов, Ж.К. Тасжурекова ИССЛЕДОВАНИЕ ИНФОРМАЦИОННЫХ СИСТЕМ И МЕТОДОВ ПРОГНОЗИРОВАНИЯ ПОЧВЕННОЙ И ПОЧВЕННОЙ ЭРОЗИИ.....	128
Л.З. Жолшиева, Т.К. Жукабаева, Ш. Тураев, М.А. Бердиева, Б.А. Ху Вен-Цен РАЗРАБОТКА ИНТЕЛЛЕКТУАЛЬНОЙ СИСТЕМЫ РАСПОЗНАВАНИЯ КАЗАХСКИХ ДАКТИЛЬНЫХ ЖЕСТОВ НА ОСНОВЕ МОДЕЛЕЙ LSTM И GRU.....	141
М.Д. Кабибуллин, Б.Б. Оразбаев, К.Н. Оразбаева, С.Ш. Исакова, Ж.Ш. Аманбаева РАЗРАБОТКА МОДЕЛЕЙ АГРЕГАТОВ СЛОЖНЫХ ХИМИКО-ТЕХНОЛОГИЧЕСКИХ СИСТЕМ В УСЛОВИЯХ ДЕФИЦИТА И НЕЧЕТКОСТИ ИСХОДНОЙ ИНФОРМАЦИИ.....	154

М.Ж. Калдарова, А.С. Аканова, М.Г. Гриф, У.Ж. Айтимова, А.С. Муканова АЛГОРИТМЫ И МЕТОДЫ ОБРАБОТКИ КОСМИЧЕСКИХ СНИМКОВ ДЛЯ ОЦЕНКИ СОСТОЯНИЯ ПОЧВ.....	172
К. Келесбаев, Ш. Раманкулов, М. Нуризинова, А. Паттаев, Н. Мұсахан ОСОБЕННОСТИ ПРОЕКТНОГО ОБУЧЕНИЯ STEM В ПОДГОТОВКЕ БУДУЩИХ СПЕЦИАЛИСТОВ ПО ФИЗИКЕ.....	193
А.Е. Кулакаева, Е.А. Дайнеко, А.З. Айтмагамбетов, А.Т. Жетписбаева, Б.А. Кожаметова О ВЛИЯНИИ ХАРАКТЕРИСТИК ОРБИТЫ МАЛОГО КОСМИЧЕСКОГО АППАРАТА НА ПАРАМЕТРЫ СИСТЕМЫ СПУТНИКОВОГО РАДИОМОНИТОРИНГА.....	208
А.Е. Назырова, Г.Т. Бекманова, А.С. Муканова, Н. Амангелді, М.Ж. Калдарова, РАЗРАБОТКА АВТОМАТИЗИРОВАННОЙ СИСТЕМЫ ДЛЯ ОБРАЗОВАТЕЛЬНЫХ ПРОГРАММ.....	221
А.Б. Токгарова, Б.С. Омаров, Ж.Ж. Ажибекова, Г.И. Бейсенова, Р.Б. Абдрахманов АНАЛИЗ НЕОБРАЗНЫХ СЛОВ В ОНЛАЙН-КОНТЕНТЕ С ПОМОЩЬЮ DATA MINING.....	237
Ә.Б. Тынымбаев, К.С. Байшоланова, К.Е. Кубаев ПРИМЕНЕНИЕ НАИВНОГО БАЙЕСОВСКОГО КЛАССИФИКАТОРА В СИСТЕМАХ ЗАЩИТЫ ИНФОРМАЦИИ.....	252
Г.Қ. Шаметова, А.Ә. Шәріпбай, Б.Ғ. Сайлау АНАЛИЗ КРИПТОГРАФИЧЕСКИХ СХЕМ РАСПРЕДЕЛЕНИЯ СЕКРЕТОВ В СИСТЕМАХ УПРАВЛЕНИЯ ДОСТУПОМ.....	261
Г.Б. Абдикеримова, А.А. Шекербек, М.Г. Байбулова, С.К. Абдикаримова, Ш.Ш. Жолдасова ОПРЕДЕЛЕНИЕ ГРУДНОЙ ПАТОЛОГИИ С ПОМОЩЬЮ ФУНКЦИИ АВТОКОРРЕЛЯЦИИ.....	274

CONTENTS

A. Adamova, T. Zhukabayeva, Y. Mardenov INTERNET OF THINGS: STATUS AND PROSPECTS FOR THE DEVELOPMENT OF LIGHTWEIGHT ALGORITHMS.....	5
G. Alpysbay, A. Bedelbayev, O. Ussatova, A. Zhumabekova, Edzard Höfig APPLICATION OF MACHINE LEARNING ALGORITHM IN THE ANALYSIS OF MALICIOUS SOFTWARE.....	21
A.U. Altaeva, A.S. Kaipova, A.U. Mukhamejanova, G.K. Ospanova PROSPECTS OF USING CHATBOTS IN MEDICINE.....	32
G.A. Anarbekova, N.N. Ospanova, D.Zh. Anarbekov NORMALIZED INPUT VECTORS: THE PRIMARY STAGE OF DATA PREPARATION.....	40
A.E. Abzhanova, A.I. Takuadina, S.K. Sagnaeva, S.K. Serikbayeva, G.T. Azieva THE USE OF INFORMATION SYSTEMS IN THE METHODS OF TECHNICAL SOIL RECLAMATION.....	55
K. Alibekova, Zh. Alimzhanova, S.S. Baizakova RATING VALUATION OF BLOCK CIPHERS FOR WIRELESS SENSOR NETWORKS.....	70
K.B. Bagitova, Sh.Zh. Mussiraliyeva, M.A. Bolatbek, R.K. Ospanov DEVELOPMENT OF EXWEB SOFTWARE FOR DETECTING EXTREMIST CONTENT ON THE INTERNET.....	81
A.Sh. Barakova, O.A. Usatova, A.S. Orynbaeva DIGITAL RESOURCES ON WEBSITES MODEL OF PROTECTION BY STEGANOGRAPHY.....	96
A.S. Omarbekova, A.E. Nazyrova, N. Tasbolatuly, B.Sh. Razakhova ONTOLOGICAL MODEL OF AN INTELLIGENT E-LEARNING SYSTEM AND LEARNING OUTCOMES.....	108
M. Bolsynbek, G. Abdikerimova, S. Serikbayeva, A. Tanirbergenov, Zh. Taszhurekova RESEARCH OF INFORMATION SYSTEMS AND METHODS OF FORECASTING SOIL AND SOIL EROSION.....	128
L. Zholshiyeva, T. Zhukabayeva, Sh. Turaev, M. Berdieva, B. Khu Ven-Tsen DEVELOPMENT OF AN INTELLECTUAL SYSTEM FOR RECOGNIZING KAZAKH DACTYL GESTURES BASED ON LSTM AND GRU MODELS.....	141
M. Kabibullin, B. Orazbayev, K. Orazbayeva, S. Iskakova, Zh. Amanbayeva DEVELOPMENT OF MODELS OF UNITS OF COMPLEX CHEMICAL-TECHNOLOGICAL SYSTEMS UNDER CONDITIONS OF DEFICIENCY AND FUZZY OF INITIAL INFORMATION.....	154
M.Zh. Kaldarova, A.S. Akanova, M.G. Grif, U.Zh. Aitimova, A.S. Mukanova ALGORITHM AND METHOD OF PROCESSING SPACE PHOTOS FOR ASSESSMENT OF SOIL.....	172

K. Kelesbaev, Sh. Ramankulov, M. Nurizinova, A. Pattaev, N. Mussakhan FEATURES OF STEAM PROJECT TRAINING IN THE PREPARATION OF FUTURE SPECIALISTS IN PHYSICS.....	193
A.E. Kulakayeva, Y.A. Daineko, A.Z. Aitmagambetov, A.T. Zhetpisbaeva, B.A. Kozhakhmetova ABOUT THE INFLUENCE OF THE ORBIT CHARACTERISTICS OF A SMALL SPACECRAFT ON THE PARAMETERS OF THE SATELLITE RADIO MONITORING SYSTEM.....	208
A.E. Nazyrova, G.T. Bekmanova, A.S. Mukanova, N. Amangeldi, M.Zh. Kaldarova DEVELOPMENT OF AN AUTOMATED SYSTEM FOR EDUCATIONAL PROGRAMS.....	221
A.B. Toktarova, B.S. Omarov, Zh.Zh. Azhibekova, G.I. Beissenova, R.B. Abdrakhmanov ANALYSIS OF HATE SPEECH WORDS IN ONLINE CONTENT BY USING DATA MINING.....	237
A.B. Tynymbayev, K.S. Baisholanova, K.Ye. Kubaev APPLICATION OF NAVIVE BAYESIAN CLASSIFIER IN INFORMATION PROTECTION SYSTEMS.....	252
G.K. Shametova, A.A. Sharipbay, B.G. Sailau ANALYSIS OF CRYPTOGRAPHIC SECRET DISTRIBUTION SCHEMES IN ACCESS CONTROL SYSTEMS.....	261
G.B. Abdikerimova, A.A. Shekerbek, M.G. Baibulova, S.K. Abdikarimova, Sh.Sh. Zholdassova CHEST PATHOLOGY DETERMINATION THROUGH AUTOCORRELATION FUNCTION.....	274

**Publication Ethics and Publication Malpractice
the journals of the National Academy of Sciences of the Republic of Kazakhstan**

For information on Ethics in publishing and Ethical guidelines for journal publication see <http://www.elsevier.com/publishingethics> and <http://www.elsevier.com/journal-authors/ethics>.

Submission of an article to the National Academy of Sciences of the Republic of Kazakhstan implies that the described work has not been published previously (except in the form of an abstract or as part of a published lecture or academic thesis or as an electronic preprint, see <http://www.elsevier.com/postingpolicy>), that it is not under consideration for publication elsewhere, that its publication is approved by all authors and tacitly or explicitly by the responsible authorities where the work was carried out, and that, if accepted, it will not be published elsewhere in the same form, in English or in any other language, including electronically without the written consent of the copyright-holder. In particular, translations into English of papers already published in another language are not accepted.

No other forms of scientific misconduct are allowed, such as plagiarism, falsification, fraudulent data, incorrect interpretation of other works, incorrect citations, etc. The National Academy of Sciences of the Republic of Kazakhstan follows the Code of Conduct of the Committee on Publication Ethics (COPE), and follows the COPE Flowcharts for Resolving Cases of Suspected Misconduct (http://publicationethics.org/files/u2/New_Code.pdf). To verify originality, your article may be checked by the Cross Check originality detection service <http://www.elsevier.com/editors/plagdetect>.

The authors are obliged to participate in peer review process and be ready to provide corrections, clarifications, retractions and apologies when needed. All authors of a paper should have significantly contributed to the research.

The reviewers should provide objective judgments and should point out relevant published works which are not yet cited. Reviewed articles should be treated confidentially. The reviewers will be chosen in such a way that there is no conflict of interests with respect to the research, the authors and/or the research funders.

The editors have complete responsibility and authority to reject or accept a paper, and they will only accept a paper when reasonably certain. They will preserve anonymity of reviewers and promote publication of corrections, clarifications, retractions and apologies when needed. The acceptance of a paper automatically implies the copyright transfer to the National Academy of Sciences of the Republic of Kazakhstan.

The Editorial Board of the National Academy of Sciences of the Republic of Kazakhstan will monitor and safeguard publishing ethics.

Правила оформления статьи для публикации в журнале смотреть на сайтах:

www.nauka-nanrk.kz

<http://physics-mathematics.kz/index.php/en/archive>

ISSN 2518-1726 (Online),

ISSN 1991-346X (Print)

Заместитель директор отдела издания научных журналов НАН РК *Р. Жалиқызы*

Редакторы: *М.С. Ахметова, Д.С. Аленов*

Верстка на компьютере *Г.Д. Жадыранова*

Подписано в печать 12.06.2023.

Формат 60x881/8. Бумага офсетная. Печать – ризограф.

19,0 п.л. Тираж 300. Заказ 2.